

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 733 279**

51 Int. Cl.:

**C07K 14/14** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.07.2014 PCT/EP2014/064652**

87 Fecha y número de publicación internacional: **15.01.2015 WO15004158**

96 Fecha de presentación y número de la solicitud europea: **08.07.2014 E 14736835 (1)**

97 Fecha y número de publicación de la concesión europea: **03.04.2019 EP 3019518**

54 Título: **Partículas de rotavirus con proteínas de superficie quiméricas**

30 Prioridad:

**08.07.2013 EP 13175637**

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

**28.11.2019**

73 Titular/es:

**NOVARTIS AG (33.3%)  
Lichtstrasse 35  
4056 Basel, CH;  
CHILDREN'S MEDICAL CENTER CORPORATION  
(33.3%) y  
BRANDEIS UNIVERSITY (33.3%)**

72 Inventor/es:

**DORMITZER, PHILIP R.;  
GRIGORIEFF, NIKOLAUS;  
HARRISON, STEPHEN;  
PAN, JUNHUA y  
SETTEMBRE, ETHAN**

74 Agente/Representante:

**IZQUIERDO BLANCO, María Alicia**

**ES 2 733 279 T3**

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

**DESCRIPCIÓN**

Partículas de rotavirus con proteínas de superficie quiméricas

**5 REFERENCIA CRUZADA A APLICACIONES RELACIONADAS**

Esta solicitud reivindica el beneficio de la Solicitud de Patente Europea No. 13175637.1, presentada el 8 de julio de 2013.

**10 LISTADO DE SECUENCIAS**

Se adjunta una copia electrónica de una lista de secuencias que se incorpora aquí como referencia y forma parte de la solicitud tal como se presenta.

**15 APOYO DEL GOBIERNO**

Esta invención se realizó con el apoyo del gobierno de los Estados Unidos. Con los números de subvención P01-GM062580 y AI-89618 otorgados por los Institutos Nacionales de la Salud. El gobierno de los Estados Unidos tiene ciertos derechos en la invención.

**20 CAMPO DE LA INVENCION**

La presente invención se refiere al uso de partículas de rotavirus para exhibir una proteína heteróloga, sola o en complejo con otra molécula. La invención se refiere adicionalmente a métodos que emplean estas partículas de rotavirus modificadas para determinar rápidamente la estructura de la proteína heteróloga o el complejo usando microscopía de crioelectrones (crio-EM). La invención también se refiere a un método para inmunizar un paciente, en donde el método comprende administrar al paciente las partículas de rotavirus modificadas de la invención.

**30 ANTECEDENTES**

La familia de *Reoviridae* comprende un grupo estructuralmente relacionado de virus. Los miembros de esta familia de virus pueden provocar infecciones del sistema gastrointestinal y los tractos respiratorios en los mamíferos y aves. Algunos virus de esta familia también pueden infectar plantas.

Los Reoviridae son virus de ARN de doble hebra, no envueltos constituidos de una cápside icosaédrica que se compone típicamente de una capa exterior de proteína y una o más capas interiores de proteína. El género Rotavirus forma parte de la familia Reoviridae. Los rotavirus forman partículas de virus de tres capas. La capa exterior de una partícula de rotavirus consiste de una proteína de envoltura VP7, y una proteína pico, VP4; la capa intermedia consiste de VP6; y la capa interior se forma por VP2. La partícula de virus de tres capas es infecciosa, después de la activación por tripsina. Durante la entrada de células de rotavirus, la capa exterior de las partículas del virus se remueve por un mecanismo llamado "eliminación del recubrimiento", y la partícula de doble capa entra al citoplasma. Durante la entrada de las células, el rotavirus penetra una membrana celular, posiblemente una membrana endosómica. La partícula de rotavirus de doble capa está transcripcionalmente activa. En el citoplasma, transcribe el genoma, extruyendo el ARN a través de los poros en su superficie. Los nuevos transcritos se usan para fabricar proteínas de rotavirus y segmentos de genoma de ARN de doble hebra nuevos, que se empaquetan en las partículas reovíricas recientemente formadas.

Para algunos Reoviridae, el proceso de eliminación del recubrimiento se puede simular in vitro usando partículas reovíricas purificadas. En el caso de los rotavirus, la quelación de calcio o choque térmico provoca la eliminación del recubrimiento de las partículas reovíricas in vitro. Una vez que la capa exterior de proteínas se ha removido, las partículas de doble capa de rotavirus resultante se pueden "recubrir" con versiones recombinantemente expresadas de las proteínas que forman la capa exterior. Al recubrir con VP4 y VP7 recombinantes, se pueden formar partículas de rotavirus recubiertas infecciosas que se asemejan estrechamente a viriones maduros. Al recubrir con VP7 se pueden formar partículas sin picos, de tres capas, solas, que son mínimamente infecciosas. Al modificar genéticamente las regiones de codificación de las proteínas individuales de la capa exterior y al expresarlas recombinantemente, las propiedades de las proteínas que constituyen la capa exterior se pueden estudiar en experimentos de recubrimiento sin requerir la producción de virus completamente recombinantes. Para describir estos tipos de experimentos, se ha acuñado el término "genética de recubrimiento".

En el caso de los reovirus, el estado metaestable de la partícula subvívica infecciosa ha permitido el uso de microscopía de crioelectrones (crio-EM) en elucidar la base estructural para el mecanismo cebador que se usa por los reovirus para entrar a sus células hospedadoras. Similarmente, se han usado partículas de rotavirus recubiertas in vitro para estudiar las interacciones moleculares en el ensamblaje de rotavirus y eliminación del recubrimiento usando la crio-EM. La presente invención se basa en los avances tecnológicos que se llevaron a cabo por estos estudios de recubrimiento y crio-EM.

Peralta et al., 2009, Virology Journal 6:192 describes the use of rotavirus V2 and VP6 proteins to display epitopes on the surface of double layered VLPs.

## 5 BREVE DESCRIPCION DE LA INVENCION

La combinación de crio-EM de alta resolución en combinación con genética de recubrimiento ha hecho posible hacer avanzar el uso del modelo de rotavirus más allá del estudio de los rotavirus mismos. Los inventores han encontrado que las partículas de rotavirus se pueden usar para la rápida determinación de estructura de las proteínas triméricas heterólogas que se muestran sobre la superficie de las partículas de rotavirus modificadas. Las partículas de rotavirus modificadas para mostrar proteínas triméricas heterólogas sobre su superficie también son útiles en producir una amplia gama de vacunas novedosas.

La presente invención proporciona partículas de rotavirus que comprenden una proteína de superficie química compuesta de una proteína de superficie de rotavirus ligada a una proteína heteróloga. Un aspecto de la invención, la proteína de superficie de rotavirus se liga a la proteína heteróloga a través de una secuencia ligadora. En otro aspecto de la invención, la proteína de superficie de rotavirus se liga a la proteína heteróloga a través de un sistema adaptador. En un aspecto preferido de la invención, la proteína heteróloga se liga no covalentemente a la proteína de superficie de rotavirus por un sistema adaptador de dos partes, donde una parte del sistema adaptador se liga a la proteína de superficie de rotavirus y la otra parte del sistema adaptador se liga a la proteína heteróloga, mediante lo cual ambas partes del sistema adaptador forman un complejo estable uniéndose de esta manera no covalentemente la proteína heteróloga a la proteínas de superficie de rotavirus. La proteína de superficie química puede ser parte de la capa exterior de la partícula de rotavirus al recubrir in vitro las partículas de rotavirus de doble capa (DLPs) con ella.

La invención proporciona además un primer ácido nucleico que comprende un marco de lectura abierto que codifica una proteína de superficie de rotavirus modificada que comprende una proteína de superficie de rotavirus, un primer polipéptido adaptador y, opcionalmente, una secuencia ligadora. La invención también proporciona un segundo ácido nucleico que comprende un marco de lectura abierto que codifica una proteína de fusión que comprende una proteína heteróloga, un segundo polipéptido adaptador y, opcionalmente, una secuencia ligadora. El primer polipéptido adaptador y el segundo polipéptido adaptador son capaces de formar un complejo estable. Típicamente, el marco de lectura abierto en los ácidos nucleicos de la invención se liga operacionalmente a una secuencia promotora. En algunas modalidades, el polipéptido adaptador comprende una secuencia de repetición héptada. Una célula que contiene un ácido nucleico de la invención también se proporciona. Los ácidos nucleicos y las células que contienen los mismos se pueden usar para preparar las proteínas de superficie químicas de la invención.

La invención se refiere adicionalmente a un kit que comprende la primera y segunda secuencias de ácido nucleico de la invención. En algunas modalidades, un kit de acuerdo con la invención puede comprender un primer ácido nucleico que codifica una proteína de superficie de rotavirus modificada que comprende una proteína de superficie de rotavirus y un primer polipéptido adaptador, y un segundo ácido nucleico que comprende una secuencia de nucleótidos que codifica un segundo polipéptido adaptador y un sitio de clonación múltiple, mediante lo cual la inserción de una región de codificación para una proteína heteróloga en el sitio de clonación múltiple produce un marco de lectura abierto que codifica una proteína de fusión que comprende la proteína heteróloga y el segundo polipéptido adaptador, en donde el primer polipéptido adaptador y el segundo polipéptido adaptador son capaces de formar un complejo estable. Los kits de la invención pueden comprender además una partícula de rotavirus. La partícula de rotavirus se puede derivar de la misma especie de rotavirus de la cual la proteína de superficie de rotavirus se derivó o de una especie diferente. Por ejemplo, el rotavirus rhesus VP7 se puede usar para recubrir las DLPs preparadas de rotavirus de bovino y viceversa.

Las partículas de rotavirus que comprenden una proteína de superficie química se pueden preparar por varios métodos. Un método preferido de la invención implica hacer propagar una partícula de rotavirus nativa que comprende una capa exterior y una célula desarrollada en un medio de cultivo, purificar las partículas del medio de cultivo, remover la capa exterior de la partícula para obtener un DLP de rotavirus, y recubrir la DLP rotavirus con una o más proteínas de superficie químicas para formar una partícula de rotavirus que comprende la proteína o proteínas de superficie químicas. Una partícula de rotavirus nativa es una partícula de tres capas, en donde la tercera capa más exterior forma la cubierta exterior de la partícula de rotavirus.

En un aspecto particular, la invención se refiere una primera proteína de fusión que comprende una proteína de superficie de rotavirus que forma trímeros, una secuencia de repetición héptada, y opcionalmente una secuencia ligadora. La invención proporciona además una segunda proteína de fusión que comprende una proteína heteróloga que forma trímeros, una secuencia de repetición héptada y opcionalmente una secuencia ligadora. La primera proteína de fusión y la segunda proteína de fusión son capaces de formar un complejo estable a través de las secuencias de repetición héptadas comprendidas en cada una de ellas. En algunas modalidades, la invención proporciona una proteína de superficie química formada por la primera proteína de fusión y la segunda proteína de

fusión. La proteína de superficie quimérica se puede mostrar sobre la superficie de una partícula de rotavirus.

En un aspecto de la invención, las partículas de rotavirus que comprenden una proteína de superficie quimérica se usan para determinar la estructura de una proteína heteróloga que forma una porción de la proteína de superficie quimérica. Por ejemplo, la invención proporciona un método para obtener un modelo tridimensional de una proteína de superficie quimérica, en donde el método comprende los pasos de (i) recubrir un DLP de rotavirus con una proteína de superficie quimérica que comprende todo o parte de una proteína heteróloga para producir una suspensión de partículas de rotavirus que muestran la proteína de superficie quimérica, (ii) congelar la suspensión, (iii) formar en imagen las partículas de rotavirus usando crio-EM para obtener una pluralidad de micrografías, y (iv) analizar la pluralidad de micrografías para obtener un modelo tridimensional de la proteína de superficie quimérica. En algunas modalidades, el paso (i) se puede subdividir en dos pasos, particularmente (a) un paso de recubrimiento en el cual la DLP de rotavirus se recubre con una proteína de superficie de rotavirus que comprende un primer adaptador para formar partículas de rotavirus y (b) un paso de ligación en el cual la DLP de rotavirus se incuba en presencia de una proteína heteróloga que comprende un segundo adaptador mediante lo cual se forma un complejo entre el primer adaptador y el segundo adaptador que da por resultado una proteína de superficie quimérica que se muestra en las partículas de rotavirus. En algunos casos, el método se puede modificar para determinar la estructura de una proteína heteróloga ligada a una molécula que se liga específicamente a la proteína heteróloga. El método modificado comprende los pasos de (i) recubrir una DLP de rotavirus con una proteína de superficie quimérica que comprende todo o parte de una proteína heteróloga para producir una suspensión de partículas de rotavirus que muestran la proteína de superficie quimérica, (ii) agregar a la suspensión una molécula que se liga específicamente a la proteína heteróloga, en donde la molécula forma un complejo con la proteína de superficie quimérica, (iii) congelar la suspensión, (iv) formar en imagen las partículas de rotavirus usando crio-EM para obtener una pluralidad de micrografías, y (v) analizar la pluralidad de micrografías para obtener un modelo tridimensional de la proteína de superficie quimérica combinada con la molécula. La molécula puede ser una molécula proteínica. Por ejemplo, la molécula proteínica puede ser todo o parte de un receptor o anticuerpo que liga específicamente la proteína heteróloga. Alternativamente, la molécula proteínica puede ser un polipéptido o péptido que se liga específicamente a la proteína heteróloga. En otras modalidades, la molécula es una molécula no proteínica. Por ejemplo, la molécula no proteínica puede ser un ácido nucleico

En un aspecto adicional, las partículas de rotavirus preparadas de acuerdo con la invención se usan como un medicamento. En una modalidad, la invención se refiere a una composición inmunogénica que comprende una partícula de rotavirus que comprende la proteína de superficie quimérica de la invención. La invención se refiere además a un método para tratar a un paciente en necesidad del mismo que comprende administrar al paciente la partícula de rotavirus de la invención, por ejemplo en la forma de una composición inmunogénica de la invención.

El concepto subyacente de la invención se puede extender a otras proteínas heterólogas (triméricas o no triméricas) y otros virus (en reovirus particulares) o ensamblajes (por ejemplo partículas similares a virus, jaulas de ferritina, etc.) al adaptar los métodos descritos en la presente.

## DESCRIPCION DETALLADA DE LA INVENCION

### Partículas víricas

El Rotavirus es un género de virus de ARN de doble hebra en la familia Reoviridae que comprende cinco tipos conocidos (Rotavirus A-E). Los rotavirus son una causa principal de gastroenteritis de la niñez. El rotavirus es una partícula icosaédrica de tres capas, no envuelta. Una partícula de tres capas infecciosa (TLP) o virión se forma de una partícula de doble capa no infecciosa (DLP) al recubrir DLP con una proteína de envoltura VP7 y la proteína pico VP4. La VP7 es un trímero, y 260 de tales trímeros decoran el exterior de la partícula recubierta. La DLP, que se compone de capas de proteínas icosaédricas VP2 y VP6 concéntricas es de aproximadamente 700 Å en diámetro y encapsida 11 segmentos de genoma de ARN de doble hebra, la polimerasa vírica (VP1), y una enzima de terminación (VP3). Durante la infección de una célula, las proteínas VP4 y VP7, que forman la capa exterior de la partícula de rotavirus, se disocian de la DLP en un medio ambiente bajo en calcio (probablemente en el compartimiento endosómico) - un proceso referido como "eliminación del recubrimiento" - y suministran el DLP que contiene la maquinaria de transcripción de ARN viral en el citoplasma. En ese punto, la DLP se sintetiza, se termina y libera copias de las 11 especies de ARNm.

La disociación o "eliminación del recubrimiento" de VP4 y VP7 de una partícula de rotavirus se puede llevar a cabo in vitro al incubar partículas de rotavirus en presencia de un quelante de calcio tal como EDTA o EGTA o por choque térmico. Las DLPs resultantes se pueden recubrir in vitro con VP4 y VP7 recombinantemente expresadas para formar partículas de rotavirus completamente infecciosas. Las partículas recubiertas de esta forma son muy ordenadas y proporcionan imágenes de crio-EM de alta resolución y mapas de densidad. Usando las TLPs reconstituidas in vitro, se ha usado el crio-EM para estudiar las interacciones moleculares en el ensamblaje de rotavirus y la separación en una resolución comparable con aquella de cristalografía de rayos X (ver las referencias 1 y 2).

Los inventores muestran en la presente que una proteína trimérica heteróloga (tal como hemaglutinina de

influenza) se puede unir a la proteína VP7 trimérica, formando de esta manera una proteína de superficie quimérica que puede proyectarse desde la superficie de una DLP de rotavirus adecuadamente recubierta, por lo tanto haciendo posible determinar la estructura de la proteína heteróloga con los mismos métodos de crió-EM que han proporcionado previamente estructuras de resolución casi atómicas del rotavirus y sus subpartículas. El método hace posible por primera vez desarrollar un ensayo de alto rendimiento para la determinación de estructura de los complejos de antígeno-anticuerpo. Esto no fue previamente posible debido a las limitaciones de la cristalografía de rayos X, que ha sido el método de elección para determinar la estructura de los complejos de antígeno-anticuerpo.

En principio, cualquier partícula de virus icosaédrica no envuelta que comprende una capa interior y una capa exterior se puede usar para practicar los métodos de la invención. Por ejemplo, la muestra de una proteína de superficie quimérica que comprende todo o parte de la proteína heteróloga es casi imposible usando cualquiera virus icosaédrico por el cual un sistema de genética inversa para la producción de las partículas de virus se ha establecido. Un sistema de genética inversa basada en plásmido que consiste de diez constructos de ADNc de reovirus se ha establecido para los reovirus de mamífero (ver referencia 3). La simetría octamérica de las jaulas de ferritina también tiene ejes de simetría de 3 veces y es adecuado para la práctica de esta invención. Otras proteínas que se ensamblan en las estructuras regulares, particuladas con ejes de simetría de 3 veces pueden ser adecuadas para la práctica de esta invención.

Idealmente, la capa exterior de la partícula de virus de un virus icosaédrico no envuelto se puede remover o separar, por ejemplo por tratamiento por proteasa o bajo condiciones bajas en calcio, para producir partículas subvéricas que comprenden la capa(s) interior solamente y se pueden recubrir con proteínas de capa exterior recombinantemente producidas. El re-ensamblaje o recubrimiento in vitro de las partículas subvéricas para completar las partículas reovéricas por la adición de la proteína(s) de capa exterior recombinantemente expresada es particularmente ventajoso en casos donde la presencia de las proteínas de superficie quimérica interferiría con el ensamblaje apropiado del virus propagado en el cultivo celular usando un procedimiento de genética inversa. Además, solo los vectores de expresión para las proteínas de capa exterior necesitan ser construidas, removiendo la necesidad por un sistema de genética inversa basada en plásmidos eficiente para estar en su lugar. El virus nativo se puede propagar simplemente en células de cultivo de tejido, y las partículas reovéricas se pueden separar de la capa exterior y volverse a ensamblar in vitro usando proteínas de capa exterior recombinantemente expresadas. De esta manera, usando un método que depende del recubrimiento in vitro de las partículas subvéricas para completar las partículas reovéricas elimina la necesidad para transfectar grandes números de plásmidos y remueve el paso de propagación adicional necesario típicamente para producir grandes números de partículas de virus del sistema de genética inversa basado en plásmido, haciendo tal método más amable para aplicaciones de alto rendimiento.

De acuerdo con el conocimiento de los inventores, todos los virus icosaédricos no envueltos conocidos donde la capa exterior de la partícula de virus se puede separar y la partícula subvérica se puede recubrir con proteínas de capa exterior recombinantemente expresadas para formar partículas de virus infecciosas que pertenecen a la familia de Reoviridae. Esta familia se subdivide de dos subfamilias, Sedoreovirinae y Spinareovirinae, que comprenden seis y nueve géneros, respectivamente. La subfamilia Sedoreovirinae comprende los géneros Cardoreovirus, Mimoreovirus, Orbivirus, Phytoreovirus, Rotavirus, y Seadornavirus. La subfamilia Spinareovirinae comprende los géneros Aquareovirus, Coltivirus, Cypovirus, Dinovernavirus, Fijivirus, Idnoreovirus, Mycoreovirus, Orthoreovirus, y Oryzavirus.

De manera preferente, la partícula de virus usada en practicar la invención tiene tres o algunas proteínas de capa exterior. De manera más preferente, la capa exterior de la partícula de virus se puede formar por una proteína de capa exterior individual, que es la proteína de superficie exterior usada en la preparación de la proteína de superficie quimérica. Un bajo número de proteínas de capa exterior es ventajoso debido a que algunas proteínas necesitan ser expresadas recombinantemente para recubrir las partículas subvéricas después de la separación.

El rotavirus tiene dos proteínas de capa exterior, VP4 y VP7, pero solamente la VP7 es una proteína de envoltura que es requerida para formar la capa exterior de una partícula de rotavirus. En muchos casos, solo el recubrimiento de las DLPs de rotavirus con una proteína de capa exterior, VP7, es suficiente a fin de practicar la invención.

### **Complejos de proteínas quiméricas**

Como concepto fundamental de la invención, la proteína de superficie de rotavirus se liga no covalentemente a la proteína heteróloga a través de un sistema adaptador. En un aspecto preferido de la invención, la proteína heteróloga se liga no covalentemente a la proteína de superficie de rotavirus por un sistema adaptador de dos partes, donde una parte del sistema adaptador se liga a la proteína de superficie de rotavirus y la otra parte del sistema adaptador se liga a la proteína heteróloga, mediante lo cual ambas partes del sistema adaptador forman un complejo estable uniendo no covalentemente de esta manera la proteína heteróloga a la proteína de superficie de rotavirus. El sistema adaptador se compone típicamente de un primer polipéptido adaptador y un segundo polipéptido adaptador. El primer polipéptido adaptador se fusiona a la proteína de superficie de rotavirus, opcionalmente a través de la secuencia ligadora. El segundo polipéptido adaptador se fusiona a la proteína

heteróloga, opcionalmente a través de una secuencia ligadora. El primero y segundo polipéptidos adaptadores interactúan entre sí para formar un complejo estable uniendo por lo tanto no covalentemente la proteína de superficie de rotavirus a la proteína heteróloga, formando de esta manera la proteína de superficie quimérica.

## 5 Proteína de superficie vírica

La capa exterior de una partícula de virus adecuada para practicar la invención comprende típicamente de varias proteínas de superficie exterior diferentes. La proteína de superficie principal es particularmente adecuada para mostrar una proteína heteróloga, ya que cubre la mayoría de la superficie de la partícula de virus. Por ejemplo, la superficie exterior del rotavirus (excluyendo los picos) se forma por 780 copias de la proteína VP7, que forma homotrímeros. La elección de una partícula vírica con una proteína de superficie principal que forma homotrímeros es particularmente preferida para practicar la invención.

En una modalidad específica, la proteína de superficie vírica es una proteína de superficie de rotavirus. En otra modalidad específica, la proteína de superficie vírica es una glicoproteína.

Aunque no se reivindica como parte de la invención, en principio, otros virus con proteínas de la superficie viral pueden ser adecuados para los métodos descritos en el presente documento, incluidos los ortoreovirus. Por ejemplo, la cápside viral exterior de la aquareovirus se forma por 200 trímeros de una proteína designada VP5. En los ortoreovirus de mamífero, la capa exterior de una partícula reoviral infecciosa contiene 600 copias de la proteína de penetración de membrana trimérica  $\mu$ , que es llamada con la proteína chaperona  $\sigma$ 3 (también presente en 600 copias) formando de esta manera un heterohexámero.

## 25 Proteína Heteróloga

Muchas proteínas heterólogas se pueden mostrar usando una proteína de superficie de rotavirus trimérica, proporcionaron un ligador de longitud apropiada y/o un sitio de inserción apropiado dentro de la proteína de superficie que se eligen para evitar el impedimento estérico entre los monómeros durante el ensamblaje del trímero. No existe límite de tamaño superior para la proteína heteróloga siempre y cuando la proteína heteróloga no forme un volumen que se traslape con los volúmenes vecinos de otras proteínas heterólogas montadas. De manera preferente se usa un sistema adaptador para mostrar la proteína heteróloga sobre la proteína de superficie de rotavirus trimérica. Las proteínas heterólogas que no son solubles, forman oligómeros de orden superior o agregados que no se consideran típicamente adecuados en practicar la invención. Cualquiera de los problemas de tamaño debido al impedimento estérico se puede superar al dirigir un ligador apropiado o sistema adaptador.

En el contexto de la proteína heteróloga, el término "heterólogo" significa típicamente que la proteína no es una proteína de rotavirus. En algunas modalidades, la expresión "proteína heteróloga" puede significar que la proteína no se derive de la misma cepa de rotavirus que se usa para mostrar la proteína.

Las proteínas de superficie triméricas son particularmente útiles para mostrar proteínas heterólogas triméricas. Ejemplos de proteínas heterólogas triméricas incluyen proteínas de entrada de células víricas triméricas y otras proteínas de superficie vírica triméricas. En particular aquellas que se orientan por anticuerpos neutralizantes. Ejemplos específicos son la hemaglutinina de influenza (HA), virus de inmunodeficiencia humana (HIV) gp140, la glicoproteína del virus del Ébola, la glicoproteína del virus de la rabia (RVG), la proteína Env de virus de encefalitis de artritis caprina, la proteína F del virus sincitial respiratorio (RSV), la proteína gB y sus complejos encontrados en los virus de herpes simple humanos y citomegalovirus humano (HCMV) etc.

En algunas modalidades, la proteína de superficie de rotavirus y/o la proteína heteróloga incluye una etiqueta de trimerización para ayudar en ensamblaje del complejo de heterohexámeros formado por la proteína de superficie de rotavirus trimérica y la proteína heteróloga trimérica. La etiqueta de trimerización, etiquetas de trimerización basadas particularmente en bobina enrollada (por ejemplo GCN4, [10]), también puede servir como módulos estructurales para extender el espacio disponible para una proteína heteróloga que se muestra sobre la superficie de una partícula de rotavirus recubierta con una proteína VP7 de rotavirus modificada. Otra etiqueta de trimerización adecuada se puede derivar de la fibritina T4 de bacteriófago [11].

Las proteínas heterólogas que no son triméricas pueden interactuar con una etiqueta de trimerización presente en las proteínas de superficie de rotavirus. Como ejemplo, si la proteína heteróloga tiene una  $\alpha$ -hélice accesible que podría interactuar con la  $\alpha$ -hélice es de una etiqueta de trimerización basada en una súper hélice, esa proteína heteróloga podría ligar la etiqueta de trimerización. Si no existe impedimento estérico, hasta tres de tales proteínas heterólogas podrían ligar la etiqueta de trimerización simultáneamente, formando un atado de seis hélices. Si menos de tres de tales proteínas heterólogas se ligan, el atado de seis hélices sería incompleto, conteniendo las tres hélices interiores contribuidas por la etiqueta de trimerización presente en la proteína de rotavirus pero solo una o dos de las hélices exteriores del atado. La proteína heteróloga no trimérica podría tener 1, 2, 4, o más unidades (es decir ser un monómero, dímero, tetrámero, u otro oligómero), con la condición de que al menos una de las subunidades tuvieron una  $\alpha$ -hélice que podría interactuar con la etiqueta de trimerización presente en la proteína de

superficie de rotavirus.

Una proteína heteróloga monomérica puede incluir múltiples  $\alpha$ -hélices arregladas tal que más de una hélice podría interactuar con las  $\alpha$ -hélices de la etiqueta de trimerización presente en la proteína de superficie de rotavirus, formando un atado de seis hélices completo o parcial.

Tetrámeros (o estructuras de orden superior) en donde al menos tres de los monómeros contienen  $\alpha$ -hélices adecuadamente arregladas podrían asociarse con la etiqueta de trimerización unida a la proteína de superficie de rotavirus para formar un atado de seis hélices. En ciertas modalidades, los tetrámeros (o estructuras de orden superior), en donde al menos dos de los monómeros contiene  $\alpha$ -hélices adecuadamente arregladas, podrían asociarse con la etiqueta de trimerización asociada con la proteína de superficie de rotavirus para formar un atado de seis hélices parcial, perdiendo una de las hélices exteriores. Las subunidades de la proteína heteróloga pueden ser idénticas (es decir, un homodímero, o motrímero, homotetrámero, etc.), o no pueden ser idénticas (es decir, un heterodímero, heterotrímero, heterotetrámero, etc.).

Si la etiqueta de trimerización presente en la proteína de superficie de rotavirus no es una súper hélice  $\alpha$ -helicoidal, el elemento estructural de interacción en la proteína heteróloga podría tener una estructura secundaria que liga la etiqueta de trimerización pero no es  $\alpha$ -helicoidal.

De manera preferente, la proteína heteróloga no es un anticuerpo que se liga específicamente a una proteína de superficie de rotavirus.

#### *Sistema adaptador*

Un sistema adaptador para mostrar la proteína heteróloga en una partícula de rotavirus. El sistema adaptador se compone típicamente de dos moléculas adaptadoras. La primera molécula adaptadora se liga covalentemente a la proteína de superficie de rotavirus que se eligió para mostrar una proteína heteróloga. La segunda molécula adaptadora se liga covalentemente a la proteína heteróloga. La primera molécula adaptadora y la segunda molécula adaptadora interactúan entre sí para formar un complejo estable.

Los sistemas adaptadores se componen típicamente de dos polipéptidos adaptadores. El primer polipéptido adaptador se fusiona a la proteína de superficie de rotavirus que se eligió para mostrar una proteína heteróloga. El segundo polipéptido adaptador se fusiona a la proteína heteróloga. El primer polipéptido adaptador y el segundo polipéptido adaptador pueden interactuar entre sí para formar un complejo estable.

Son conocidos muchos polipéptidos que se asocian entre sí para formar un complejo estable. En algunos casos, estos polipéptidos se derivan de diferentes proteínas, por ejemplo un receptor y un ligando o un anticuerpo y un antígeno. En otros casos, estos polipéptidos se pueden derivar de la misma proteína tal como las dos secuencias de repetición héptadas de VIH gp41.

Los sistemas adaptadores adecuados que no se dependen de las interacciones entre dos polipéptidos adaptadores también se pueden contemplar. Por ejemplo, la proteína de superficie de rotavirus se puede modificar para incluir un sitio de glicosilación específico, y la proteína heteróloga se puede modificar para incluir un dominio de lectina que reconoce específicamente el glicano en el sitio de glicosilación. Ejemplos de lectinas que reconocen estructuras de glicano específicas son bien conocidos en la técnica.

Otro sistema adaptador adecuado que no depende de las interacciones entre dos polipéptidos adaptadores es el sistema de estreptavidina-biotina. La estreptavidina monomérica que se ha mutado para prevenir la formación de tetrámeros y para mejorar la solubilidad se usa de manera preferente [12]. La estreptavidina monomérica se puede fusionar a la proteína de superficie de rotavirus que se eligió para mostrar una proteína heteróloga. Para que se presente la formación compleja, la proteína heteróloga se biotinizila. La biotinylación enzimática es preferida ya que permite que la biotina se ligué específicamente a un residuo de aminoácido presente en la proteína que se va a biotinilar. Por ejemplo, la proteína heteróloga se puede modificar por la inserción de un "AviTag" o "Péptido Aceptor" (AP), que se puede biotinilar específicamente por una biotina ligasa (por ejemplo, BirA) en presencia de biotina y ATP (ver referencia 13 para detalles).

Otro sistema adaptador que no depende de las interacciones entre dos polipéptidos adaptadores puede incluir un anticuerpo o ligado a un hapteno (por ejemplo, ácido dietilentriaminopentaacético (DTPA), ácido 1,4,7,10-tetraazaciclododecano-1,4,7,10-tetraacético (DOTA)). Alternativamente, los quelantes DTPA y DOTA se pueden coordinar con iones de metal para formar complejos.

El uso de un sistema adaptador no es esencial para practicar la invención. Sin embargo, el uso de un sistema adaptador se puede referir para conservar óptimamente las características estructurales de la proteína heteróloga, en particular de las proteínas de superficie triméricas de otros virus que se van a demostrar en una partícula de rotavirus. El uso de un sistema adaptador es particularmente ventajoso debido a que la proteína de

superficie de rotavirus que se ha modificado contiene un primer adaptador necesario para ser preparado solo una vez y se puede usar subsecuentemente con cualquier proteína heteróloga que se ha modificado para contener un segundo adaptador compatible.

5 Una ventaja adicional de usar un sistema adaptador es que la expresión y purificación se pueden optimizar separadamente para cada uno de los componentes de la proteína de superficie quimérica. En la mayoría de casos, la inserción de una secuencia de polipéptidos adaptadora no cambiará esas características de la proteína de superficie o de la proteína heteróloga que son relevantes para su expresión o purificación. De esta manera los sistemas de expresión existentes y los métodos de purificación se pueden usar típicamente sin modificación para preparar grandes cantidades de la proteína de superficie vírica que comprende el primer polipéptido adaptador y la proteína heteróloga que comprende el segundo polipéptido adaptador. En contraste, una proteína de superficie quimérica compuesta de una proteína de superficie vírica ligada covalentemente a una proteína heteróloga tendrá probablemente las características que difieren sustancialmente de las características de la proteína de superficie vírica o la proteína heteróloga cada una por sí misma.

15 Las proteínas de superficie de rotavirus y muchas proteínas de superficie de reovirales forman trímeros. De esta manera, en un aspecto de la invención, el sistema adaptador de elección forma de manera preferente un complejo trimérico. Se prefieren particularmente sistemas adaptadores en donde el primero polipéptido adaptador y el segundo polipéptido adaptador comprenden secuencias de repetición héptadas y forman estructuras  $\alpha$ -helicoidales que permiten la formación de atados de seis hélices. De manera más preferente, tres copias del segundo polipéptido adaptador forman un trímero de súper hélice interior, mientras que tres copias del primer polipéptido adaptador se empaican en las hendiduras en la superficie de este trímero completando el atado de seis hélices. Sin embargo, también puede ser posible lo inverso, en cuyo caso tres copias del primer polipéptido adaptador forma un trímero de súper hélice interior, mientras que tres copias del segundo polipéptido adaptador se empaican en las hendiduras en la superficie de este trímero. Varios pares de polipéptidos que comprenden secuencias de repetición héptadas y forman atados de seis hélices son conocidos en la técnica. Estos se derivan típicamente de proteínas de fusión vírica. Debido a la simetría y estabilidad rotacional de sus atados de seis hélices, las proteínas de fusión vírica son la fuente preferida para las secuencias de polipéptidos adaptadoras.

30 Una de las proteínas de fusión víricas más completamente estudiadas es la glicoproteína de envoltura del virus de inmunodeficiencia humano 1 (VIH-1). El ectodominio de la glicoproteína de envoltura VIH -1 consiste de gp120 y gp41. gp41 media la fusión entre las membranas virales y celulares. gp41 comprende dos secuencias de repetición héptadas, "región helicoidal 1" (HR1) y "región helicoidal 2" (HR2), que pueden formar un atado de seis hélices en la proteína nativa [14]. En una modalidad preferida, el primer polipéptido adaptador, y el segundo polipéptido adaptador del sistema adaptador corresponden a gp41 HR2 (SEQ ID NO: 1) y gp41 HR1 (SEQ ID NO: 2), respectivamente. En una modalidad preferida adicional, el primer polipéptido adaptador y el segundo polipéptido adaptador del sistema adaptador corresponden a la proteína del virus Nipah F HR2 (SEQ ID NO: 3) y HR1 (SEQ ID NO: 4), respectivamente. Sin embargo, se conocen varias otras proteínas de fusión vírica que pueden ser una fuente de polipéptidos HR1 y HR2 de interacción que se pueden usar para preparar el primero y segundo polipéptidos adaptadores. Estos incluyen la hemaglutinina 2 de influenza (HA2), las subunidad transmembrana (TM) del virus de leucemia de murino Moloney (Mo-MLV), la proteína F paramixovirus (incluyendo proteína F paramixovirus aviar, por ejemplo proteína F del virus Hendra) por ejemplo del virus sincitial respiratorio (RSV) y virus de enfermedad de Newcastle (NDV), la proteína pico de coronavirus (por ejemplo las proteínas pico de virus de hepatitis de ratón y SARS-CoV), la proteína ZEBRA del virus de Epstein-Barr, la proteína de fusión 5 del virus de simio etc. (ver referencias 14, 15, 16 y 17).

50 Donde la proteína heteróloga sola comprende secuencias de repetición héptadas y estas secuencias son compatibles con la secuencia de repetición héptada comprendida en una proteína de superficie de rotavirus modificada, la adición de las secuencias de repetición héptadas a la proteína heteróloga pueden no ser necesarias. En tal caso, las secuencias de repetición héptadas nativas de la proteína heteróloga pueden formar un atado de seis hélices con la secuencia de repetición héptada de la proteína de superficie de rotavirus modificada y la proteína heteróloga puede no ser necesaria ser modificada por la adición de una secuencia de repetición héptada que es extraña a la proteína heteróloga.

55 Por ejemplo, cualquiera de las proteínas de fusión víricas listadas en lo anterior contienen secuencias de repetición héptadas de interacción, y estas secuencias de repetición héptadas nativas pueden ser capaces de interactuar con las secuencias de repetición héptadas en una proteína de superficie de rotavirus modificada para contener las secuencias de repetición héptadas. En algunas modalidades, puede ser preferido depender de las secuencias de repetición héptadas heterólogas a la proteína de fusión vírica, en particular donde el sistema adaptador de dos partes se eligen para permitir un procedimiento modular donde la misma proteína de superficie de rotavirus modificada se usa para mostrar varias proteínas heterólogas diferentes.

65 Son conocidos otros dominios de proteínas que comprenden secuencias de repetición héptadas que forman atados de seis hélices. Estas incluyen el dominio CARD de Apaf-1 y RAIDD, dominio efector de muerte de FADD y el dominio de muerte de p75 y Fas. Otras fuentes de secuencias de péptido que forman atados de seis hélices incluyen



las SNAREs y GCN4-pII. Cualquiera de estos dominios o subsecuencias de péptidos se pueden adaptar del mismo modo para el uso en la presente invención. Por ejemplo, las tres hélices N-terminal y las tres hélices C-terminal del dominio CARD se pueden usar para el primero polipéptido adaptador y el segundo polipéptido adaptador, respectivamente, para formar un sistema adaptador adecuado para las proteínas de superficie monoméricas o diméricas.

Otros sistemas adaptadores basados en péptidos también se pueden usar. Por ejemplo, una porción de ligación a anticuerpo o antígeno del mismo (por ejemplo, un fragmento Fab, un scFv, un anticuerpo de dominio (DAb)) que liga la proteína heteróloga que se desea mostrar en la superficie del virus se puede insertar en la proteína de superficie de rotavirus. El anticuerpo o porción de ligación a antígeno del mismo puede reconocer un epítipo específico que ya sea se puede insertar en la proteína heteróloga (por ejemplo, en la forma de una etiqueta de péptido) o puede estar naturalmente presente en la proteína heteróloga.

En algunas modalidades, un sistema adaptador de dos partes adecuado requiere la modificación de la proteína heteróloga solamente. Por ejemplo, en una modalidad de la invención, la proteína de superficie de rotavirus es una glicoproteína. En tal caso, una proteína heteróloga se puede modificar para contener un dominio de lectina que liga específicamente el glicano en la glicoproteína. Por ejemplo, la proteína VP7 del rotavirus se puede modificar por la introducción de un solo sitio de glicosilación en la porción de superficie expuesta de la proteína de modo que el dominio de lectina de la proteína heteróloga modificada se liga específicamente al glicano en la proteína VP7 modificada.

Alternativamente, la proteína heteróloga se puede modificar para contener un dominio de ligación a antígeno de un anticuerpo que es específico para un epítipo en el antígeno de superficie de rotavirus. Por ejemplo, la proteína heteróloga se puede fusionar a un fragmento Fab de un anticuerpo que liga el antígeno de superficie de rotavirus con alta afinidad.

#### *Secuencia ligadora*

Las secuencias ligadoras separan los dominios derivados de diferentes proteínas y permiten que estos dominios se plieguen apropiadamente. Son conocidas en la técnica muchas secuencias ligadoras (ver referencia 18). Una secuencia ligadora se puede incluir para separar la proteína de superficie de rotavirus y/o la proteína heteróloga de la secuencia adaptadora. Alternativamente, donde la proteína de superficie de rotavirus y la proteína heteróloga forman dos partes de la misma proteína, ambas partes se pueden separar entre sí por una secuencia ligadora.

Tanto los ligadores rígidos como flexibles son conocidos. Una secuencia típica de un ligador flexible se compone de repeticiones de los aminoácidos G y S. Por ejemplo, el ligador puede tener la siguiente secuencia: GS, GSG, SGG, GGSGG (SEQ ID NO: 5) o GSGSGSGT (SEQ ID NO: 6). En algunas modalidades, la misma secuencia se repite múltiples veces (por ejemplo dos, tres, cuatro, cinco o seis veces) para crear un ligador más largo. En otras modalidades, un solo aminoácido tal como S o G se pueden usar como un ligador. Un ligador rígido se puede componer de varias repeticiones de la secuencia de aminoácidos EAAAR (SEQ ID NO: 7).

Cuando se elige una secuencia ligadora, se debe tomar cuidado para seleccionar un ligador hidrofílico para evitar la agregación de la proteína de superficie de rotavirus modificada o la proteína heteróloga.

Típicamente, el ligador es insensible a la proteasa pero en algunas modalidades el ligador contiene un sitio de escisión de proteasa. Los sitios de escisión de proteasa pueden ser útiles para remover etiquetas que se incluyen para supresión/purificación de por ejemplo una proteína de superficie de rotavirus modificada de la invención. Los sitios de escisión de proteasa también pueden ser útiles para exponer una secuencia de aminoácidos que se ha insertado en una proteína de superficie de rotavirus modificada para que la secuencia de aminoácidos sea más accesible en la superficie exterior de la proteína de superficie de rotavirus modificada después de la escisión con una proteasa específica de sitio. Por ejemplo, un sitio de escisión de proteasa se puede usar para exponer un polipéptido adaptador para mejorar su ligación al polipéptido adaptador correspondiente de un sistema adaptador de dos partes que se ha fusionado a una proteína heteróloga que se va a mostrar sobre la superficie de una partícula de rotavirus. En algunas modalidades, el uso de una escisión de proteasa para exponer un polipéptido adaptador que forma parte de una proteína de superficie de rotavirus modificada puede evitar la necesidad de secuencias ligadoras adicionales y de esta manera reducir el número de secuencias de aminoácidos adicionales que necesitan ser insertadas en una proteína de superficie de rotavirus modificada a fin de formar un complejo estable con una proteína heteróloga.

Los sitios de escisión de proteasa para una proteasa específica se pueden encontrar en una variedad de proteínas. Por ejemplo, la cascada de coagulación de sangre y la cascada de complementación contiene una variedad de proteasas muy específicas que reconocen los sitios de escisión en las proteínas más allá de la cadena abajo en la cascada. Usualmente, las enzimas en las etapas tempranas de una cascada son más específicas que las son las posteriores. Por ejemplo, el factor X es más específico que la trombina. Si la trombina se usa, los sitios de escisión sensibles a la trombina más preferida son aquellos encontrados en el fibrinógeno, factor X, y protrombina.

Ejemplos adicionales de proteasas de la cascada de coagulación de sangre, sus proteínas objetivo y sitios de escisión específicos se listan en la siguiente Tabla 1. La porción subrayada de la secuencia mostrada en la Tabla 1 es el sitio de escisión mínimo que necesita ser incluido para la que la proteasa reconozca el objetivo.

5 Tabla 1

Proteasa	Objetivo	Sitio(s) de Escisión
Factor humano XI	Factor humano IX	OTSKLTRA <b>E</b> AVF (SEO ID NO: 8) y SFNDFTRWG <b>G</b> E (SEQ ID NO: 9)
Calicreína humana	Factor humano XII	LFSSMTRWGGLV (SEO ID NO: 10)
Factor humano XII	Factor humano XI	KIKPRIVGGT (SEO ID NO: 11)

Otras proteasas que se han usado para escindir las proteínas de fusión incluyen enterocinasa, colagenasa, quimosina, urocinasa, renina, proteasa de Rinovirus 3C, proteasa del Virus de Marca del Tabaco (TEV), factor Xa, trombina, furina, y ciertas peptidasas señal (ver por ejemplo referencia 5).

De manera preferente el sitio de escisión se coloca de tal manera en el constructo final que cualquier etiqueta que se ha agregado a la proteína de superficie de rotavirus o la proteína heteróloga se puede remover fácilmente.

En algunas modalidades, el ligador contiene una etiqueta para detección y/o purificación. Muchas etiquetas para facilitar la detección de las proteínas son conocidas en la técnica. Las etiquetas de péptido frecuentemente usadas incluyen etiqueta FLAG (DYKDDDDK; SEQ ID NO: 12), etiqueta HA (YPYDVPDYA; SEQ ID NO: 13), etiqueta His (por ejemplo HHHHHH; SEQ ID NO: 14), etiqueta Myc (EQKLISEEDL; SEQ ID NO: 15), etiqueta Strep I (AWRHPQFGG; SEQ ID NO: 16), etiqueta Strep II (NWSHPQFEK; SEQ ID NO: 17), y etiqueta de proteína C (EDQVDPRLIDGK; SEQ ID NO: 18). Las etiquetas His son preferidas ya que permiten la fácil detección por los anticuerpos anti-His y permiten la purificación de la proteína etiquetada usando una columna de níquel. La etiqueta Strep II permite una purificación simple y fácil de las proteínas recombinantemente expresadas que usan columnas de estreptavidina. En algunos casos, se usan etiquetas de proteínas. Por ejemplo, una etiqueta de glutatona-S-transferasa se puede incluir para permitir la fácil purificación de una proteína de la invención usando una columna que comprende glutatona inmovilizada. Se puede usar una etiqueta de proteína fluorescente verde si se requiere una fácil detección por microscopia de fluorescencia.

Una etiqueta (por ejemplo una etiqueta de proteína C) se incluye como parte de la secuencia ligadora pero las etiquetas usadas típicamente por ejemplo para la detección y purificación de proteínas no interfieren con la fusión y plegamiento de la proteína etiqueta y se exponen generalmente en la superficie. Por lo tanto una etiqueta puede proporcionar propiedades ventajosas sobre otras secuencias ligadoras artificialmente diseñadas.

En otras modalidades, una secuencia ligadora puede comprender una secuencia de epítomos. La inclusión de una secuencia de epítomos pueden ser útiles para empacar los fragmentos de anticuerpos que reconocen el epítopo para estabilizar adicionalmente el complejo formado por una proteína de superficie de rotavirus modificada y una proteína heteróloga a través de un sistema adaptador de dos partes. La estabilización del complejo puede ser particularmente importante para lograr imágenes de alta resolución para estudios estructurales.

#### 50 *Péptido señal*

La proteína de superficie de rotavirus y la proteína heteróloga se modifican adicionalmente para comprender una secuencia de péptido señal heteróloga, reemplazando de manera preferente la secuencia de péptido señal nativa. El uso de un péptido señal heterólogo puede ser ventajoso para lograr niveles de expresión más altos en el sistema de expresión usado para preparar grandes cantidades de la proteína de superficie de rotavirus modificada y la proteína heteróloga para recubrir las DLPs de rotavirus. Por consiguiente, el péptido señal heterólogo se derivará de una proteína que se sabe se va a expresar en altos niveles en el sistema de expresión elegido. Por ejemplo, la secuencia señal consenso de HIV o péptido señal del activador de plasminógeno de tejido humano (tPA) son particulares adecuadas para la expresión en las células humanas. Para la expresión en las células de insecto, el péptido señal gp64 de Baculovirus o la secuencia señal de melitina de abeja de la miel se pueden usar. Típicamente, un ligador se coloca después de la secuencia de péptido señal heteróloga a fin de garantizar la eficiente escisión de péptido señal. Generalmente, el péptido señal se remueve por una peptidasa señal endógena al sistema de expresión elegido para que no esté presente en la proteína final (es decir la proteína de superficie de rotavirus y la proteína heteróloga recuperada del sistema de expresión).

**Ácidos nucleicos**

La invención también se refiere a ácidos nucleicos que comprenden un marco de lectura abierto que codifica una proteína de superficie quimérica de la invención ligada operacionalmente a una secuencia promotora tal que la proteína de superficie quimérica se expresa en grandes cantidades en un sistema de expresión.

La invención también se refiere a un constructo de ácido nucleico que codifica una proteína de superficie de rotavirus modificada, que comprende todo o una porción de una proteína de superficie de rotavirus, un primer polipéptido adaptador y, opcionalmente, una secuencia ligadora. La invención se refiere además a un ácido nucleico que comprende un marco de lectura abierto que codifica una proteína de fusión que comprende una proteína heteróloga, un segundo polipéptido adaptador, y, opcionalmente, una secuencia ligadora, en donde el marco de lectura abierto se liga operacionalmente a una secuencia promotora tal que la proteína de fusión se expresa en grandes cantidades en un sistema de expresión. En una modalidad, la invención se refiere a un constructo de ácido nucleico que comprende una secuencia de nucleótidos que codifica un segundo polipéptido adaptador, opcionalmente, una secuencia ligadora y un sitio de clonación múltiple, en donde la inserción de una región de codificación para una proteína heteróloga en el sitio de clonación múltiple produce un marco de lectura abierto que codifica una proteína de fusión que comprende la proteína heteróloga y el segundo polipéptido adaptador. El constructo de ácido nucleico comprende además una secuencia promotora que puede conducir la expresión de la proteína de fusión que comprende la proteína heteróloga y el segundo polipéptido adaptador en un sistema de expresión.

El primer polipéptido adaptador y el segundo polipéptido adaptador forman parte de un sistema adaptador de dos partes de modo que la proteína de fusión que comprende la proteína heteróloga y el segundo polipéptido adaptador y la proteína de superficie de rotavirus modificada que comprende el primer polipéptido adaptador forman un complejo estable entre sí.

En un aspecto particular, la invención se refiere a una composición de ácido nucleico que comprende

- a) un ácido nucleico que comprende un marco de lectura abierto que codifica una proteína de superficie de rotavirus modificada que comprende una proteína de superficie de rotavirus VP7, un primer polipéptido adaptador y, opcionalmente, un conector; en el que opcionalmente el marco de lectura abierto está vinculado operativamente a una secuencia promotora; y
- (b) un ácido nucleico que comprende un marco de lectura abierto que codifica una proteína de fusión que comprende una proteína heteróloga, un segundo polipéptido adaptador y, opcionalmente, un enlazador; en el que opcionalmente el marco de lectura abierto está vinculado operativamente a una secuencia promotora;

en donde los polipéptidos adaptadores primero y segundo son capaces de formar un complejo estable entre sí.

**Sistemas de expresión**

La invención también se refiere a sistemas de expresión para expresar las proteínas codificadas por los ácidos nucleicos de la invención.

En una modalidad, se usa un primer sistema de expresión para expresar una proteína de superficie de rotavirus modificada que comprende todo o una porción de una proteína de superficie de rotavirus. Un primer polipéptido adaptador y, opcionalmente, una secuencia ligadora. Un segundo sistema de expresión se usa para expresar una proteína de fusión que comprende una proteína heteróloga, un segundo polipéptido adaptador, y, opcionalmente, una secuencia ligadora. Opcionalmente, se usa un tercer sistema de expresión para expresar una o más proteína(s) de rotavirus que, junto con la proteína de superficie de rotavirus, forma(n) la capa exterior de una partícula de rotavirus. El primero y segundo polipéptidos adaptadores interactúan entre sí para formar un complejo estable. El primer sistema de expresión comprende un primer constructo de ácido nucleico que comprende un marco de lectura abierto que codifica la proteína de superficie de rotavirus modificada, en donde el marco de lectura abierto se liga operacionalmente a una secuencia promotora. El segundo sistema de expresión comprende un segundo constructo de ácido nucleico que comprende un marco de lectura abierto que codifica la proteína de fusión, en donde el marco de lectura abierto se liga operacionalmente a una secuencia promotora. El tercer sistema de expresión que comprende uno o más vector(s) de expresión para la una o más proteína(s) de rotavirus. La proteína de superficie de rotavirus modificada y la proteína de fusión y opcionalmente la una o más proteína(s) de rotavirus después se purifican. La proteína de superficie de rotavirus modificada y la proteína de fusión se pueden mezclar en relaciones apropiadas para formar una proteína de superficie quimérica. La proteína de superficie quimérica y opcionalmente la una o más proteína(s) de rotavirus después se usa para recubrir las DLPs de rotavirus para formar partículas de rotavirus que muestran la proteína heteróloga sobre su superficie. Alternativamente, las DLPs de rotavirus se recubren con la proteína de superficie de rotavirus modificada y opcionalmente la una o más proteína(s) de rotavirus para formar partículas de rotavirus. Las partículas de rotavirus después se pueden mezclar con las proteínas de fusión para permitir la formación de un complejo entre el primer polipéptido adaptador y el segundo polipéptido

adaptador que produce una proteína de superficie química de modo que la proteína heteróloga se muestra sobre la superficie de las partículas de rotavirus.

5 En otra modalidad, la invención se refiere a un sistema de expresión que comprende (i) un primero constructo de ácido nucleico que comprende un marco de lectura abierto que codifica una proteína de superficie de rotavirus modificada que comprende todo o una porción de una proteína de superficie de rotavirus, un primer polipéptido adaptador y, opcionalmente, una secuencia ligadora, en donde el marco de lectura abierto se liga operacionalmente a una secuencia promotora, y (ii) un segundo constructo de ácido nucleico que comprende un marco de lectura abierto que codifica una proteína de fusión que comprende una proteína heteróloga, un segundo polipéptido adaptador, y, opcionalmente, una secuencia ligadora, en donde el marco de lectura abierto se liga operacionalmente a una secuencia promotora. En algunos casos, el sistema de expresión comprende además un vector de expresión para una o más proteína(s) de rotavirus que, junto con la proteína de superficie de rotavirus, forma(n) la capa exterior de la partícula de rotavirus.

15 El sistema de expresión puede ser una célula bacteriana, una célula de levadura, una célula de protozooario, una célula de insecto o una célula de mamífero. El uso de células bacterianas o células de levadura como sistemas de expresión es menos preferido, y en particular donde la glicosilación apropiada de las proteínas expresadas se desea.

## 20 Kits

La invención proporciona adicionalmente kits que comprende un primer constructo de ácido nucleico que codifica una proteína de superficie de rotavirus modificada que comprende todo o una porción de una proteína de superficie de rotavirus y un primer polipéptido adaptador, y un segundo constructo de ácido nucleico, en donde el segundo constructo de ácido nucleico comprende una secuencia de nucleótidos que codifica un segundo polipéptido adaptador y un sitio de clonación múltiple, y en donde la inserción de una región de codificación para una proteína heteróloga en el sitio de clonación múltiple produce un marco de lectura abierto que codifica una proteína de fusión que comprende la proteína heteróloga y un segundo polipéptido adaptador, en donde el primer polipéptido adaptador de la proteína de fusión química y el segundo polipéptido adaptador de la proteína de fusión interactúan entre sí para formar un complejo estable.

La invención se refiere adicionalmente a kits que comprenden un primer constructo de ácido nucleico que codifica una proteína de superficie de rotavirus modificada que comprende todo o una porción de una proteína de superficie de rotavirus y un primer polipéptido adaptador, y un segundo constructo de ácido nucleico que codifica una proteína de fusión que comprende todo o una porción de una proteína heteróloga y un segundo polipéptido adaptador, en donde el primer polipéptido adaptador de la proteína de superficie de rotavirus modificada y el segundo polipéptido adaptador de la proteína de fusión interactúan entre sí para formar un complejo estable, produciendo de esta manera una proteína de superficie química.

40 Los kits pueden comprender además una partícula de rotavirus. La partícula de rotavirus puede ser ya sea de la misma especie de la cual la proteína de superficie de rotavirus se derivó o de una especie diferente. Por ejemplo, la VP7 de rotavirus de rhesus se puede usar para recubrir las DLPs preparadas de rotavirus de bovino y viceversa. El rotavirus se le puede eliminar el recubrimiento y recubrir con la proteína de superficie química. Alternativamente, los kits pueden comprender DLPs para recubrimiento con la proteína de superficie química. Las DLPs de rotavirus se pueden preparar al eliminar el recubrimiento de las partículas de rotavirus nativas o al expresar recombinantemente las proteínas de envoltura interior de rotavirus VP2 y VP6.

## Expresión recombinante de proteínas de capa exterior

50 El recubrimiento de las DLPs de rotavirus requiere solamente una o dos proteínas virales recombinantes, la proteína de capa exterior VP7 o la proteína de capa exterior VP7 junto con la proteína pico de capa exterior VP4. La expresión y purificación de VP4 y VP7 se describen con detalle en las referencias 20 y 21, respectivamente.

55 Las proteínas de capa exterior incluyendo la proteína de superficie química de la invención se pueden producir usando sistemas de expresión convencionales conocidos por la persona experta. A fin de garantizar el plegamiento correcto, y en algunos casos la glicosilación apropiada, los sistemas de expresión diferentes a los sistemas de expresión procariotas o de levadura son preferidos. Por ejemplo, las células de mamífero tales como células CHO o células 293 se pueden usar para sobreexpresar las proteínas de capa exterior necesarias para el recubrimiento de las DLPs de rotavirus. Alternativamente, el protozooario *Leishmania tarentolae* se puede usar para expresar las proteínas de capa exterior. Los sistemas de células de insecto también son adecuados para la expresión de las proteínas de capa exterior. Por ejemplo, las líneas de células de insecto Sf9, Sf21 y Hi-5 son adecuadas para la sobreexpresión de las proteínas glicosiladas. En algunos casos, son preferidos los sistemas de células de insecto basados en baculovirus. Los sistemas de expresión descritos en las referencias 2, 4 y 5 son particularmente adecuados en practicar la invención.

65

Varias formas para recubrir y purificar las proteínas de capa exterior sobreexpresadas son conocidas en la técnica. Típicamente se usa una serie de pasos cromatográficos para purificar las proteínas sobreexpresadas de los extractos citoplásmicos o el sobrenadante de las células que se usó como el sistema de expresión. Por ejemplo la afinidad de lecitina, inmunoafinidad y cromatografía de exclusión de tamaño se pueden usar. Si la proteína de cubierta exterior se ha etiquetado con una etiqueta de péptido o proteína, esta etiqueta se puede usar ventajosamente para purificación. Sin un sitio de escisión de proteasa está presente en la secuencia que precede la etiqueta, la etiqueta se puede remover después de la purificación usando una proteasa que reconoce específicamente el sitio de escisión de proteasa.

En algunos casos, una preparación cruda de las proteínas de capa exterior recombinante expresadas se puede usar para la reacción de recubrimiento. Por ejemplo, los lisados de células usados para expresar las proteínas de capa exterior se pueden preparar usando una solución amortiguadora de lisis y/o la alteración mecánica de las células (por ejemplo, por raspado o sonicación de las células). Cualquier resto celular se remueve por centrifugación, y el sobrenadante que contiene una preparación cruda de las proteínas de capa exterior recombinantes se puede usar en una reacción de recubrimiento. La preparación cruda se puede concentrar usando ultrafiltración antes de ser usada en una reacción de recubrimiento.

Donde una proteína heteróloga se liga no covalentemente a una proteína de superficie de rotavirus por un sistema adaptador de dos partes que comprende un primer polipéptido adaptador y un segundo polipéptido adaptador, la proteína de superficie de rotavirus fusionada al primer polipéptido adaptador y la proteína heteróloga fusionada al segundo polipéptido adaptador se expresan separadamente en diferentes células. La expresión separada puede ser preferible debido a que ambas proteínas se pueden purificar separadamente usando protocolos de purificación conocidos para cada una de las proteínas. Después de la purificación, las proteínas se pueden mezclar en las relaciones apropiadas requeridas para el recubrimiento de las DLPs de rotavirus. Por ejemplo, la relación molar de VP7 y VP4 en la capa exterior de rotavirus es 13:1. La relación molar de la proteína VP7 modificada que comprende el primer polipéptido adaptador y la proteína heteróloga que comprende el segundo polipéptido adaptador es típicamente 1:1. Alternativamente, las DLPs se recubren con la proteína de capa exterior de rotavirus para formar partículas de rotavirus. Las partículas de rotavirus después se mezclan con la proteína heteróloga para permitir la formación de un complejo entre el primer polipéptido adaptador y el segundo polipéptido adaptador para que la proteína heteróloga se muestre sobre la superficie de las partículas de rotavirus.

Alternativamente, tanto la proteína de superficie de rotavirus como la proteína heteróloga se pueden expresar en la misma célula. La co-expresión da por resultado la formación de un complejo estable de la proteína de superficie de rotavirus y la proteína heteróloga mediada por el sistema adaptador de dos partes (ver lo anterior para detalles). Donde la secuencia ligadora que conecta la secuencia adaptadora a la proteína de superficie de rotavirus o a la proteína heteróloga que contiene etiqueta, esta etiqueta se puede usar para purificar el complejo de las células/sobrenadante celular del sistema de expresión.

Para ser capaz de recubrir una partícula de núcleo retroviral con una proteína de superficie quimérica que comprende todo o parte de una proteína heteróloga, no todas las proteínas necesitan ser proporcionadas para formar una capa exterior. Por ejemplo, la capa exterior de un retrovirus de mamífero se forma de  $\mu 1$ ,  $\sigma 1$  y  $\sigma 3$ . A fin de recubrir las partículas de núcleo reovíricas para los propósitos de la invención, proporcionar típicamente una versión recombinante de  $\mu 1$  es suficiente para formar las partículas reovíricas. Opcionalmente, las versiones recombinantes de  $\sigma 1$  y  $\sigma 3$  también se pueden proporcionar. Las proteínas recombinantes se pueden preparar con uno de los sistemas de expresión mencionados en lo anterior y después se pueden purificar usando protocolos conocidos.

### **Propagación de partículas reovíricas**

Los rotavirus se han descubierto en un gran número de especies animales incluyendo ganado, cerdos, caballos, conejos, ratones, perros, gatos, aves y especies de animales exóticos tales como addax, saiga, gnu de cola blanca, oso grizzli y canguro rojo. Por consiguiente, la idoneidad de un cierto tipo de célula para su propagación depende de intervalo de hospedero del virus elegido. De manera preferente, se elige un virus que permite la propagación de alto rendimiento en los sistemas de cultivo celulares establecidos.

Por ejemplo, la línea de células MA104 se puede usar para propagar el rotavirus de rhesus. El rotavirus se puede recubrir al lisar las células infectadas, por ejemplo por congelamiento o descongelamiento de las células MA104 infectadas en el medio. Los restos celulares se eliminan de los lisados por centrifugación de baja velocidad, y las partículas reovíricas se concentran al formarlas en pelotillas usando ultracentrifugación o por ultrafiltración. La suspensión concentrada de partículas reovíricas se puede purificar adicionalmente usando la centrifugación en gradiente CsCl [2]. Una solución amortiguadora adecuada para la preparación del gradiente de CsCl es TNC (Tris 20 mM pH 8.0, NaCl 100 mM, CaCl<sub>2</sub> 1 mM).

De manera preferente, se usa una célula bien caracterizada para la propagación de rotavirus. Los rotavirus humanos, de bovino y de mono rhesus se pueden propagar en las células Vero. Una línea de células Vero

adecuadas (CCL81) se puede obtener de la Recolección de Cultivo de Tipo Americano (ATCC).

5 Similarmente, los miembros de la familia Reoviridae se han aislado de una gama de mamíferos, aves, reptiles, peces, crustáceos, insectos, garrapatas, arácnidos, plantas y hongos. Los sistemas de cultivo también se han establecido para otros virus en la familia Reoviridae. Por ejemplo, el virus de enanismo desigual del arroz stunt virus (RRSV) se puede preparar en cantidades suficientes de hojas de arroz infectadas [7]. Las células C6/36 se pueden usar para propagar el virus Banna (BAV) [22]. Las células L de ratón, en particular la línea de células L929, y las células de eritroleucemia de murino (MEL) se pueden usar para propagar las cepas de reovirus de mamífero tipo 1 Lang (TIL) y tipo 3 Dearing (T3D) [5]. El aquareovirus de la carpa herbívora (GCRV) se puede propagar en el cultivo de células de riñón de *Ctenopharyngodon idellus* (CIK) [8].

15 Las partículas de virus se pueden recubrir típicamente de los sobrenadantes celulares o células lisadas. Por ejemplo, cualquier resto celular se puede remover por centrifugación de baja velocidad, las partículas reovíricas se forman en pelotillas del sobrenadante usando ultracentrifugación. Las pelotillas contienen partículas de virus que se pueden purificar adicionalmente si es necesario.

### Preparación de la DLP

20 Las DLPs de rotavirus se preparan al remover las proteínas de capa exterior de las partículas de rotavirus purificadas. Las partículas de rotavirus se les pueden eliminar el recubrimiento al incubar varias partículas en presencia de un quelante de calcio tal como EDTA o EGTA. Una solución amortiguadora adecuada para eliminar el recubrimiento de las partículas de rotavirus contiene Tris 20 mM, pH 8, NaCl 100 mM y EDTA 1 mM [2]. Las partículas de rotavirus también se les pueden eliminar el recubrimiento por choque térmico. Las DLPs resultantes se pueden purificar para agrupamiento en los gradientes CsCl preformados sucesivos ( $\rho = 1.25$  a  $1.50$   $35 \text{ g/cm}^3$ ).

25 El método de elección para remover las proteínas de capas exterior depende de la partícula vírica elegida. Por ejemplo, las partículas de reovirus de mamífero se pueden separar de núcleo al incubar la partícula con  $\alpha$ -quimotripsina (CHT) durante dos horas a  $37^\circ\text{C}$  [4].

30 Los métodos anteriores se pueden adaptar en otros virus en la familia Reoviridae dependiendo de la sensibilidad de sus proteínas de capas exteriores al tratamiento de proteasa o condiciones bajas en calcio.

35 En algunos casos, la partícula subvívica se puede recubrir al lisar las células que se usaron para propagar el virus. La partícula subvívica después se puede recubrir de los lisados celulares, por ejemplo, por un gradiente CsCl como se describe para el virus Banna [22].

40 Alternativamente, las DLPs o partículas subvívicas se pueden preparar al expresar recombinantemente las proteínas virales relevantes en uno de los sistemas de expresión descritas en lo anterior. Por ejemplo, las proteínas de envoltura interior de rotavirus VP2 y VP6 se pueden expresar recombinantemente de partículas similares a virus que se asemejan a las DLPs.

### Recubrimiento

45 El recubrimiento de las DLPs rotavirus se presenta en un intervalo de pH incluyendo pH 4.5 y 6.5, pero es más eficiente entre pH 5 y 5.5 (de manera preferente pH 5.2). El recubrimiento se presenta en temperaturas incluyendo  $4^\circ\text{C}$  y  $37^\circ\text{C}$ , pero es más eficiente entre  $4^\circ\text{C}$  y  $30^\circ\text{C}$  que en  $37^\circ\text{C}$ .

50 El recubrimiento de las partículas de núcleo reovíricas se puede hacer típicamente bajo condiciones que imitan las condiciones bajo las cuales el ensamblaje de las partículas reovíricas se presenta en la célula. Por ejemplo, las partículas de núcleo de ortoreovirus de mamífero se pueden recubrir a  $37^\circ\text{C}$ . Sin embargo, los parámetros individuales tales como pH y temperatura pueden requerir optimización adicional para el re-ensamblaje eficiente de las partículas reovíricas de los núcleos virales purificados y las proteínas de capa exterior recombinantemente expresadas.

55 Las partículas de virus recubiertas se pueden recubrir por centrifugación en gradiente CsCl. La purificación en gradiente CsCl se puede repetir una segunda vez para incrementar la pureza de las partículas de virus recubiertas.

60 Donde un sistema adaptador se usa para mostrar una proteína heteróloga sobre la superficie de una partícula de rotavirus recubierta, las DLPs de rotavirus se pueden recubrir con una proteína de superficie de rotavirus que comprende un primer polipéptido adaptador. La proteína heteróloga que comprende un segundo polipéptido adaptador, que forma un complejo estable con el primer polipéptido adaptador, se puede agregar subsecuentemente a las partículas de virus recubiertas. Esto puede ser ventajoso donde el recubrimiento con el complejo formado por la proteína heteróloga y la proteína de superficie de rotavirus sería menos eficiente, por ejemplo debido al tamaño de la proteína heteróloga. Alternativamente, un complejo entre la proteína heteróloga y la proteína de superficie de

65

rotavirus a través de los polipéptidos adaptadores se pueden formar primero, y este complejo se puede agregar a las DLPs de rotavirus para el recubrimiento.

### Formación de complejos

5 En algunos aspectos de la invención, la proteína de superficie quimérica se usa para estudiar la estructura de los complejos formados entre la proteína heteróloga que forma parte de la proteína de superficie quimérica (por ejemplo al ser ligado no covalentemente a una proteína de superficie de rotavirus a través de un sistema adaptador de dos partes) y una molécula que se liga específicamente a la proteína heteróloga.

10 En un aspecto la molécula es una molécula proteínica. La molécula proteínica es típicamente de otra proteína tal como un receptor o ligando que interactúa con la proteína heteróloga o un anticuerpo o fragmento de anticuerpo que reconoce un epítipo encontrado sobre la superficie de la proteínas heteróloga. Por ejemplo, un complejo hexamérico de una proteína de superficie de rotavirus trimérica ligada a una proteína de entrada de célula vírica trimérica (por ejemplo, proteína VIH gp140 o RSV F) que se muestra sobre la superficie de la partícula de rotavirus se puede usar para estudiar la interacción entre la proteína de entrada de célula vírica y su receptor de superficie de célula hospedera.

15 Las condiciones óptimas para la formación de complejos entre la proteína heteróloga y la molécula proteínica dependen de la naturaleza de la interacción. Las interacciones de receptor-ligando pueden requerir diferentes condiciones de interacciones de anticuerpo-antígeno. Típicamente, la formación de complejos se lleva a cabo a temperatura ambiente en una solución amortiguada (por ejemplo, solución salina amortiguada con fosfato).

20 En una modalidad, la molécula proteínica se agrega a una suspensión de las partículas de rotavirus recubiertas que muestran la proteína de superficie quimérica. La solución amortiguada puede contener componentes adicionales tales como Ca<sup>2+</sup> para prevenir la eliminación del recubrimiento de las partículas de rotavirus y opcionalmente uno o más inhibidores de proteasa para bloquear la degradación de las proteínas. Después de la incubación de la molécula proteínica en presencia de las partículas de rotavirus recubiertas, los complejos recientemente formados de las moléculas proteínicas ligadas a las partículas de rotavirus se pueden separar de cualquiera de las moléculas proteínicas no ligadas por centrifugación o ultrafiltración. Sin embargo, la remoción de moléculas proteínicas no ligadas puede no es necesario si las partículas de rotavirus recubiertas se usan para análisis crio-EM.

25 Alternativamente, la molécula proteínica y la proteína de superficie quimérica se incuban conjuntamente para permitir que se presente la formación de complejos. Primero, la proteína heteróloga y el antígeno de superficie de rotavirus se pueden incubar conjuntamente para formar la proteína de superficie quimérica. Una vez que la proteína de superficie quimérica se ha formado, se agrega la molécula proteínica. El complejo de la molécula proteínica y la proteína de superficie quimérica después se pueden agregar a las DLPs de rotavirus para formar las partículas de rotavirus, opcionalmente en presencia de cualquiera de las proteínas de rotavirus adicionales que, conjuntamente con la proteína de superficie de rotavirus, forman la capa exterior de una partícula de rotavirus nativa.

30 En un aspecto particular de la invención, la proteína de superficie quimérica de la invención se usa para determinar la estructura de un complejo de antígeno-anticuerpo. En este aspecto de la invención, la proteína heteróloga se puede derivar de un patógeno tal como un virus o una bacteria. Por ejemplo, un antígeno inmunodominante se puede elegir como la proteína heteróloga para estudiar qué partes de la proteína se fijan como objetivo por los anticuerpos durante una respuesta inmunitaria contra el antígeno. De manera preferente, la proteína heteróloga es una proteína de superficie vírica trimérica tal como hemaglutinina de virus de influenza, virus sincitial respiratorio F, o VIH gp140. Para determinar la estructura de un complejo de antígeno-anticuerpo, el uso de un fragmento Fab en lugar del anticuerpo de longitud completa se prefiere típicamente para evitar el impedimento estérico entre las proteínas de superficie quiméricas vecinas y para garantizar la ocupación máxima del epítipo encontrado en la proteína heteróloga.

35 En otro aspecto de la invención, la formación de complejos entre una proteína heteróloga y una molécula no proteínica se puede estudiar. Una molécula no proteínica puede ser un ácido nucleico (por ejemplo, ARN o ADN), un polisacárido u oligosacárido (por ejemplo, un glicano). Por ejemplo, la proteína heteróloga puede ser un factor de transcripción u otra proteína de ligación a ADN que forma un complejo con una secuencia de ADN específica. Alternativamente, la proteína heteróloga puede ser una lecitina que forma un complejo con un glicano.

### Crio-EM

40 Usando la crio-EM para la determinación de estructura tiene varias ventajas sobre procedimientos más tradicionales tales como cristalografía de rayos X. En particular, la crio-EM pone menos requisitos riguroso en la muestra que se va a analizar con respecto a la pureza, homogeneidad y cantidad. De manera importante, la crio-EM se puede aplicar para objetivos que no forman cristales adecuados para la determinación de la estructura.

65

Una suspensión de partículas de rotavirus recubiertas no purificadas o purificadas, ya sea solas o en complejo con una molécula proteínica tal como un anticuerpo o molécula no proteínica tal como un ácido nucleico, se puede aplicar a rejillas de carbono para formación en imágenes por crió-EM. Las rejillas recubiertas se congelan instantáneamente, usualmente en etanol líquido, para conservar las partículas en la suspensión en un estado congelado-hidratado. Las partículas más grandes se pueden verificar por criofijación. La muestra vitrificada se puede cortar en secciones delgadas (típicamente de 40 a 200 nm de grueso) en un crió-ultramicrotomo, y las secciones se pueden colocar en rejillas de microscopio electrónico para formar en imágenes.

La calidad de los datos obtenidos de las imágenes se puede mejorar al usar iluminación paralela y mejor alineación de microscopio para obtener resoluciones tan altas como  $\approx 3.3$  Å. En tal alta resolución, es posible la construcción del modelo ab initio de las estructuras de átomo completo. Sin embargo, la formación en imágenes en menor resolución podrían ser suficientes en donde los datos estructurales en la solución atómica en la partícula de rotavirus elegida o estrechamente relacionada y la proteína heteróloga seleccionada o un homólogo cercanos son disponibles para la modelación comparativa limitada (ver a continuación).

Para mejorar adicionalmente la calidad de los datos, el microscopio se puede alinear cuidadosamente para revelar los anillos de función de transferencia de constante visibles (CTF) más allá de  $1/3$  Å<sup>-1</sup> en la transformada de Fourier de las imágenes de la película de carbono registradas las mismas condiciones usadas para la formación de imágenes. Los valores de desenfoque para cada micrografía después se pueden determinar usando el software tal como CTFFIND [23]. El tamaño del pixel final del mapa de densidad se puede calibrar usando, por ejemplo, el virus del mosaico de tabaco (TMV).

Las descripciones útiles para aplicar la crió-EM a los estudios estructurales de las partículas de rotavirus se encuentran en las referencias 24 y 25.

### **Análisis de formación de imágenes y determinación de estructura**

Las imágenes obtenidas por crió-EM se analizan para identificar las micrografías de partículas individuales. La selección de partículas individuales se puede hacer con la ayuda de herramientas de software tal como SIGNATURE [26]. El desenfoque astigmático, eje de inclinación de espécimen, y ángulo de inclinación para cada micrografía se puede determinar usando el programa de Computadora CTFTILT [23]. La obtención de los valores de desenfoque separados para cada partícula de acuerdo con su coordenada de la imagen original mejora la calidad de los datos del mapa de densidad crió-EM que se obtiene al promediar las micrografías de partícula individual de las partículas de rotavirus.

El ajuste de los modelos atómico conocidos dentro de un mapa de densidad de crió-EM es un procedimiento común para construir modelos de estructuras complejas tales como partículas reovíricas. Son disponibles una variedad de herramientas de ajuste de computación que varían de localización de cuerpo rígido simple de estructuras de proteínas, tales como [27], Foldhunter [28] y Mod-EM [29], para la formación de complejos y algoritmos de ajuste flexibles dinámicos similares a NMFF [30], Flex-EM [31], MDFF [32] y DireX [33, 34], que transforman las estructuras conocidas a un mapa de densidad.

Cuando un modelo atómico no es conocido, se pueden usar mapa de densidad crió-EM de la construcción y/o evaluación de los modelos estructurales de una galería de modelos potenciales que se construyen in silico (ver referencias 29, 35, 36, 37 y 38). Una estructura de plantilla relacionada puede ser conocida para la moderación comparativa limitada o, para la modelación ab initio limitada, el plegamiento que se modela debe ser relativamente pequeño. Por ejemplo, una estructura inicial se puede obtener usando IMIRS [39]. La alineación y reconstrucción adicional se puede llevar a cabo con FREALIGN [40] usando una estructura de rotavirus conocida y una estructura conocida de la proteína heteróloga o un homólogo cercano como plantilla.

La información estructural y funcional significativa se puede obtener directamente del mapa de densidad mismo. Por ejemplo, en resoluciones de 5-10 Å, algunos elementos de estructura secundarios son visibles en los mapas de densidad crió-EM: las  $\alpha$ -hélices parecen como cilindros, mientras que las  $\beta$ -hojas parecen como placas curvadas, delgadas. Estos elementos de estructura secundarias se pueden identificar y cuantificar confiablemente usando herramientas de reconocimiento de característica para describir una estructura de proteína o inferior en la función de las proteínas individuales. En las resoluciones casi atómicas (3-5 Å), la inclinación de las  $\alpha$ -hélices, separación de las  $\beta$ -hebra, así como las densidades que las conectan, se pueden visualizar no ambiguamente (ver, por ejemplo, las referencias 41, 42, 43 y 44).

La construcción de modelo de novo en crió-EM comprende el reconocimiento de característica, análisis de secuencia, correspondencia de elemento de estructura secundaria, colocación de C $\alpha$  y optimización del modelo. Se pueden usar varias aplicaciones de software, por ejemplo, EMAN para la segmentación y manipulación del mapa de densidad [45], SSEHunter [46] para detectar los elementos de estructura secundarios, la visualización en la quimera de UCSF [47] y la manipulación atómica en Coot [48,49].



Los programas de identificación de estructura secundaria como SSEHunter proporcionan un mecanismo semi-automatizado para detectar y mostrar elementos de estructura secundaria visualmente observables en un mapa de densidad [46]. El registro de los elementos de estructura secundaria en una secuencia y estructura, combinados con la información geométrica y biofísica, se puede usar para anclar la cadena principal de la proteína en el mapa de densidad [41, 43]. Esta correspondencia de secuencia a estructura se relaciona con los elementos de estructura secundarios observados en la densidad para aquellos predichos en la secuencia. El kit de herramienta de modelación GORGON acopla la predicción de estructura secundaria basada en secuencia con la detección de característica y las técnicas de modelación geométricas para generar los modelos de cadena principal de proteína iniciales [50]. Los métodos de modelación automáticos tales como EM-IMO (optimización modular de microscopía electrónica, -iterativa) se pueden usar para construir, modificar y refinar las estructuras locales de los modelos de proteínas usando mapas de crio-EM como una limitante [51].

Una vez que se ha determinado la correspondencia usando el elemento de estructura secundaria, los átomos C $\alpha$  se pueden asignar a la densidad que comienza con las  $\alpha$ -hélices y es seguido por las  $\beta$ -hebras y bucles. Por ejemplo, al tomar ventaja de los golpes claros para los átomos de C $\alpha$ , los modelos C $\alpha$  se pueden construir usando la utilidad Baton build en los programas de cristalográficos O [52] y/o Coot [48]. Las posiciones de C $\alpha$  se pueden ajustar iterativamente tal que ajustan la densidad mientras que mantienen óptimamente las geometrías razonables y eliminan los choques dentro del modelo. Los modelos de átomo completo grueso se pueden refinar en una manera pseudocristalográfica usando CNS [53]. Los modelos se pueden optimizar adicionalmente usando software de modelación computacional tal como Rosetta [36]. Los modelos de átomo completa también se pueden construir con la ayuda de otras herramientas computacionales tal como REMO [54]. La calidad de un modelo se puede confirmar por comparación visual del modelo con el mapa de densidad. El análisis de factor R/Rfree pseudocristalográfico [55] proporciona una medición del acuerdo entre las amplitudes del factor de estructura observadas y calculadas y se puede usar para confirmar que el modelo atómico obtenido proporciona un buen ajuste a los mapas de densidad crio-EM. La geometría de modelo de proteínas se puede verificar por PROCHECK [56].

#### Métodos de clasificación para el diseño de inmunógeno

Usando la crio-EM en lugar de la cristalografía de rayos X facilita la rápida determinación de estructura. En comparación con la cristalografía de rayos X, la crio-EM pone menos requisitos astringentes en la pureza, homogeneidad y cantidad de la muestra que se va a analizar. Estas características hacen la crio-EM atractiva en el contexto del diseño de inmunógeno para vacunas, particularmente contra patógenos que se someten a la deriva antigénica.

Teniendo más información estructural especialmente acerca de los epítomos compartidos por las variantes del mismo patógeno o patógenos estrechamente relacionados puede permitir el diseño racional de inmunógenos, que se pueden usar en vacunas que proporcionan amplia protección contra un gran número de variantes de un gran número de patógeno o patógenos estrechamente relacionados. Tales así llamadas vacunas "universales" se cree que son más económicas que las vacunas tradicionales debido a que harían superfluo e incluir varias variantes del mismo patógeno o patógenos estrechamente relacionados en la misma composición de vacuna (como es el caso, por ejemplo, para las vacunas de la polio actualmente disponibles así como para vacunas de conjugados estreptocócicos y meningocócicos) o proporcionan más composiciones de vacuna cada año para tomar en cuenta la deriva antigénica que se presentó en la población de patógenos en la estación previa (como es el caso para las vacunas de influenza).

El diseño racional de un inmunógeno para vacunación contra un patógeno diverso implica la identificación de esas regiones de una proteína inmunodominante que se conservan entre diversas variantes/subtipos del mismo patógeno o entre los patógenos estrechamente relacionados. El uso de partículas de rotavirus para mostrar proteínas heterólogas en las proteínas de superficie de rotavirus quimérica hace posible determinar rápidamente un gran número de estructuras en un tiempo relativamente corto. Por lo tanto los métodos de la invención pueden ser particularmente útiles en la identificación de epítomos conservados de proteínas inmunogénicas.

Al determinar la estructura de inmunógenos en el complejo con anticuerpos o fragmentos de ligación a antígeno correspondientes que se han encontrado que son inducidos contra un número de variantes/subtipos del mismo patógeno y/o contra patógenos estrechamente relacionados, regiones conservadas en inmunógeno inmunodominante de esta patógeno y/o patógenos estrechamente relacionados se pueden identificar. En una modalidad, la invención por lo tanto se refiere a un método para obtener un modelo tridimensional de un inmunógeno combinado con un anticuerpo en donde el método comprende los pasos de (i) recubrir una DLP de rotavirus con una proteína de superficie quimérica que comprende el inmunógeno para producir una suspensión de partículas de rotavirus que muestran la proteína de superficie quimérica, (ii) agregar a la suspensión o anticuerpo o fragmento de anticuerpo que se liga específicamente al inmunógeno, en donde el anticuerpo o fragmento forma un complejo con la proteína de la superficie quimérica, (iii) congelar la suspensión, (iv) formar en imágenes las partículas de rotavirus usando crio-EM para obtener una pluralidad de micrografías, y (vi) analizar la pluralidad de micrografías para obtener un modelo tridimensional del inmunógeno formado en complejo con el anticuerpo. El inmunógeno es típicamente heterólogo a la partícula de rotavirus. El modelo tridimensional del inmunógeno formado en complejo con el

anticuerpo se puede usar para definir el epítipo sobre la superficie del inmunógeno que es reconocido por el anticuerpo.

5 Este método se puede usar para identificar el epítipo ligado por un anticuerpo que tiene actividad  
ampliamente neutralizante contra un número de patógenos relacionados o variantes de un patógeno que surgen  
debido a la deriva antigénica. Alternativamente, este método se puede usar para identificar el epítipo ligado por un  
anticuerpo que puede reconocer un epítipo que interfiere por ejemplo con la función de una proteína de entrada de  
células víricas y se puede usar para prevenir o inhibir la propagación del virus. Un anticuerpo que se liga a una  
10 configuración tridimensional específica de una proteína también puede ser útil como un reactivo de control de calidad  
durante la fabricación de la proteína, por ejemplo para excluir lotes de una proteína fabricada que contiene un  
número grande de copias desnaturalizadas o plegadas de manera errónea de esta proteína. Este método se puede  
emplear adicionalmente para identificar epítipos que son particularmente útiles cuando se ligan por un anticuerpo  
que se usa como un reactivo de diagnóstico (por ejemplo, epítipos que son específicos a un patógeno particular en  
15 grupo de patógenos estrechamente relacionados).

Al repetir el método para obtener un modelo tridimensional de un inmunógeno formado en complejo con un  
anticuerpo para un panel de anticuerpos o sus fragmentos de ligación a antígeno correspondientes que se ha  
encontrado que se inducen en respuesta a la infección por un número de variantes/subtipos del mismo patógeno y/o  
20 contra patógenos estrechamente relacionados, el epítipo de cada anticuerpo sobre la superficie del inmunógeno se  
puede identificar. Una vez que el epítipo de cada anticuerpo en el panel se ha identificado, esta información se  
puede usar para el diseño racional de un inmunógeno. Para ejemplo, el método se puede usar para identificar los  
epítipos de un panel de anticuerpos que se ha descubierto tienen actividad neutralizante contra un número de  
variantes/subtipos del mismo patógeno y/o contra patógenos estrechamente relacionados para diseñar un  
inmunógeno que proporciona actividad neutralizante amplia contra un gran número de variantes/subtipos del mismo  
25 patógeno y/o patógenos estrechamente relacionados.

Alternativamente, el método descrito en lo anterior se puede emplear para mapear el repertorio de epítipos  
reconocidos por los anticuerpos que se inducen en respuesta a la inmunización con una vacuna existente. Al mapear  
los epítipos reconocidos más comúnmente por estos anticuerpos, se pueden identificar epítipos  
30 inmunodominantes. Esta información se puede usar ventajosamente para utilizar adicionalmente una vacuna  
existente. Por ejemplo, donde una vacuna existente se compone de un patógeno inactivado, una versión optimizada  
de la vacuna también puede incluir solamente las porciones inmunodominantes del patógeno (por ejemplo, en la  
forma de una vacuna de subunidad que comprende versiones recombinantes de los antígenos o epítipos  
inmunodominantes identificados). El entendimiento de los determinantes estructurales de los epítipos de  
35 anticuerpos inmunodominantes, obtenidos usando esta técnica, se puede aplicar para preparar antígenos que se  
han diseñado para hacer los epítipos más útiles (tales como epítipos ampliamente neutralizantes)  
inmunodominantes.

Ejemplo de mapeo de epítipos para ayudar en el diseño de inmunógenos son conocidos en la técnica pero  
40 se han limitado generalmente a un pequeño número de epítipos en un inmunógeno debido a los esfuerzos  
implicados en obtener datos estructurales para los complejos de inmunógeno-anticuerpo por cristalografía de rayos  
X. Solo frecuentemente se somete a prueba un anticuerpo (ver referencias 57, 58 y 59). La presente invención por  
primera vez permite obtener información estructural para complejos de inmunógenos-anticuerpo usando un  
procedimiento de alto rendimiento. La invención es particularmente útil para anticuerpos que solo se asocian  
45 débilmente con el inmunógeno y de esta manera presentan mayores retos para la cristalización.

Varios enfoques para diseñar inmunógenos que sean ampliamente protectores contra un número de  
variantes del mismo patógeno o patógenos estrechamente relacionados son conocidos en la técnica. En años  
recientes, se han hecho muchos esfuerzos para diseñar vacunas "universales" contra el virus de influenza y VIH (ver  
50 referencias 60, 61 y 62). El diseño racional de un inmunógeno modificado se ha impedido por la falta de datos  
estructurales para guiar el proceso. El uso de partículas de rotavirus para mostrar proteínas heterólogas en las  
proteínas de superficie de rotavirus quiméricas hace posible determinar rápidamente como las modificaciones de la  
secuencia de aminoácidos de una proteína afectan su estructura tridimensional. Las modificaciones de aminoácidos  
que cambian la estructura de una proteína pueden afectar su capacidad de formar cristales. De esta manera, la  
55 información estructural para evaluar el impacto de estas modificaciones puede no ser fácilmente disponible debido a  
que la proteína modificada puede no cristalizarse bajo las mismas condiciones como la proteína no modificada.  
Puesto que el uso de crió-EM no depende de la formación de los cristales para la determinación de las estructuras  
de proteína y, además, no requieren ninguna adaptación de la configuración experimental básica para diferentes  
tipos de proteínas, es posible la rápida caracterización estructural de las proteínas modificadas.  
60

Una vez que un epítipo en una proteína inmunogénica de un patógeno se ha identificado que está presente  
en un número de variantes/subtipos del mismo patógeno y/o en muchos patógenos estrechamente relacionados,  
esta información se puede usar para diseñar una vacuna universal. En algunos casos, esto puede requerir la  
modificación del inmunógeno, particularmente donde una región conservada comprende un epítipo compartido no  
65 es fácilmente accesible a los anticuerpos en la proteína nativa, ya que la pobre accesibilidad del epítipo se traduce

usualmente en pobre inmunogenicidad. Al modificar el inmunógeno de tal manera que retenga el epítipo nativo que se puede reconocer por un anticuerpo, pero haciendo el epítipo más accesible, el inmunógeno modificado puede producir una respuesta de anticuerpo que es protectora contra un amplio intervalo de variantes/subtipos del mismo patógeno o de patógenos estrechamente relacionados. Además del mapeo de los epítopos de los anticuerpos madurados por afinidad, los epítopos reconocidos por los ancestros no mutados de estos anticuerpos, una vez deducido de los experimentos de clonación de repertorios de células B, se puede mapear por esta técnica. Esta información podría ayudar en el diseño de inmunógenos que amplifican selectivamente un anticuerpo ancestral no mutado deseado y después dirigir su maduración por afinidad a un anticuerpo amplia y potentemente neutralizante.

Los inmunógenos candidato que se han modificado de un inmunógeno nativo (a fin de preparar una vacuna universal o por otras razones) se puede someter a prueba para presencia de un epítipo nativo de interés usando los métodos de la invención. En una modalidad particular, la invención por lo tanto se refiere adicionalmente a un método para determinar las diferencias estructurales entre dos variantes de una proteína heteróloga, donde el método comprende los pasos de (i) recubrir una DLP de rotavirus con una primera proteína de superficie quimérica que comprende una primera variantes de la proteína heteróloga para producir una suspensión de una primera partícula de rotavirus que muestra la primera proteína de superficie quimérica, (ii) congelar la suspensión, (iii) formar en imagen la primera partícula de rotavirus usando crio-EM para obtener una pluralidad de macrófagos, (iv) analizar la pluralidad de macrófagos para obtener un modelo tridimensional de la primera proteína de superficie quimérica, (v) repetir los pasos (i)-(iv) con una segunda de proteína de superficie quimérica que comprende una segunda variante de la proteína heteróloga en donde la primera y segunda proteína de superficie quimérica son idénticas entre sí aparte de la diferencia entre la primera variante y la segunda variante, y (vi) comparar el modelo tridimensional de la primera proteína de superficie quimérica al modelo tridimensional de la segunda proteína de superficie quimérica para determinar las diferencias estructurales entre la primera variante y la segunda variante de la proteína heteróloga. Este método se puede repetir para variantes adicionales de la misma proteína heteróloga para proporcionar un arreglo de clasificación para la identificación de una variante con ciertas características estructurales deseadas.

En algunos casos, las estructuras del inmunógeno nativo y de una pluralidad de inmunógenos modificados cada uno formado en complejo con un anticuerpo que reconoce un epítipo se puede determinar. La comparación de la estructura de los inmunógenos modificados al inmunógeno nativo ayudará a seleccionar inmunógenos candidatos que conservar mejor el epítipo nativo para prueba adicional in vivo.

### Composiciones inmunogénicas

En un aspecto adicional, la presente invención se refiere al uso de partículas de rotavirus que comprenden la proteína de superficie quimérica de la invención en la preparación de un medicamento. En particular, las partículas de rotavirus que comprenden la proteína de superficie quimérica de la invención se pueden usar en composiciones inmunogénicas adecuadas para vacunación humana. Por ejemplo, la proteína de superficie quimérica se puede diseñar para mostrar una proteína heteróloga que induce una respuesta inmunitaria protectora cuando se administra a un paciente.

El uso de virus quiméricos para inducir una respuesta inmunitaria a un antígeno derivado de patógenos es bien conocido en la técnica. Tradicionalmente, este procedimiento requiere el rediseño del genoma del virus hospedero que se selecciona como un vector para antígeno derivado de patógeno. Tal procedimiento tiene un número de limitaciones. Típicamente, el genoma del virus se modifica para contener un gen derivado de patógeno que codifica el antígeno. El tamaño del genoma del virus de elección puede limitar por lo tanto el tamaño del antígeno derivado de patógeno que se puede expresar. Similarmente, la inserción de un gen que codifica el antígeno derivado de patógeno puede interferir con la propagación del virus sin limitar los rendimientos que se pueden lograr. Por lo tanto la preparación de virus quiméricos necesita ser optimizada para antígeno derivado de patógeno.

La creación de un virus quimérico con una nueva proteína de superficie que usa genética recombinante para alterar el genoma del virus hospedero al insertar una región de codificación para un antígeno derivado de patógeno podría cambiar potencialmente el intervalo de hospedero del virus quimérico y por lo tanto tiene consecuencias impredecibles en la patogenicidad de los virus recientemente creados. Por lo tanto se tienen que tomar pasos adicionales para proporcionar un virus quimérico suficientemente atenuado que no sea capaz de replicarse en el sujeto que se va a vacunar.

La presente invención supera los problemas asociados con el procedimiento tradicional debido a que permite la preparación de una clase de proteína de superficie quimérica que se puede usar para recubrir DLPs de rotavirus. Las partículas de rotavirus recubiertas no contienen ninguna información genética para el antígeno derivado de patógeno. De esta manera después de la infección inicial de una célula hospedera con las partículas de rotavirus recubiertas, no se forman virus de progenie que llevan el antígeno derivado de patógeno.

El uso de un sistema adaptador de dos partes reduce adicionalmente el número de pasos de optimización. Una vez que una proteína de rotavirus modificada fusionada a un primer polipéptido adaptador se ha preparado,

cualquier antígeno derivado de patógeno conocido se puede fusionar a un segundo polipéptido adaptador que forma un complejo estable con el primer polipéptido adaptador. Esto elimina cualquier necesidad para la ingeniería genética del rotavirus. En muchos casos, la expresión y purificación del antígeno derivado de patógeno no se afecta por la presencia del segundo polipéptido adaptador en el antígeno derivado de patógeno, de modo que no se requiere modificación de métodos de expresión y purificación existentes para el antígeno. Los protocolos existentes para recubrir las DLPs de rotavirus con una proteína de superficie quimérica que comprende la proteína de rotavirus modificada ligada al antígeno derivado de patógeno a través del sistema adaptador de dos partes se puede usar para preparar partículas de rotavirus que comprenden la proteína de superficie quimérica. Estas partículas de rotavirus después se pueden incluir en composiciones inmunogénicas para inducir una respuesta inmunitaria contra el antígeno derivado de patógeno.

La plataforma de vacuna es particularmente útil al preparar composiciones inmunogénicas en las cuales las partículas de rotavirus comprenden una proteína de superficie quimérica trimérica que contiene una proteína de entrada de células víricas triméricas heterólogas. Las proteínas de entrada de células víricas triméricas son los antígenos de superficie inmunodominantes de muchos virus. Con las excepciones del antígeno de superficie de hepatitis B y la proteína L1 de virus de papiloma humano, ambas de las cuales forman partículas similares a virus, e influenza HA, que forma rosetas, vacunas de subunidad basadas en los antígenos de superficie vírica inmunodominantes han fallado típicamente para dar por resultado vacunas clínicamente efectivas. Al mostrar estos antígenos de superficie vírica en su conformación nativa en las partículas de rotavirus, se pueden preparar composiciones inmunogénicas que inducen una respuesta inmunitaria protectora contra el virus del cual el antígeno de superficie se derivó originalmente.

Ejemplos específicos de la proteína de entrada de células víricas triméricas incluyen influenza HA, HIV gp140, la glicoproteína del virus del Ébola, glicoproteína del virus de la rabia, la proteína Env del virus de encefalitis de artritis caprina, la proteína RSV F, el gB y opcionalmente su complejo con otras proteínas del virus de herpes simple humano y de HCMV.

#### *Partículas de rotavirus*

Los rotavirus son particularmente útiles debido a que existen varias cepas animales que tienen un intervalo de hospedero definido. Por ejemplo, algunas cepas son específicas para vacas y otras para monos. La mayoría de cepas animales son antigénicamente distintas de aquellas cepas que infectan típicamente los humanos y por lo tanto son mucho más incapaces de provocar una enfermedad en los humanos. Por lo tanto los humanos no tienen típicamente ninguna inmunidad preexistente a estas cepas de rotavirus que podrían interferir con una respuesta inmunitaria contra la proteína de superficie quimérica cuando estas cepas animales se usan como vectores en las composiciones inmunogénicas para la vacunación humana. Además, estos virus son naturalmente atenuados en los humanos.

Además, dos vacunas de rotavirus atenuado, vivo patentadas se comercializan actualmente: RotarixMR y RotaTaqMR. RotarixMR contiene la cepa de rotavirus humano atenuado RIX4414, que se hizo pasar 26 veces en células de Riñón de Mono Verde Primario (AGMK) y se propagó en células Vero. RotaTaqMR contiene una combinación de cinco cepas reordenantes de rotavirus de bovino (BRV)-rotavirus humano (HRV), diseñadas como G1, G2, G3, G4, y P1, respectivamente, que también se propagan en las células Vero. Todos los reordenantes se componen del fondo del genoma de la cepa BRV WC3 (G6P7[5]) que expresa proteínas VP7 o VP4 humanas.

La vacuna basada en rotavirus rhesus previamente patentada (RRV) RotaShieldMR consiste de rotavirus reordenantes RRV (G3P5B[3]) y 3 RRV-HRV. Cada reordenante deriva 10 genes de RRV y un solo gen de HRV que codifica una proteína VP7 para la especificidad 1, 2, o 4 del serotipo G (G1P5B[3], G2P5B[3], y G4P5B[3]).

Adicionalmente, la seguridad de las cepas de rotavirus de bovino (NCDV RIT4237 G6P6) y de simio (cepa RRV MMU18006 G3P5B) monovalentes para el uso como vacuna se ha establecido.

La inmunización, seguridad clínica y experiencia de fabricación con estas cepas de vacuna se puede aplicar directamente a las partículas de rotavirus de la invención. Por ejemplo, cualquiera de los rotavirus contenidos en las vacunas de rotavirus patentadas se podría usar para preparar las DLPs para el recubrimiento con la proteína de superficie quimérica de la invención.

#### *Preparación*

Las partículas de rotavirus modificadas incluidas en las composiciones inmunogénicas se pueden preparar usando cualquiera de los métodos descritos en lo anterior. Las partículas de rotavirus nativas se pueden propagar y purificar. Las partículas de rotavirus purificadas después se les pueden eliminar el recubrimiento para preparar las DLPs que después se pueden recubrir con una proteína de superficie quimérica que comprende un antígeno derivado de patógeno.

Típicamente, una proteína de rotavirus modificada fusionada a un primer polipéptido adaptador se agrega a las DLPs para formar partículas de rotavirus. La adición de proteínas de rotavirus adicionales que, junto con la proteína de superficie de rotavirus, forman la capa exterior de una partícula de rotavirus nativa es opcional. Las partículas de rotavirus después se vuelven a incubar con un antígeno derivado de patógeno fusionado a un segundo polipéptido adaptador que forma un complejo estable con el primer polipéptido adaptador para proporcionar la proteína de superficie quimérica.

#### Formulaciones

La composición inmunogénica de la invención se puede proporcionar en la forma de un kit que comprende un primer recipiente que comprende partículas de rotavirus liofilizadas y un segundo recipiente que comprende una solución para la resuspensión extemporánea de las partículas de rotavirus liofilizadas. Las partículas de rotavirus liofilizadas pueden comprender uno o más lioprotector tales como sacarosa, dextrano, sorbitol y aminoácidos para estabilizar las partículas de rotavirus durante la liofilización.

Alternativamente, la composición inmunogénica se proporciona en un solo recipiente que comprende las partículas de rotavirus en suspensión. Donde la composición inmunogénica es para inyección, se puede proporcionar en una jeringa.

Cualquier solución puede contener uno o más excipiente(s).

Las soluciones son típicamente basadas en agua. Por lo tanto el agua purificada puede formar el excipiente principal. Por ejemplo, la dilución de las partículas de rotavirus para proporcionar la concentración final deseada se llevará a cabo usualmente con agua para inyección (WFI).

La solución contiene típicamente una solución amortiguadora. Por lo tanto los excipientes adicionales incluyen agentes amortiguadores y reguladores de pH tales como citrato de sodio, monohidrato de fosfato diácido de sodio, e hidróxido de sodio. Los antiácidos tales como carbonato de calcio se pueden agregar para prevenir la inactivación del virus durante el pasaje a través del estómago si la composición inmunogénica se administra oralmente. Un regulador de acidez tal como adipato de disodio también se puede incluir, de manera preferente en lugar del antiácido.

En algunos casos, un agente espesante tal como xantano puede estar presente como un excipiente adicional.

Un agente tensoactivo, en particular un agente tensoactivo no iónico tal como polisorbato 80, también puede estar presente.

Otros excipientes incluyen, sacarosa, sorbitol, sales inorgánicas, aminoácidos y vitaminas.

Las composiciones tendrán generalmente una osmolalidad de entre 200 mOsm/kg y 400 mOsm/kg, de manera preferente entre 240-360 mOsm/kg, y estarán de manera más preferente dentro del intervalo de 280-320 mOsm/kg.

El pH de una composición de la invención será generalmente entre 5.0 y 7.5, y más típicamente entre 5.0 y 6.0 para estabilidad óptima.

Las composiciones de la invención contienen de manera preferente <1 EU (unidad de endotoxina, una medición estándar; 1 EU es igual a 0.2 ng de la Endotoxina de estándar de referencia FDA EC 2 "RSE") por dosis, y de manera preferente <0.1 EU por dosis. Las composiciones de la invención están de manera preferente sin gluten. Adicionalmente, las composiciones de la invención son preferiblemente estériles.

#### Administración

Las composiciones inmunogénicas se pueden usar para estimular una respuesta inmunitaria de la mucosa. Por lo tanto las composiciones inmunogénicas se pueden administrar oral o intratraquealmente. Otras vías de administración tales como inyección intramuscular se pueden elegir dependiendo del antígeno derivado de patógeno mostrado sobre la superficie de la partícula de rotavirus modificada incluida en la composición inmunogénica de la invención.

#### General

El término "que comprende" abarca "que incluye" así como "que consiste" por ejemplo una composición "que comprende" X puede consistir exclusivamente de X o puede incluir algo adicional por ejemplo X + Y.

La palabra “sustancialmente” no excluye “completamente” por ejemplo una composición que esta “sustancialmente libre” de Y puede estar completamente libre de Y. Donde sea necesario, la palabra “sustancialmente” se puede omitir de la definición de la invención.

5 El término “aproximadamente” en relación con un valor numérico x significa, por ejemplo,  $x \pm 10\%$  o x + dos derivaciones estándares del valor. En ciertas modalidades, “aproximadamente” se entiende como variación aceptable y tolerancia dentro de la técnica específica. A menos que sea claro a partir del contexto, todos los términos numéricos en la presente se entiende que se modifican por aproximadamente. El término “anticuerpo” incluye fragmentos de anticuerpo tales como fragmentos de ligación a antígeno (Fabs), fragmentos variables de cadena individual (scFv), etc. El término “porción de ligación a antígeno” de un anticuerpo (o “porción de anticuerpo”) incluye fragmentos de un anticuerpo que retiene la capacidad para ligarse específicamente a un antígeno. Se ha mostrado que la función de ligación a antígeno de un anticuerpo se puede llevar a cabo por fragmentos de un anticuerpo de longitud completa. Ejemplos de fragmentos de ligación abarcados dentro del término “porción de ligación a antígeno” de un anticuerpo incluyen (i) un fragmento Fab, un fragmento monovalente que consiste de los dominios VL, VH, CL y CH1; (ii) un fragmento F(ab')<sub>2</sub>, un fragmento bivalente que comprende dos fragmentos Fab ligados por un puente de disulfuro en la región bisagra; (iii) un fragmento Fd que consiste de los dominios VH y CH1; (iv) un fragmento Fv que consiste de los dominios VL y VH de un brazo individual de un anticuerpo, (v) un fragmento dAb [63], que consiste de un dominio VH; y (vi) una región determinadora de complementariedad aislada (CDR). Adicionalmente, aunque los dos dominios del fragmento Fv, VL y VH, se codifican por genes separados, se puede unir, usando métodos recombinantes, por un ligador sintético que les permite ser fabricados como una cadena de proteínas individual en la cual las regiones VL y VH se porean para formar moléculas monovalentes (conocida como Fv de cadena individual (scFv); ver por ejemplo referencias 64 y 65. Tales anticuerpos de cadena individual también se proponen para ser abarcados dentro del término “porción de ligación a antígeno” de un anticuerpo.

25 Como se usa en la presente, “o” se usa para ser exclusiva y es intercambiable con “y/o”, al menos que se indique claramente de otra manera por el contexto.

30 Como se usa en la presente “un” y “el” se entiende que incluyen tanto singular como plural a menos que se indique claramente de otra manera por el contexto.

A menos que se establezca específicamente, un proceso que comprende un paso para mezclar dos o más componentes no requiere ningún orden específico de mezclado. Estos componentes se pueden mezclar en cualquier orden. Donde existan tres componentes entonces dos componentes se pueden combinar entre sí, y después la combinación se puede combinar con el tercer componente, etc.

35 Donde los materiales animales (y particularmente de bovino) se usan en el cultivo de células para la preparación de material para la administración a animales, especialmente humanos, se deben obtener de fuentes que se estén libres de encefalopatías espongiformes admisibles (TSEs), y en particular libres de encefalopatía espongiforme de bovino (BSE). En general, cuando los productos se preparan para administración a animales, y especialmente a humanos, se prefiere cultivar células en ausencia total de materiales derivados de animales.

40 Donde un sustrato de células se usa para el reordenamiento o procedimientos de genética inversa, es de manera preferente que haya sido aprobado para uso en la producción de vacunas humanas por ejemplo como en Ph Eur general capítulo 5.2.3 [66],

45 La identidad entre las secuencias de polipéptidos se determina de manera preferente por el algoritmo de búsqueda de homología de Smith-Waterman como es implementado en el programa MPSRCH (Oxford Molecular), usando una búsqueda de espacio de afinidad con parámetros de penalización de abertura de espacio =12 y penalización de =1.

## 50 BREVE DESCRIPCION DE LAS FIGURAS

Figuras 1A: Vista general esquemática de las modificaciones hechas a (A) la proteína VP7 de rotavirus y (B) la proteína heteróloga. HR1 y HR2 forman parte de un sistema adaptador de dos partes para ligar no covalentemente la proteína heteróloga a la proteína VP7 de rotavirus. Las secuencias de péptido señal heterólogas sirven el propósito de lograr altos niveles de expresión en los sistemas de expresión elegidos. Las etiquetas de afinidad permiten la fácil purificación. El sitio de reconocimiento de proteasa P1 se puede usar para remover la etiqueta de afinidad A1 después de la purificación. El segundo conjunto de sitios de reconocimiento de afinidad y proteasa (A2 y P2) se pueden reemplazar por las secuencias ligadoras adecuadas y pueden servir como una secuencia espaciadora que puede ser necesaria para mostrar la proteína heteróloga grande. La etiqueta de trimerización es opcional y se puede usar en ayudar en la formación de trimeros de algunas proteínas heterólogas que después facilita la ligación a la proteína VP7 de rotavirus trimérica. La etiqueta de trimerización, particularmente las etiquetas de trimerización basadas particularmente en superhélice (por ejemplo GCN4), también pueden servir como módulos estructurales para extender el espacio disponible para una proteína heteróloga que se muestra sobre la superficie de una partícula de rotavirus recubierta con una proteína VP7 de rotavirus modificada.

Figura 2: Vista general esquemática del complejo formado por una proteína VP7 de rotavirus trimérica modificada y una proteínas heteróloga trimérica. La proteína VP7 de rotavirus se liga no covalentemente a la proteína heteróloga a través de un sistema adaptador de dos partes, donde una parte del sistema adaptador (HR2) se liga a la proteína VP7 de rotavirus y la otra parte del sistema adaptador (HR1) se liga a la proteína heteróloga. HR1 y HR2 forman un atado de seis hélices que da por resultado un complejo estable para unir no covalentemente la proteína heteróloga a la proteína VP7 de rotavirus. La proteína de superficie quimérica puede ser parte de la capa exterior de la partícula de rotavirus por las de DLPs de recubrimiento in vitro.

Figuras 3: Micrografías electrónicas de (Figura 3A) DLPs purificadas, (Figura 3B) DLPs recubiertas con la proteína VP7 de rotavirus y (Figura 3C) DLPs recubiertas con la proteína VP7 que muestra el virus de influenza HA como la proteína heteróloga. Las partículas se tificaron negativamente antes de la adquisición de la imagen.

Figuras 4: (Figura 4A) Gel de acrilamida teñido con Coomassie después de SDS-PAGE. La línea 1 muestra el marcador de peso molecular (MW). La línea 2 se cargó con las DLPs purificadas usadas en la reacción de recubrimiento (DLP). La línea 3 se cargó con el complejo de proteínas VP7-HA purificado (pro) usado para el recubrimiento de las DLPs. La línea 4 se cargó con la mezcla de entrada de DLPs, a la proteína VP7 modificada y la proteína HA para la reacción de recubrimiento (inp). Las líneas 5-7 se cargaron con las bandas observadas después de la purificación de las partículas recubiertas en un gradiente CsCl (bandas 1 y 2 t banda superior). Las líneas 8-14 se cargan con las mismas muestras en el mismo orden como en las líneas 1-7, pero no se agregó agente reductor a las muestras antes de la SDS-PAGE. (Figura 4B) las DLPs de rotavirus recubiertas con el complejo de proteína VP7-HA en una gradiente CsCl. Las bandas 1 y 2 y la banda superior corresponden a las muestras cargadas en las líneas 5-7 en el panel A. Las posiciones donde las partículas recubiertas y las DLPs típicamente migrarían, se indican por flechas punteadas.

Figuras 5: Sección transversal a través de una reconstrucción tridimensional de la DLP de rotavirus recubierta con una proteína VP7 modificada (Figura 5A) que muestra HA, (Figura 5B) que muestra HA con fragmentos ScFv ligados del anticuerpo CR6261, y (Figura 5C) que muestra HA con fragmentos de Fab ligados del anticuerpo CR6261. El CR6261 reconoce una región helicoidal sumamente conservada en el tallo próximo a la membrana de HA1/HA2. La reconstrucción se basa en las imágenes adquiridas al llevar a cabo la crio-EM en las DLPs recubiertas.

Figura 6: Superposición de una DLP (blanco) en una reconstrucción tridimensional de una DLP recubierta con la proteína VP7 de rotavirus modificada y el virus de influenza HA como la proteína heteróloga modificada. La reconstrucción se basa en las imágenes adquiridas al llevar a cabo la crio-EM en las DLPs recubiertas.

Figura 7: Detalle de la reconstrucción tridimensional de una partícula de rotavirus que muestra el virus de influenza HA ligado a los fragmentos Fab del anticuerpo CR6261. Las partículas de rotavirus se prepararon al recubrir las DLPs con una proteína VP7 modificada que contiene una secuencia adaptadora (HR2). La proteína HA no se ligó covalentemente a la proteína VP7 modificada a través de una secuencia de repetición héptada HR1 C-terminal fusionada que forma un atado de seis hélices con las secuencias de repetición héptadas HR2 de la proteína VP7. CR6261 se liga en el extremo próximo a la membrana de cada subunidad HA (HA1 y HA2). La reconstrucción se basa en las imágenes adquiridas al llevar a cabo la crio-EM en las DLPs recubiertas.

## MODOS PARA LLEVAR A CABO LA INVENCION

### Ejemplo 1: Construcción del vector de expresión para las proteínas VP7, HA y HIV gpl40 recombinantes modificadas

El gen de VP7 de tipo natural de la cepa G3 de rotavirus rhesus (RRV) se clonó en un vector pFastBacDual (InvitrogenMR) entre los sitios de restricción BamH I y Not I a través de procedimientos estándares. Una secuencia Kozak (GCCACC; SEQ ID NO: 19) se diseñó en el extremo 5' antes del codón de inicio de la secuencia de codificación VP7. Las modificaciones de la secuencia de codificación VP7 se llevaron a cabo por PCR invertida. Se diseñaron cebadores de temperaturas de recocido apropiadas que contienen las modificaciones relevantes en los extremos de los cebadores. La Figura 1A muestra las diversas modificaciones que se agregaron a la secuencia de codificación de rotavirus.

La proteína final codificada por la secuencia de codificación VP7 modificada tiene las siguiente siguientes características: los primeros 21 aminoácidos codifican la secuencia señal de melitina de abeja para la expresión de la proteína modificada en las células de insecto. El péptido señal de proteína VP7 se ha removido. Se pueden usar otras secuencias señal como sea apropiado, tal como la secuencia señal consenso HIV o el péptido señal del activador de plasminógeno de tejido humano (htPA) para la expresión de células humanas. (b) Residuos de aminoácidos 22 a 33 forman una etiqueta de afinidad opcional (Strep-tag II más un ligador para una escisión de péptido señal más eficiente) para los propósitos de purificación de proteínas y se puede reemplazar por cualquiera

de otras etiquetas de afinidad, tales como His-tag, etiqueta HA, etiqueta FLAG, etc. (c) Residuos de aminoácidos 34 a 40 forman una secuencia de reconocimiento de proteasa TEV para la escisión de la etiqueta de afinidad y se pueden reemplazar por las secuencias de reconocimiento de cualquiera de otras proteasas, tales como la proteasa PreScissionMR (es decir proteasa de Rhinovirus 3C), factor Xa, enterocinasa, trombina, furina, etc. (d) Residuos de aminoácidos 41 a 136 son las secuencias de la porción N-terminal VP7 de rotavirus de rhesus (residuos de aminoácidos VP7 51 a 146). (e) El residuo de aminoácido 137 es un ligador de un aminoácido y se puede reemplazar por otra secuencia ligadora apropiada. (f) Residuos de aminoácidos 138 a 165 son parte de la repetición C-héptada de la cepa HIV gp41 HXB2 y se podría reemplazar por secuencias de repetición héptadas de otros retrovirus, paramixovirus, etc., siempre y cuando la secuencia de reemplazo sea compatible con las secuencias de repetición héptadas usadas en los constructos de proteína HA o gp140 modificados (ver a continuación). (g) Residuos de aminoácidos 166 a 187 son una secuencia de reconocimiento de factor Xa proteasa seguido por una etiqueta C de proteína flanqueada por ligadores; esta modificación es opcional y se puede reemplazar por otras secuencias, tal como un epítipo diseñado o simplemente una secuencia ligadora tal como GGSGGSGGSGGSGGS (SEQ ID NO: 20) o GGSGGSGGSGGSGGSGG (SEQ ID NO: 21). La presentación de un epítipo puede ser útil para empacar fragmentos de anticuerpo que reconocen el epítipo para estabilizar adicionalmente los ensamblajes mostrados (incluyendo la proteína heteróloga) para el propósito de estudios estructurales. Como se dice, el diseño para los estudios estructurales es muy complejo que el diseño para presentar antígenos. (h) Los residuos de aminoácidos 188 a 367 son la porción C-terminal de la CP7 de rotavirus de rhesus (residuos de aminoácidos VP7 147 a 326). La proteína VP7 modificada codificada por esta secuencia de codificación modificada tiene la secuencia de aminoácidos de SEQ ID NO: 22.

Se prepararon constructos alternativos que contienen algunas de las modificaciones descritas en el párrafo anterior. Por ejemplo, algunos de estos constructos incluyen una etiqueta de epítipo adicional que es reconocida por el anticuerpo anti-HIV 2F5 (por ejemplo SEQ ID NO: 29 y 31). En otros constructos, la repetición C-héptada de la cepa HIV gp41 HXB2 se reemplazó con la repetición C-héptada del virus Nipah (por ejemplo SEQ ID NO: 32-41). La longitud de la secuencia de longitud héptada se varió en algunos constructos (por ejemplo SEQ ID NO: 27 y 28). Similarmente, el ligador que conecta la secuencia de repetición C-héptada y el resto de la secuencia de codificación de la proteína VP7 se acortó en algunos de los constructos (ver por ejemplo SEQ ID NO: 24-26). Se resumen las diversas variantes en la Tabla 1. La secuencia de proteína VP7 modificada de SEQ ID NO: 22 se incluye como secuencia de referencia. Las secuencias en la Tabla 1 se muestran de la N-terminal a la C-terminal.

Tabla 1

Descripción del constructo	Secuencia
ligador 1 (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a del Factor Xa- ligador 3b de etiqueta de Proteína C (SGG))-VP7 (147-326) (SGG))-VP7 (147-326)	SEQ ID NO: 22
ligador 1 (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a del Factor Xa- ligador 3b de etiqueta de Proteína C (SGG))-VP7 (147-326) (G))-VP7 (147-326)	SEQ ID NO: 23
ligador 1 (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a (SGG)-ligador 3b de etiqueta de Proteína C (SGG))- VP7 (147-326)	SEQ ID NO: 24
ligador 1 (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a (SGG)-ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 25
ligador 1 (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a (G)-ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 26



(Continuado)

	<b>Descripción del constructo</b>	<b>Secuencia</b>
5	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 42 a.a.) (ligador 3a (SGG)-ligador 3b de etiqueta de Proteína C (SGG))- VP7 (147-326)	SEQ ID NO: 27
10	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 42 a.a.) (ligador 3a (G)-ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 28
15	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a (G)-Ligador 3b de epítotope 2F5 de anticuerpo de HIV (SGG)-Ligador 3c de etiqueta de Proteína C (SGG))-VP7 (147-326)	SEQ ID NO: 29
20	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a (G)-Ligador 3b de epítotope 2F5 de anticuerpo de HIV (G)-Ligador 3c de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 30
25	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a (G)-Ligador 3b de epítotope 2F5 de anticuerpo de HIV (GS)-Ligador 3c del Factor Xa (SGG)-Ligador 3d de etiqueta de Proteína C (SGG))-VP7 (147-326)	SEQ ID NO: 31
30	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (K)-Ligador 3b del Factor Xa (SGG)-Ligador 3c de etiqueta de Proteína C (SGG))-VP7 (147-326)	SEQ ID NO: 32
35	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (K)-Ligador 3b del Factor Xa (SGG)-Ligador 3c de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 33
40	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (SGG)-ligador 3b de etiqueta de Proteína C (SGG))- VP7 (147-326)	SEQ ID NO: 34
45	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (SGG)-ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 35
50	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (G)-ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 36

55

60

65

## (Continuado)

5	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) ligador 3a (epítotope 2F5 de anticuerpo HIV (14 a.a.) (SGG)-ligador 3b de etiqueta de Proteína C (SGG))-VP7 (147-326)	SEQ ID NO: 37
10	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) ligador 3a (epítotope 2F5 de anticuerpo HIV (14 a.a.) (G) ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 38
15	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (G)-Ligador 3b (epítotope 2F5 de anticuerpo HIV (9 a.a.) (SGG)-Ligador 3c de etiqueta de Proteína C (SGG))-VP7 (147-326)	SEQ ID NO: 39
20	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (G)-Ligador 3b (epítotope 2F5 de anticuerpo HIV (9 a.a.) (G)-Ligador 3c de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 40
25	ligador I (melitina de abeja) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (G)-Ligador 3b (epítotope 2F5 de anticuerpo HIV (9 a.a.) (GS)-Ligador 3c del Factor Xa (SGG)-Ligador 3d de etiqueta de Proteína C (SGG))-VP7 (147- 326)	SEQ ID NO: 41
30	ligador 1 (htPA) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a del Factor Xa- ligador 3b de etiqueta de Proteína C (SGG))-VP7 (147-326) (SGG)- VP7 (147-326)	SEQ ID NO: 42
35	ligador 1 (htPA) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep ligador 3 (G)-HR2 (HIV, 28 a.a.) (ligador 3a (G)-ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 43
40	ligador 1 (htPA) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (K)-Ligador 3b del Factor Xa (SGG)-Ligador 3c de etiqueta de Proteína C (SGG))-VP7 (147-326)	SEQ ID NO: 44
45	ligador 1 (htPA) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (G)-ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 45
50	ligador 1 (htPA) señal ligador 2 II-TEV-VP7 (51-146) etiqueta (EDSA)-estrep (G)-HR2 ligador 3 (Nipah, 30 a.a.) (ligador 3a (G)-ligador 3b de etiqueta de Proteína C (G))-VP7 (147-326)	SEQ ID NO: 45

55 Los ADNs que consisten de tanto el vector como los genes modificados se generaron por PCR. Los productos de PCR se purificaron en gel y se sometieron a un tratamiento de T4 polinucleótido cinasa para generar extremos fosforilados. La ligación de extremo redondo después se llevó a cabo para el ADN resultante al incubar el ADN con cantidades apropiadas de T4 ligasa durante 2 horas o durante la noche a temperatura ambiente. Los productos de ligación después se crearon con Dpn I durante 30 minutos a 37°C antes de ser usados para transformar las células DH5a a través de los protocolos estándares.

60 Típicamente tres a cuatro colonias se eligieron y se desarrollaron cultivos celulares durante la noche para preparar ADN plásmido usando el Kit Miniprep (QiagenMR). Los plásmidos se examinaron por electroforesis en gel de agarosa y se confirmaron las secuencias correctas por secuenciación de ADN. Los plásmidos de las secuencias correctas se usaron para transformar las células competentes DHIOBac a través de los procedimientos estándares. 65 Dos a tres colonias blancas se seleccionaron para cada constructo y se desarrollaron cultivos celulares durante la

noche para la extracción del ADN bÁcrido recombinante por precipitación de isopropanol/etanol (Solución I: Tris 15 mM, pH 8.0, EDTA 10 mM, y 100 µg/mL de RNasa A; Solución II: NaOH 0.2 M y 1% de SDA; y Solución III: acetato de potasio 3 M, pH 5.5; todos esterilizados por filtro). Los bÁcridos purificados se examinaron por PCR usando los cebadores MI3 y los ADN correctos se usaron para transfectar las monocapas de las células sf9 en placas de 6 pocillos, cada pocillo sembrado con 1 millón de células. Los virus P1 se recolectaron 5 días postransfección y los virus P2 se produjeron al infectar las células sf9 (densidad de 1.5-2 millón/mL) con 0.05-0.1% de virus P1. Los virus P2 se cosecharon 5-7 días pos-infección y se usaron para la expresión de la proteína.

Los constructos de proteínas VP7 modificados descritos en lo anterior se pueden usar para mostrar las proteínas heterólogas formadoras de trímeros con la proteína heteróloga formadora de trímeros correspondientemente modificada (ver Figura 1B para una vista general esquemática de las modificaciones adecuadas). Como se muestra en la Figura 2, las secuencias de repetición héptadas HR2 de la proteína VP7 modificada y las secuencias de repetición héptadas HR1 de la proteína heteróloga formadora de trímeros correspondientemente modificada forman un complejo estable a través de un atado de seis hélices de las cuales se diseñó en el extremo 5' del codón de inicio antes de la proteína heteróloga en la proteína VP7 de rotavirus. La proteína de hemaglutinina de influenza A (HA) y la proteína de fusión gp140 de HIV-1 se eligieron como ejemplos para las proteínas heterólogas formadoras de trímeros.

El gen HA del virus de influenza A H1N1 Solomon Islands 2006 se clonó en un vector pFastBac LICMR (Life TechnologiesMR) por medio del método de clonación independiente de ligación (LIC). El vector pFastBac LICMR se creó al insertar un sitio LIC en un vector pFastBacI (Invitrogen). Una secuencia Kozak (SEQ ID NO: 19) se diseñó en el extremo 5' del codón de inicio antes de la secuencia de codificación. La modificación en la secuencia de codificación del gen HA se introdujo por PCR invertida.

La proteína final codificada por la secuencia de codificación HA modificada tiene las siguientes características: (a) los primeros 38 aminoácidos codifican el péptido señal Baculoviral gp64 para la expresión de la proteína modificada en las células de insecto. El péptido señal de la proteína HA se ha removido. Otra secuencia señal se puede usar como apropiadas, tal como la secuencia señal consenso de HIV o el péptido señal de la tPA humana para la expresión en las células humanas. (b) Residuos de aminoácidos 42 a 47 forman una etiqueta de afinidad opcional (His-tag más un ligador para una escisión de péptido de señal más eficiente para propósitos de purificación de proteínas y se puede reemplazar por cualquier otra etiqueta de afinidad, tal como etiqueta HA, etiqueta FLAG, etc. (c) Los residuos de aminoácidos 48 a 54 forman una secuencia de reconocimiento de proteasa TEV para la escisión de etiqueta de afinidad. Una secuencia ligadora también se incluye. La secuencia de reconocimiento de proteasa TEV se puede reemplazar por la secuencia de reconocimiento de cualquier otra proteasa, tal como la proteasa PreScission™ (Le. Proteasa de Rhinovirus 3C, GE HealthcareMR, Life Sciences), factor Xa, enterocinasa, trombina, furina, etc. (d) Residuos de aminoácidos 55 a 381 son las secuencias de la porción HA1 de la hemaglutinina de la cepa Solomon Islands 2006 de influenza A, que se podría modificar/mutar como sea apropiado o ser reemplazado por aquel de cualquier otra de las cepas de influenza como se muestra a continuación. (e) Residuos de aminoácidos 382 a 386 son la secuencia de reconocimiento de la enterocinasa diseñada entre HA1 y HA2 para propósitos de escisión. Este sitio es opcional y se puede reemplazar por la secuencia de reconocimiento del Factor Xa, proteasa TEV, u otros como sea apropiado. (f) Residuos de aminoácidos 387 a 565 son la secuencia de la porción HA2 de HA. (g) Residuos de aminoácidos 566 a 595 son una etiqueta de trimerización de la fibrina del bacteriófago T4 (Foldon), que es opcional y se puede reemplazar por cualquier otra etiqueta de trimerización o secuencias ligadoras. (h) Residuos de aminoácidos 596 a 629 son parte de la N-héptada de la cepa HIV gp41 HXB2 y se puede reemplazar por otras secuencias de repetición héptadas adecuadas por ejemplo de cualquier otro retrovirus, paramixovirus, etc. La numeración de las secuencias de aminoácidos HA usadas en la descripción de la secuencia de codificación HA modificada es de acuerdo con GenBank ID ABU50586.1 (incorporada en la presente a manera de referencia en la versión disponible en la fecha de la presentación de la solicitud de prioridad). La secuencia de aminoácidos de ese constructo se muestra en SEQ ID NO: 46.

Se prepararon constructos alternativos que contienen algunas de las modificaciones descritas en el párrafo anterior. Por ejemplo, el dominio de trimerización se omitió de los constructos alternativos (por ejemplo SEQ ID NO: 48 y 49). En algunos constructos, la secuencia de reconocimiento de enterocinasa se reemplazó por la secuencia de reconocimiento de Factor Xa (por ejemplo SEQ ID NO: 52). Este constructo tampoco incluyó un ligador entre la secuencia de codificación HA y la secuencia de repetición C-héptada. En otros constructos, tanto el dominio de trimerización como la secuencia de reconocimiento de enterocinasas se omitieron (SEQ ID NO: 51). La repetición N-héptada del virus Nipah se usó algunas veces en lugar de la repetición N-héptada de la cepa HIV gp41 HXB2 (ver por ejemplo SEQ ID NO: 47 y 49). La longitud de la secuencia de repetición héptada varió en algunos constructos (por ejemplo SEQ ID NO: 54, 56 y 58). En otros constructos, la longitud del ligador que conecta la secuencia de codificación HA a la secuencia de repetición N-héptada varió (ver por ejemplo SEQ ID NO: 56 y 58). Los genes HA modificados de H3 Wisconsin 2005 (ver SEQ ID NOs: 60 y 61) y H5 Vietnam 2004 (ver 15 SEQ ID NOs: 62 y 63) se optimizaron con codón y se sintetizaron por GeneArtMR (Life TechnologiesMR) antes de que se incluyera en los constructos modificados. La numeración de las secuencias de aminoácidos HA usadas en las descripciones de estos constructos es de acuerdo con GenBank IDs AAT73274.1 y ACV49644.1 (ambos incorporados en la presente a

manera de referencia en versiones disponibles en la fecha de la presentación de la solicitud de pérdida), respectivamente. La estructura de estas secuencias de codificación HA modificadas se resume en la Tabla 2. La secuencia de proteína HA modificadas de SEQ ID NO: 46 se excluye como la secuencia de referencia. Las secuencias en la Tabla 2 se muestran de N-terminal a la C-terminal.

5

**Tabla 2**

	<b>Descripción del constructo</b>	<b>Secuencia</b>
10	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-ligador 3 enterocinasa (DDDDK)-HA (344-519, es decir HA2) (RSL)-Foldon (Pequeña cantidad de Fibrina T4 C-terminal)-HRI (HIV, 34 a.a.)	SEQ ID NO: 46
15	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-ligador 3 enterocinasa (DDDDK)-HA (344-519, es decir HA2) (RSL)-Foldon (Pequeña cantidad de Fibrina T4 C-terminal)-HRI (Nipah, 34 a.a.)	SEQ ID NO: 47
20	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-enterocinasa (DDDDK)-HA (344-519, es decir HA2)- HR1 (HIV, 34 a.a.)	SEQ ID NO: 48
25	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-ligador 3 enterocinasa (DDDDK)-HA (344-519, es decir HA2) (RS)-HRI (Nipah, 34 a.a.)	SEQ ID NO: 49
30	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-519)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 50
35	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV ligador 3 (GYLLE)-HA (18-519)(RS)-HRI (Nipah, 34 a.a.)	SEQ ID NO: 51
40	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-Factor Xa (IEGR)-HA (344-519, es decir HA2)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 52
45	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-Factor Xa (IEGR)-HA ligador 3 (344-519, es decir HA2) (RS) HR1 (Nipah, 34 a.a.)	SEQ ID NO: 53
50	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-Enterocinasa (DDDDK)-HA ligador 3 (344-511, es decir HA2) (IGE)-HRI (HIV, 39 a.a.)	SEQ ID NO: 54
55	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-Enterocinasa (DDDDK)-HA ligador 3 (344-511, es decir HA2) (IGEARQ)-HRI (Nipah, 34 a.a.)	SEQ ID NO: 55
60	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-Ligador 3 enterocinasa (DDDDK)-HA (344-519, es decir HA2) (RSI)-HRI (HIV, 38 a.a.)	SEQ ID NO: 56
65	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-Ligador 3 enterocinasa (DDDDK)-HA (344-519, es decir HA2) (RSIRQ)-HRI (Nipah, 34 a.a.)	SEQ ID NO: 57

(Continuado)

5	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-Ligador 3 enterocinasa (DDDDK)-HA (344-519, es decir HA2) (RSIKKLIGE)-HRI (HIV, 39 a.a.)	SEQ ID NO: 58
10	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (18-339, es decir HAI)-Ligador 3 enterocinasa (DDDDK)-HA (344-519, es decir HA2) (RSIKKLIGEARQ)-HRI (HIV, 34 a.a.)	SEQ ID NO: 59
15	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (17-341, es decir HAI)-Enterocinasa (DDDDK)-HA ligador 3 (347-522, es decir HA2) (RSL)-HRI (HIV, 38 a.a.)	SEQ ID NO: 60
20	ligador 1 señal (péptido señal gp64 de Baculovirus) ligador 2 (ADP)-6xHis etiqueta-TEV (GYLLE)-HA (17-341, es decir HAI)-Enterocinasa (DDDDK)-HA ligador 3 (347-522, es decir HA2) (RSL)-Foldon (Pequeña cantidad de fibritina T4 C-terminal)-HRI (HIV, 35 a.a.)	SEQ ID NO: 61
25	ligador señal (péptido señal gp64 de Baculovirus)-HA (17-20)-etiqueta 6xHis-TEV-HA (21-345, es decir HA1 )-Enterocinasa (DDDDK)-HA (346-521, es decir HA2)(RSL)-HRI (HIV, 38 a.a.)	SEQ ID NO: 62
30	ligador señal (péptido señal gp64 de Baculovirus)-HA (17-20)-etiqueta 6xHis-TEV-HA (21-345, es decir HA1 )-Enterocinasa (DDDDK)-HA (346-521, es decir HA2)(RSL)-Foldon (Pequeña cantidad de fibritina T4 C-terminal)-HRI (HIV, 35 a.a.)	SEQ ID NO: 63

35 Cada uno de los tres genes HA modificados se subclonaron en pFastBacDualMR (Life TechnologiesMR) entre los sitios Sal I y Not I. Los baculovirus recombinantes que llevan los genes HA modificados se crearon siguiendo los procedimientos similares como aquellos descritos para VP7.

40 Los genes gpl40 modificados del serotipo HIV-1 clado A 1992 Uganda 037.8 y clado C 1997 se utilizaron con codón y se sintetizaron por GeneArtMR (Life TechnologiesMR). Los genes modificados se subclonaron en pFastBacDualMR (Life TechnologiesMR) entre los sitios Sal I y Not I.

45 La proteína final codificada por la secuencia de codificación gpl40 modificada de la cepa HIV-1 clado A tiene las siguientes características: (a) Los primeros 21 aminoácidos que codifican el péptido señal de melitina de abeja para la expresión de la proteína modificada en las células de insecto. El péptido señal gpl40 se ha removido. Otra secuencia señal se pueden usar como apropiadas, tal como la secuencia señal de consenso HIV o el péptido señal de tPA humana para la expresión en las células de mamífero/humanas. (b) Residuos de aminoácidos 22 a 29 forman una etiqueta de afinidad opcional (etiqueta His más un ligador para la escisión de péptido señal más eficiente) para propósitos de purificación de proteínas y se pueden reemplazar por cualquier otras etiquetas de afinidad, tales como etiqueta HA, etiqueta FLAG, etc. (c) Residuos de aminoácidos 30 a 36 forman una secuencia de reconocimiento de proteasa TEV para la escisión de la etiqueta de afinidad y se pueden reemplazar por las secuencias de reconocimiento de cualquier otra proteasa, tal como la proteasa PreScissionMR (es decir proteasa de Rhinovirus 3C, GE HealthcareMR, Life Sciences), factor Xa, enterocinasa, trombina, furina, etc. (d) Residuos de aminoácidos 37 a 685 en la secuencia de codificación gpl40 de la cepa de VIH 1992 Uganda 037.8 que se podría modificar/mutar como sea apropiado o se reemplace por aquella de las otras cepas HIV/SIV. (e) Residuos de aminoácidos 686 a 694 son un ligador y se pueden reemplazar por cualquiera de otras secuencias ligadoras apropiadas. (f) Residuos de aminoácidos 695 a 721 son una etiqueta de trimerización de fibritina T4 de bacteriófago (Foldon), que es opcional o se puede reemplazar por cualquier otra etiqueta de trimerización o secuencias ligadoras. (g) Residuos de aminoácidos 722 a 755 son parte de la repetición N-héptada de la cepa HIV gp41 HXB2 y se puede reemplazar por otras secuencias de repetición héptadas adecuadas por ejemplo de cualquier otros retrovirus, 25 paramixovirus, etc. La secuencia de aminoácidos de este constructo se muestra en SEQ ID NO: 64.

65 Se prepararon constructos alternativos que contienen algunas de las modificaciones posibles indicadas en el párrafo anterior. Por ejemplo, algunos de los constructos no incluyen la etiqueta de trimerización, y la secuencia ligadora entre la región de codificación gpl40 y la repetición N-héptada se ha cortado (por ejemplo SEQ ID NO: 66 y

67). Otros constructos son adaptados para la expresión en las células de mamífero por el reemplazo del péptido señal (por ejemplo SEQ ID NO: 65 y 67). En algunos constructos, la repetición N-héptada de la cepa HIV gp41 HXB2 se ha reemplazado con la repetición N-héptada del virus Nipah (SEQ ID NO: 68-71 y 76-79), en otros, la secuencia de codificación gpl40 del serotipo HIV-1 clado A 1992 Uganda 037.8 se ha reemplazado con la secuencia de codificación gpl40 de HIV-1 clado C 1997 (SEQ ID NO: 72-79). La numeración para clado A y C gpl40 es de acuerdo con GenBank ID AAB05027.1 y AF286227.1 (ambos incorporados en la presente a manera de referencia en las versiones disponibles en la fecha de la presentación de la solicitud de prioridad), respectivamente.

Las diversas variantes se resumen en la Tabla 3. La secuencia de proteína gpl40 modificada de SEQ ID NO: 64 también se incluye como secuencia de referencia. Las secuencias en la Tabla 3 se muestran de N-terminal a la C-terminal.

**Tabla 3**

Descripción del constructo	Secuencia
ligador 1 señal (péptido señal de melitina de abeja) (ED)-6xHis etiqueta-TEV-Ligador 2 HIV clado A gpl40 (27-675) (ligador 2a (SR)ligador 2b del Factor Xa (GSG))-Foldon (Pequeña cantidad de fibritina T4 C-terminal)-HRI (HIV, 34 a.a.)	SEQ ID NO: 64
ligador 1 señal (péptido señal consenso HIV) (ED)-6xHis etiqueta-TEV-Ligador 2 HIV clado A gpl40 (27-675) (ligador 2a (SR)ligador 2b del Factor Xa (GSG))-Foldon (Pequeña cantidad de fibritina T4 C-terminal)-HRI (HIV, 34 a.a.)	SEQ ID NO: 65
ligador 1 señal (péptido señal de melitina de abeja) (ED)-6xHis etiqueta-TEV-Ligador 2 HIV clado A gpl40 (27-675) (GSG)-HRI (HIV, 34 a.a.)	SEQ ID NO: 66
ligador 1 señal (péptido señal consenso HIV) (ED)-6xHis etiqueta-TEV-Ligador 2 HIV clado A gpl40 (27-675) (GSG)-HRI (HIV, 34 a.a.)	SEQ ID NO: 67
ligador 1 señal (péptido señal de melitina de abeja) (ED)-6xHis etiqueta-TEV-Ligador 2 HIV clado A gpl40 (27-675) (ligador 2a (SR)ligador 2b del Factor Xa (GSG))-Foldon (Pequeña cantidad de fibritina T4 C-terminal)-HRI (Nipah, 32 a.a.)	SEQ ID NO: 68
ligador 1 señal (péptido señal consenso HIV) (ED)-6xHis etiqueta-TEV-Ligador 2 HIV clado A gpl40 (27-675) (ligador 2a (SR)ligador 2b del Factor Xa (GSG))-Foldon (Pequeña cantidad de fibritina T4 C-terminal)-HRI (Nipah, 32 a.a.)	SEQ ID NO: 69
ligador 1 señal (péptido señal de melitina de abeja) (ED)-6xHis etiqueta-TEV-Ligador 2 HIV clado A gpl40 (27-675) (S)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 70
ligador 1 señal (péptido señal consenso HIV) (ED)-6xHis etiqueta-TEV-Ligador 2 HIV clado A gpl40 (27-675) (S)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 71
ligador 1 señal (péptido señal de melitina de abeja) (ED)-6xHis etiqueta-TEV-Linker2 (AENLWV)-HIV clado C gpl40 (31-667)-linker3 (S)-Foldon (Pequeña cantidad de fibritina T4 C-terminal)-HRI (HIV, 34 a.a.)	SEQ ID NO: 72

(continuado)

5	ligador 1 señal (péptido señal consenso HIV) (ED)-6xHis etiqueta-TEV-Linker2 (AENLWV)-HIV clado C gp140 (31-667)-linker3 (S)-Foldon (Pequeña cantidad de fibrina T4 C-terminal)-HRI (HIV, 34 a.a.)	SEQ ID NO: 73
10	ligador 1 señal (péptido señal de melitina de abeja) (ED)-6xHis etiqueta-TEV-HIV clado C gp140 (31-667)-linker2 (SGI)-HRI (HIV, 34 a.a.)	SEQ ID NO: 74
15	ligador 1 señal (péptido señal consenso HIV) (ED)-6xHis etiqueta-TEV-HIV clado C gp140 (31-667)-linker2 (SGI)-HRI (HIV, 34 a.a.)	SEQ ID NO: 75
20	ligador 1 señal (péptido señal de melitina de abeja) (ED)-6xHis etiqueta-TEV-Linker2 (AENLWV)-HIV clado C gp140 (31-667)-linker3 (S)-Foldon (Pequeña cantidad de fibrina T4 C-terminal)-HRI (Nipah, 32 a.a.)	SEQ ID NO: 76
25	ligador 1 señal (péptido señal consenso HIV) (ED)-6xHis etiqueta-TEV-Linker2 (AENLWV)-HIV clado C gp140 (31-667)-linker3 (S)-Foldon (Pequeña cantidad de fibrina T4 C-terminal)-HRI (Nipah, 32 a.a.)	SEQ ID NO: 77
30	ligador 1 señal (péptido señal de melitina de abeja) (ED)-6xHis etiqueta-TEV-Linker2 (AENLWV)-HIV clado C gp140 (31-667)-linker3 (S)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 78
35	ligador 1 señal (péptido señal consenso HIV) (ED)-6xHis etiqueta-TEV-linker2 (AENLWV)-HIV clado C gp140 (31-667)-linker3 (S)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 79

40

**Ejemplo 2: Preparación de complejos VP7-HA estables**

Para producir los complejos HA-VP7, células Hi5 en la densidad de  $2 \times 10^6$ /mL se infectaron con 0.5% de baculovirus VP7 y 1.0% de baculovirus HA en presencia de 5.0% de FBS inactivado con calor (Sigma-Aldrich). El medio se recolectó 5 días pos-infección al centrifugar las células a  $4000 \times g$  durante 45 minutos. El sobrenadante se diafiltró contra 8 litros de 1x TNC (Tris 20 mM, pH 8.0, NaCl 100 mM, y  $\text{CaCl}_2$  1.0 mM; se complementó con 0.02% de azida de sodio) usando un filtro de corte de línea de 10 KDa en un Sistema de Filtración de Flujo Tangencial Cogent M (Millipore). Las muestras intercambiadas con solución amortiguadora se complementaron con PMSF 1.0 mM y se clarificaron por centrifugación durante 1 hora a 10,000 RPM en un rotor JA10 (BeckmanMR). Los complejos de proteínas se purificaron del sobrenadante por una columna StrepTactinMR (IBAMR), seguido por una columna NiNTA (QiagenMR), y una columna Superose 6 (AmershamMR).

Los complejos HA-VP7 purificados se trataron con 0.002% (p/p) de enterocinasa a 4°C durante 48 horas. La escisión de HA0 en HA1 y HA2 se examinó por SDS-PAGE. La enterocinasa se inactivó al agregar 1x de un comprimido de inhibidor de proteasa completo sin EDTA. La muestra se complementó con una solución amortiguadora de reducción oxidación (10:1 de relación molar de GSH y GSSG) en una concentración final de 0.2-0.5 mM y se trataron con proteasa TEV (1/50-200, p/p) ya sea a temperatura ambiente durante 4 horas o a 4°C durante la noche. La remoción de las etiquetas también se examinó por SDS-PAGE.

**Ejemplo 3: Preparación de complejos HIV gp140-VP7 estables**

Los genes VP7 y gp140 modificados que codifican las proteínas de SEQ ID NOs: 22, 26, 32 y 36 y SEQ ID NOs: 64, 66, 68 y 70, respectivamente se subclonaron en vectores pVRC8400 y pYRC-IRES-Puro entre Sal I y Not I para la expresión en las células de mamífero. El péptido señal para el gen YP7 se cambió a un péptido señal tPA humano y el péptido señal para el gen gp140 se cambió a la secuencia de péptido de señal consenso VIH. Las

proteínas VP7 modificadas resultantes y las proteínas gpl40 modificadas tienen las secuencias de aminoácidos de SEQ ID NOs: 42-45 y SEQ ID NOs: 65, 67, 69 y 71, respectivamente. La co-expresión y formación en complejo se sometieron a prueba para las siguientes combinaciones de la proteína VP7 modificada y la proteína gpl40 modificada: SEQ ID NOs. 42 + 65, SEQ ID NOs. 42 + 67, SEQ ID NOs. 43 + 65, SEQ ID NOs. 43 + 67, SEQ ID NOs. 44 + 69, SEQ ID NOs. 44 + 71, SEQ ID NOs. 45 + 69, SEQ ID NOs. 45 + 71.

Los complejos HIV gpl40-VP7 se expresaron al co-transfectar células 293T con un plásmido que aloja el gen gpl40-HRI y otro que aloja el gen VP7-HR2. La transfección transitoria se hizo al seguir los protocolos estándares usando polietilenimina (PEI, 25KDa, lineal o ramificada). Para el mayor rendimiento de proteínas, se seleccionaron líneas de células estables contra puromicina. En resumen, 70% de células T 293 confluentes se transfectaron con una mezcla de LipofectoamineMR 2000 (InvitrogenMR) y la versión pVRC-IRES-puro de los dos plásmidos (relación en masa 2:1 de lipofectamina al ADN total) siguiendo los procedimientos estándares. A las 20 horas pos-transfección, el medio se reemplazó con DMEM (GibcoMR) complementado con 10% de Suero Bovino Fetal (Atlas), 1% de GlutaMAXMR (GibcoMR), 100 unidades/mL de penicilina/estreptomicina (GibcoMR), y 2 µg/mL de puromicina (GibcoMR). El medio se mantuvo fresco al cambiarlo cada tres a cuatro días hasta que se forman colonias (visible a simple vista). Las colonias individuales se eligieron y se cultivaron en placas de 24 pocillos para las pruebas de expresión de escala pequeña. Las colonias que alojan tanto genes se seleccionaron basado en la detección de ambas proteínas en el análisis de western blot. La colonia con el rendimiento más alto del complejo se aumentó y se guardó como solución madre en nitrógeno líquido. Las células se aumentaron en un medio que contiene puromicina y la expresión de la proteína se podría llevar a cabo usando ya sea cultivos de células adherentes o de suspensión en un medio 293 Freestyle (GibcoMR) complementado con 5% de FBS inactivado con calor (Sigma-AldrichMR). El medio se recolectó después de 5 días al centrifugar las células a 4000xg durante 45 minutos. La diafiltración y purificación de proteínas se llevó a cabo siguiendo los mismos procedimientos como se describe para los complejos HA-VP7.

#### Ejemplo 4: Preparación de Fabs y Fvs

El anticuerpo monoclonal CR6261 reconoce una región helicoidal sumamente conservada en el tallo próximo a la membrana de la proteína HA de influenza [67]. Los factores de estructura y coordenadas para el fragmento CR6261 sHgL Fab se depositaron con el banco de datos de proteínas bajo el acceso 4EVN (incorporado en la presente a manera de referencia en la versión disponible en la fecha de la presentación de la solicitud de prioridad) y se ha descrito previamente [68].

La secuencia de codificación para el fragmento de ligación a antígeno (Fab) y el fragmento variable de cadena individual (scFv) del anticuerpo monoclonal CR6261 se clonaron en el vector pVRC-IRES-Puro entre Sal I y Not I siguiendo los procedimientos estándares como se describe en lo anterior. El péptido señal tPA humano se usó para la secreción eficiente en estos constructos en los sistemas de expresión de mamífero. Las secuencias de aminoácidos de los constructos finales se muestran en SEQ ID NOs: 81-83. Un scFv del anticuerpo monoclonal CR6261 para la expresión en las células bacterianas también se preparó (ver SEQ ID NO: 80).

Los Fabs se expresaron por transfección transitoria de las células 293T en botellas giratorias. Para cada botella giratoria, 125 µg de cada uno de ADN plásmido purificado que aloja los genes de cadena pesada y ligeras se mezclaron en 12.5 mL de DMEM no suplementado (GibcoMR) y se incubaron en la vitrina durante 15 minutos. El PEI (500µg, lineal o ramificado) también se diluyeron en 12.5 mL de DMEM no suplementado y se incubaron en la vitrina durante 15 minutos antes de ser agregados gota a gota y ser mezclados bien en la mezcla de ADN. El total de 25 mL de la mezcla se incubó en la vitrina durante 15 minutos extras para permitir un complejo de ADN-PEI para formarse antes de ser agregados a las células a 50-70% de confluencia en las botellas giratorias. Las células se incubaron por más de 5 horas con el complejo de DNA-PEI a 37°C, y el medio se reemplazó por 250 mL del medio FreeStyleMR (GibcoMR) complementado con 100 unidades/mL de penicilina/estreptomicina. El medio se recolectó 5 días pos-transfección, y los Fabs se purificaron usando una resina de afinidad CaptureSelectMR Le Kappa o Le Lambda, seguido por cromatografía de exclusión de tamaño en una columna S200 (AmershamMR).

Los ScFvs se expresaron ya sea en las células 293 T como proteínas secretadas o en el E. coli como cuerpos de inclusión, de los cuales la proteína después se extrajo y se volvió a replegar. Los mismos procedimientos de transfección transitoria como se describe en lo anterior se siguieron para la expresión de células de mamífero. Para la expresión de E. coli, la expresión de proteínas se indujo durante aproximadamente 5 horas usando IPTG 1 mM en una densidad celular de 0.6~0.8 O.D.600nm. Las células se recolectaron, se lavaron con 1x de PBS, y se lisaron por sonicación sobre hielo. Los cuerpos de inclusión se extrajeron después de la remoción de las fracciones solubles de las células por centrifugación durante 15 minutos a 20000xg. Los cuerpos de inclusión se disolvieron en Tris 100 mM, pH 7.5, urea 8 M, y β-mercaptoetanol 10 mM, se clarificaron por centrifugación (20 minutos a 40000xg), y sobrenadantes se purificó sobre una columna de NiNTA. La muestra eluida, también en la forma desnaturaliza, se diluyó en la solución amortiguadora de replegamiento (Tris 100 mM pH 7.5, arginina 1 M, NaCl 500 mM, 10% de glicerol, y EDTA 1 mM) gota a gota a 4°C a una concentración de proteínas final menor que 0.1 mg/mL. La muestra replegada se dializó contra Tris 20 mM, pH 7.5, NaCl 100 mM cuatro veces, cada vez 20 volúmenes. La muestra dializada después se clarificó a centrifugación a 10,000 rpm en un rotor JA10 (BeckmanMR) y el



sobrenadante se hizo pasar a través de una columna NINTA para concentrar la proteína. Las fracciones eluidas después se combinaron, se concentraron y se purificaron adicionalmente sobre una columna de exclusión de tamaño S200 (AmershamMR).

## 5 Ejemplo 5: Preparación de DLP

El serotipo de rotavirus rhesus G3 se amplificó al infectar células MA-104. Brevemente, las células MA104 se desarrollaron en un medio M199 (GibcoMR) complementado con 10% de suero bovino fetal (HyCloneMR Laboratories), HEPES 10 mM, pH 7.0, L-glutamina 2 mM, y 100 unidades/mL penicilina/estreptomicina. Las células se mantuvieron y se aumentaron a una cámara de cultivo de 10 células apiladas (CorningMR) o botellas giratorias. Cuando las células fueron confluentes, el medio se reemplazó por MA199 sin suero complementado con 1 µg/mL de tripsina pancreática porcina (Sigma-AldrichMR) para inoculación y amplificación del rotavirus. El medio de las células MA104 infectadas se recolectó a aproximadamente 36-42 horas pos-infección y se almacenó a -80°C para uso futuro.

El medio congelado de las células MA104 infectadas se descongeló a 4°C durante la noche. Los restos celulares se eliminaron por centrifugación a baja velocidad a  $x$  g durante aproximadamente 10 minutos. El sobrenadante resultante se filtró a través de papel de Filtro WhatmanMR para remover los restos celulares residuales antes de que se hiciera pasar adicionalmente a través de una unidad de filtro Expressplus 0.45 µm (MilliporeMR) bajo vacío. Las partículas del virus después se formaron en pelotillas a 45,000 rpm durante 1 hora a 4°C en un rotor 45Ti (BeckmanMR).

La pelotilla se volvió a suspender en un total de 10 mL de solución amortiguadora 1x TNE (Tris 20 mM, pH 8.0, NaCl 100 mM, y EDTA 1 mM), se hizo pasar por ondas sonoras brevemente, y se extrajo dos veces con Freon 113 (Sigma-AldrichMR). La fase acuosa se recubrió y se concentró en aproximadamente 1 mL en un dispositivo de filtración de centrifuga AmiconMR (línea de corte 100 KDa) a 3000xg durante aproximadamente 10 minutos. La muestra concentrada se volvió a suspender por pipeteo arriba y abajo un par de veces y se recubrió sobre un gradiente CsCl preformado en 1x TNE (1.26 a 1.45 g/mL de densidad como es determinado por la refractometría). Las muestras se centrifugaron a 55000 rpm en un rotor SW 55Ti (BeckmanMR) a 4°C durante aproximadamente 2 horas. La banda DLP se recolectó y se dializó durante la noche a 4°C contra solución amortiguadora 1x TN (Tris 20 mM, pH 8.0, NaCl 100 mM) complementado con 0.02% de azida de sodio.

Las DLPs se tiñeron negativamente y se visualizaron usando microscopia electrónica (EM; Figura 3A).

## 35 Ejemplo 6: Recubrimiento de las DLPs con el complejo HA-VP7

Las DLPs obtenidas en el Ejemplo 5 y la proteína VP7 de rotavirus modificada preparada en el Ejemplo 1 se mezclaron y se incubaron a 4°C durante al menos 1 hora en presencia de 5-10 mM de CaCl<sub>2</sub>. Las partículas resultantes se tiñeron negativamente y se visualizaron usando EM. Una capa de VP7 individual que cubre cada partícula se observó confirmando que la proteína VP7 modificada fue capaz de recubrir las DLPs (Figura 3B).

El recubrimiento de las DLPs con el complejo HA-VP7 obtenidas en el Ejemplo 2 se llevó a cabo a 4°C durante al menos 1 hora al mezclarlas en una relación molar de 1:1.5 a 1:2 en presencia de CaCl<sub>2</sub> 5-10 mM. Las partículas de rotavirus recubiertas se tiñeron negativamente y se visualizaron usando EM. Además de la capa de VP7 individual que cubre la DLP, se observó una capa adicional que circunda las partículas recubiertas sugiriendo que el complejo HA-VP7 se ensambló correctamente sobre la superficie del DLPs (Figura 3C).

Para confirmar que la capa adicional se formó por la proteína HA, las partículas recubiertas se cargaron sobre un gradiente CsCl preformado en 1x TNC (densidad 1.25 g/mL a 1.45 g/mL) y se centrifugó en un rotor SW60 (BeckmanMR) a 58,000 rpm durante 2 horas. Las bandas resultantes se recolectaron, se dializaron contra 1x TNC, y se examinaron por SDS-PAGE. Como un control, DLPs, el complejo HA-VP7 y la mezcla de ambos como se usan en la reacción de recubrimiento también se cargaron en el gel en líneas separadas. Después de la separación, las bandas separadas que corresponden a las proteínas VP2 y VP6 de las DLPs, la proteína VP7 y la proteína HA se observaron en la estequiometría correcta en la muestra que correspondió a las partículas recubiertas (Figuras 4A y 4B). Esto confirmó que la capa adicional que circunda las partículas recubiertas se compuso de la proteína HA en ocupación completa.

## Ejemplo 7: Reconstrucción de la imagen tridimensional del complejo HA-VP7-anticuerpo

La crio-EM de alta resolución tiene el potencial de reemplazar la cristalografía de rayos X como el método de elección para analizar grandes complejos moleculares, en particular en áreas donde la elucidación estructural de un gran número de complejos moleculares es necesaria. Una de tal área es diseño de vacuna de influenza, donde la deriva antigénica conduce a rápidos cambios en de la secuencia del antígeno de superficie inmunodominante, la proteína HA. Los anticuerpos monoclonales ampliamente protectores se han identificado previamente y muchos de estos anticuerpos reconocen epítomos conservados que no se someten a un cambio rápido. Para investigar su

idoneidad como agentes protectores contra las cepas de influenza que surgen recientemente, los inventores necesitan clasificar estos anticuerpos anti-HA ampliamente neutralizantes contra diferentes variantes de la proteína HA.

5 Se probó la factibilidad de tal procedimiento usando el análisis crio-EM. Las DLPs se recubrieron con la proteína VP7 modificada sola o con un complejo HA-VP7 como se describe en el Ejemplo 6. Las partículas recubiertas con el anticuerpo HA-VP7 se obtuvieron al mezclar los complejos HA-VP7 purificados con las Fabs o scFvs del anticuerpo monoclonal CR6261 preparado en el Ejemplo 4 en una relación molar de 1:1.5~2.0. La mezcla se incubó a 4°C durante 30 minutos. Los complejos se purificaron al cargar la mezcla después del final del período de incubación en una columna Superose 6. Los complejos purificados después se usaron para recubrir las DLPs siguiendo los mismos procedimientos como se describe en el Ejemplo 6. Alternativamente, los fragmentos de anticuerpos se agregaron directamente a las partículas recubiertas con los complejos HA-VP7 en una relación molar de 1:1.2~1.5 ya sea antes o después del paso de purificación en gradiente CsCl.

15 Las crio-rejillas se prepararon con un embolo Automatizado VitrobotMR (FEI). Rejillas de carbono Quantifoil Holey se les aplicó una descarga luminiscente y se dejaron a temperatura ambiente durante la noche antes del uso. Cada rejilla, 4 µL de muestra en una concentración equivalente a ~5 mg/mL de DLPs de rotavirus se aplicó a una superficie lateral de la rejilla. Durante la inmersión, la humedad de la cámara se mantuvo cercana 100% y la temperatura aproximadamente 22°C. La rejilla después se secó durante 4 segundos de ambos lados con papel de filtro, seguido inmediatamente por inmersión en etano líquido para vitrificación. Las rejillas después se almacenaron en nitrógeno líquido antes de ser usadas para recolección de datos.

25 Los datos se recolectaron en un microscopio electrónico TecnaiMR F30 (FEI) operado a 300 kV. El sistema óptimo se alineó usando procedimientos estándares (cambios de haz, inclinación del haz, altura eucéntrica, puntos de pivote, centro rotacional, astigmatismo, etc.). Los procedimientos de adquisición de imagen implicaron el descubrimiento del área de imagen deseada, ajustar el desenfoque (entre 0.6 y 3.0 µm), probar la velocidad de deriva de la rejilla (menor que 2 Å por segundo), y la exposición de imagen final. Estos procedimientos se lograron semiautomáticamente usando el programa SerialEM (Crio-electron Microscopy Facility, University of Colorado - Boulder). Los datos se registraron como películas en una cámara de detección directa K2 SummitMR (GatanMR) usando un modo de súper resolución (tamaño de pixel: 0.99 Å). Se usó el protocolo "de película" de la referencia 24. La velocidad de dosis en la muestra fue ~3 electrones/Å<sup>2</sup> por segundo. Cada marco registró una exposición de 0.5 segundos y la película final consistió de 24 marcos con una dosis acumulada de ~36 electrones/Å<sup>2</sup>.

35 Los marcos de cada película se alinearon usando los caracteres IMOD (Crio-electron Microscopy Facility, University of Colorado - Boulder) basado en la correlación cruzada de imagen. Los marcos alineados se promediaron simplemente para procesamiento de imagen inicial. Ya que la resolución se mejoró posteriormente sin embargo, las imágenes promediadas de los primeros 13 marcos (o la mejor serie de marcos) de cada película, que correspondería a ~20 de dosis de electrones, fueron usados. Las imágenes de partículas se eligieron usando e2boxer.py de la suite de procesamiento de imagen EMAN2 [69]. Las imágenes con defectos obvios, aberraciones, enfoque normal, contaminación, solapamiento, o derivaciones de muestra grandes se excluyeron manualmente después de la inspección visual. Las coordenadas de las partículas se usaron para escindir las pilas de imágenes de las partículas individuales con dimensiones de 1600 x 1600 pixeles usando proc2d de la suite de procesamiento de imagen EMAN2 [69]. Los valores de desenfoque se determinaron usando el programa CTFFIND3 [23],

45 El refinamiento de la estructura y las reconstrucciones se llevaron a cabo usando el programa FREALIGN [40]. La búsqueda de orientación inicial se llevó a cabo usando datos combinados 4x y un mapa previamente calculado de las partículas recubiertas VP7 (7RP) (ver referencia 1) como una referencia tridimensional. Los parámetros de alineación iniciales de las partículas escindidas, determinadas por una investigación sistemática (modo 3) en un intervalo angular de 10°, se refinaron adicionalmente contra las últimas reconstrucciones (modo 1) hasta que no hubo mejora adicional en la resolución. Durante la alineación, una máscara de envoltura radial entre 220 y 400 Å se aplicó para retener la densidad que corresponde a la porción de rotavirus y para excluir la densidad que corresponde a los picos de ARN y HA. Las imágenes también se filtraron en paso bajo (hasta 15 Å) para evitar el sobre ajuste posible en el proceso de alineación. Los parámetros de alineación después se usaron para calcular las reconstrucciones de la partícula completa (dentro de la envoltura radial de 600 Å o entre los radios de 220 y 600 Å para los contenidos de las proteínas).

55 Subsecuentemente, imágenes combinadas y no combinadas posteriores 2x se usaron para refinar adicionalmente los parámetros de alineación siguiendo las estrategias similares a aquellas usadas para los datos combinados 4x, es decir una búsqueda sistemática inicial (modo 3) seguido por múltiples ciclos de refinado angular y posicional (modo 1) hasta que no se observó una mejora adicional en la resolución. Durante cada paso de refinación, la máscara de envoltura entre 220 y 400 Å se aplicó los datos de restringieron a la resolución más confiable (hasta 10 Å para datos combinados 2x y hasta 8 Å para datos no combinados). Se calcularon las reconstrucciones para la partícula completa dentro de la envoltura radial de 600 Å.

65 La Figura 5A muestra una sección transversal a través de una reconstrucción tridimensional de la DLP de

rotavirus recubierta con la proteína VP7 modificada solamente. La Figura 5B muestra una sección transversal a través de una reconstrucción tridimensional de la DLP de rotavirus recubierta con un complejo HA-VP7 como se describe en el Ejemplo 6. La Figura 6 muestra una súper posición de una DLP en la reconstrucción tridimensional de la partícula recubierta mostrada en la Figura 5B. La Figura 5C muestra una sección transversal a través de una reconstrucción tridimensional de la DLP de rotavirus recubierta con la proteína VP7 modificada que muestra HA con fragmentos ScFv ligados del anticuerpo CR6261.

Para evaluar la confiabilidad del refinamiento, diferentes conjuntos de refinamiento se llevaron a cabo al enmascarar diferentes regiones de la partícula, y los parámetros de alineación se usaron para calcular las reconstrucciones para regiones omitidas que no se incluyeron durante el refinamiento. Por ejemplo, los parámetros de alineación obtenidos al refinar la porción HA (al aplicar una máscara de envoltura entre 425 y 600 Å) se usaron para calcular la reconstrucción para la porción de rotavirus (al aplicar una máscara de envoltura entre 220 y 400 Å). El mapa resultante contuvo densidades que representan claramente VP2, VP6, y VP7, sugiriendo un sesgo modelo mínimo. El arreglo casi equivalente  $T=13$  de las capas VP6 y VP7 de rotavirus se extendieron claramente al adaptador y la porción HA de la partícula. El promedio de 13 veces local no se aplicó pero se puede llevar a cabo para mejorar adicionalmente la resolución. El ajuste corporal rígido de las VP2, VP6, y VP7 de rotavirus, el atado de seis hélices, la HA, y los Fabs en los mapas de densidad de electrones se llevó a cabo usando la quimera UCSF [47].

Un ejemplo del nivel de detalle que se puede lograr al refinar adicionalmente las reconstrucciones obtenidas de crio-EM como se describe en el párrafo anterior se muestra en la Figura 7. Un detalle de la reconstrucción tridimensional de una partícula de rotavirus que muestra el virus de influenza HA ligado a los fragmentos Fab del anticuerpo CR6261 se muestra. Algunas estructuras secundarias se pueden visualizar en el mapa calculado. Como es esperado, el dominio variable de CR6261 se observa que interactúa con tanto HA1 como HA2.

#### **Ejemplo 8: Estudios de inmunización con los complejos de proteínas HA-VP7 montados en las DLPs de rotavirus**

Ratones BALB/c de 6-8 semanas de edad hembras se inmunizaron intramuscularmente (IM) tres veces en intervalos de tres semanas con una de las tres formulaciones diferentes, cada una formulada en 100  $\mu$ l de una solución amortiguadora que contiene calcio. El grupo 1 recibe los complejos de proteínas HA-VP7 formuladas montadas en las DLPs de rotavirus, el grupo 2 recibe la HA purificada, y el grupo 3 recibe la HA purificada mezclada con las DLPs recubiertas con VP7. Un cuarto grupo de ratones se inyecta de manera simulada con 100  $\mu$ l de solución amortiguadora. Los experimentos pueden incluir varias dosis para cada uno de los grupos. Cantidades equivalentes de antígenos se pueden usar para comparar la respuesta inmunitaria a las dosis equivalentes del antígeno en varios formatos. Alternativamente, las cantidades equivalentes de la proteína total se pueden usar para comparar las respuestas inmunes a varias cantidades de la proteína total.

Uno o más grupos pueden incluir un adyuvante. Los adyuvantes pueden ser particularmente útiles en estimular una respuesta inmunitaria en los ratones en el grupo 2 que reciben el antígeno HA purificado.

Cada grupo de vacuna contiene ocho ratones, y el grupo de control "solución amortiguadora sola" contiene cinco ratones. La sangre se extrae en los días -1, 20, 41, 56 y 84 de la administración de la primera inyección.

Los ratones se monitorean después de la inmunización para el cambio del peso corporal, las reacciones del sitio de la inyección, y otras observaciones clínicas. Las respuestas de anticuerpos a los antígenos administrados se analizan usando los métodos conocidos en la técnica para analizar la especificidad del anticuerpo (por ejemplo ensayo ELISA y ensayo ELISA de competición contra proteínas HA y de rotavirus incluyendo la proteína VP7) y la actividad de neutralización (por ejemplo ensayo de inhibición de hemaglutinación de influenza, ver por ejemplo referencias 70 y 71).

Los resultados muestran una respuesta inmunitaria sustancialmente más alta a los complejos de proteínas HA-VP7 montados en las DLPs de rotavirus como es comparada con la HA purificada, incluyendo la HA purificada formulada con el adyuvante. El ELISA de competición muestra que las respuestas de anticuerpos a los complejos de proteínas HA-VP7 son predominantemente a la porción HA del complejo antes de la VP7 o porción de rotavirus del complejo. Los resultados muestran títulos de anticuerpos de neutralización sustancialmente más altos como resultado de la inmunización con los complejos de proteínas HA-VP7 montados en las DLPs de rotavirus como es comparada con la HA purificada, incluyendo la HA purificada formulada con el adyuvante. Estos resultados muestran la inmunogenicidad incrementada de los antígenos de la gripe cuando se montan en la DLP recubierta con VP7 de rotavirus como es comparada con los antígenos de la gripe no montados en las DLPs recubiertas con VP7.

Todos los experimentos animales se llevaron a cabo de acuerdo con los protocolos del Comité del Cuidado y Uso Animal Institucional.

#### **Ejemplo 9: Estudios de inmunización con complejos de proteínas HA-VP7 montados en las DLPs de**

**rotavirus con y sin adyuvante**

5 Ratones BALB/c hembras se inmunizaron mediante dos inyecciones intramusculares de 50 µl bilaterales en los cuádriceps traseros en los días 0, 21, y 42 con 0.3-15 µg de los complejos de proteína de hemaglutinina (HA) H1N1 de antígeno purificado como se expone en la siguiente Tabla 4.

Los complejos se formularon con PBS o adyuvante 1-2 horas antes de la inmunización. Las vacunas de subunidad formuladas se mantienen sobre hielo hasta la administración.

10 Los ratones se monitorearon después de la inmunización para el cambio de peso corporal, las reacciones en el sitio de inyección, y otras observaciones clínicas. Se obtuvieron muestras de suero de los animales por sangrados de seno retro-orbital en el día 20 (3 semanas pos-primera inmunización) y el día 41 (3 semanas pos-segunda inmunización) y de los sangrados de los animales sacrificados en el día 63 (3 semanas pos-tercera inmunización).

15 Las respuestas de anticuerpos a los antígenos administrados se analizan usando los métodos conocidos en la técnica (por ejemplo ensayo ELISA, ensayo de neutralización). Los resultados muestran una respuesta inmunitaria sustancialmente más alta a los complejos de proteínas HA-VP7 montadas sobre las DLPs de rotavirus como es comparada con la HA purificada, incluyendo la HA purificada formulada con el adyuvante o HA purificada formulada con, pero no montada en las DLPs recubiertas con VP7.

20 Todos los experimentos animales se llevaron a cabo de acuerdo con los protocolos del Comité de Cuidado y Uso Animal Institucional.

25 **Tabla 4**

Grupo	Vacuna IM (dosis)	N
30 1	Monovolumen A/Sol (0.3lj.g)	5
2	7RP (0.3 M.g)	5
3	7RP (3 M-g)	5
35 4	7RP (15  Lig)	5
5	HA + 7RP (0.3 p,g)	5
6	HA + 7RP (3 p,g)	5
40 7	HA + 7RP (15 p,g)	5
8	HA-7RP (0.3 p,g)	5
45 9	HA-7RP (3 p,g)	5
10	HA-7RP (15 jug)	5
11	A/Sol HA (0.3 p,g)	5
50 12	A/Sol HA (3 p,g)	5
13	A/Sol HA (15 p,g)	5

55 Monovolumen A/Sol es la HA recombinante, trimerizada oligomerizada con lípidos.

7RP es las partículas recubiertas con VP7 sin HA.

HA + 7RP es partículas recubiertas con VP7 en las cuales la HA no se puede montar + HA recombinante trimerizada HA-7RP es partículas recubiertas con VP7 montadas con HA recombinante, trimerizada.

A/Sol HA es HA recombinante trimerizada.

60 Se entenderá que la invención se ha descrito por medio de ejemplo solamente y se pueden hacer modificaciones mientras que permanezcan dentro del alcance y el espíritu de la invención.

**REFERENCIAS**

65 [1] Chen y colaboradores (2009) Proc Nati Acad Sci USA 106(26): 10644-8

- [2] Trask & Dormitzer (2006) *J Virol* 80(22): 11293-304
- [3] Kobayashi y colaboradores (2007) *Cell Host Microbe* 1(2):147-157
- [4] Chandran y colaboradores (1999) *J Virol* 73(5):3941-50
- 5 [5] Chandran y colaboradores (2001) *J Virol* 75(11):5335-42
- [6] Yan y colaboradores (2011) *J Virol* 85(15):7483-95
- [7] Miyazaki y colaboradores (2008) *J Virol* 82(22): 11344-53
- [8] Zhang y colaboradores (2010) *Cell* 141(3):472-82
- [9] Biosafety in Microbiological and Biomedical Laboratories (BMBL) 5th Edition, CDC, 2009
- 10 [10] Yang y colaboradores (2000) *J Virol* 74(10):4746-54
- [11] Meier y colaboradores (2004) *J Mol Biol* 344(4): 1051-69
- [12] Wu & Wong (2005) *JBiol Chem* 280(24):23225-31
- [13] Beckett y colaboradores (1999) *Protein Sci* 8(4):921-9
- [14] Chen y colaboradores (1997) *Cell* 89(2):263-73
- 15 [15] Zhu y colaboradores (2003) *Protein Eng* 16(5):373-9
- [16] Yu y colaboradores (2002) *J Gen Virol* 83(Pt 3):623-9
- [17] Xu y colaboradores (2004) *JBiol Chem* 279(29):30514-22
- [18] <http://partsregistry.org/Protein domains/Linker>
- [19] US 4,769,326
- 20 [20] Dormitzer y colaboradores (2001) *J Virol* 75(16):7339-50
- [21] Dormitzer y colaboradores (2000) *Virology* 277(2):420-8
- [22] Mohd.Jaafar y colaboradores (2005) *J Gen Virol* 86(Pt 4): 1147-57
- [23] Mindell & Grigorieff (2003) *JStruct Biol* 142(3):334-47
- [24] Campbell y colaboradores (2012) *Structure* 20(11):1823-8
- 25 [25] Setiembre y colaboradores (2011) *EMBO .1.* 30(2):408-16
- [26] Chen & Grigorieff (2007) *J Struct Biol* 157(1): 168-73
- [27] Wriggers y colaboradores (1999) *J Struct Biol* 125(2-3):185-95
- [28] Jiang y colaboradores (2001) *J Mol Biol* 308(5):1033-44
- [29] Topf y colaboradores (2005) *J Struct Biol* 149(2): 191-203
- 30 [30] Tama y colaboradores (2004) *J Struct Biol* 147(3):315-2
- [31] Topf y colaboradores (2008) *Structure* 16(2):295-307
- [32] Trabuco y colaboradores (2009) *Methods* 49(2): 174-80
- [33] Schröder y colaboradores (2007) *Structure* 15(12): 1630-41
- [34] Zhang y colaboradores (2010) *Nature* 463(7279):379-83
- 35 [35] Baker y colaboradores (2006) *PLoS Comput Biol* 2(10):e146
- [36] DiMaio y colaboradores (2009) *J Mol Biol* 392(1): 181-90
- [37] Topf y colaboradores (2006) *JMol Biol* 357(5): 1655-68
- [38] Zhu y colaboradores (2010) *JMol Biol* 397(3):835-51
- [39] Liang y colaboradores (2002) *JStruct Biol* 137(3):292-304
- 40 [40] Grigorieff (2007) *JStruct Biol* 157(1):117-25
- [41] Cheng y colaboradores (2010) *JMol Biol* 397(3):852-63
- [42] Jiang y colaboradores (2008) *Nature* 451(7182): 1130-4
- [43] Ludtke y colaboradores (2008) *Structure* 16(3):441-8
- [44] Yu y colaboradores (2008) *Nature* 453(7193):415-9
- 45 [45] Ludtke y colaboradores (1999) *J Struct Biol* 128(l):82-97
- [46] Baker y colaboradores (2001) *Structure* 15(1):7-19
- [47] Pettersen y colaboradores (2004) *JComput Chem* 25(13):1605-12
- [48] Emsley & Cowtan (2004) *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-32
- [49] Emsley y colaboradores (2010) *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501
- 50 [50] Baker y colaboradores (2011) *JStruct Biol* 174(2):360-73
- [51] Zhuy colaboradores (2010) *JMol Biol* 397(3):835-51
- [52] Jones y colaboradores (1991) *Acta Crystallogr A* 47 (Pt 2): 110-9
- [53] Brünger y colaboradores (1998) *Acta Crystallogr D Biol Crystallogr* 54(Pt 5):905-21
- [54] Li & Zhang (2009) *Proteins* 76(3):665-76
- 55 [55] Brünger (1992) *Nature* 355(6359):472-5
- [56] Laskowski y colaboradores (1993) *JAppl Cryst* 26:283-91
- [57] Whittle y colaboradores (2011) *Proc Nati Acad Sci USA* 108(34): 14216-21
- [58] Aoki y colaboradores (2009) *Science* 324(5933):1444-7
- [59] Krause y colaboradores (2011) *J Virol* 85(20): 10905-8
- 60 [60] Bommakanti y colaboradores (2010) *Proc Nati Acad Sci USA* 107(31): 13701-6
- [61] Chakraborty y colaboradores (2006) *Biochem J* 399(3):483-91
- [62] Bhattacharyya y colaboradores (2010) *J Biol Chem* 285(35):27100-10
- [63] Ward y colaboradores (1989) *Nature* 341:544-546
- [64] Bird y colaboradores (1988) *Science* 242:423-426
- 65 [65] Huston y colaboradores (1988) *Proc. Nati. Acad. Sci. USA* 85:5879-5883
- [66] Council of Europe (2014) *European pharmacopoeia*

- [67] Ekiert y colaboradores (2009) Science 324(5924): 246-251  
 [68] Lingwood y colaboradores (2012) Nature 489(7417):566-70  
 [69] Tang y colaboradores (2007) JStruct Biol 157(l):38-46  
 [70] Salk y colaboradores (1945) Am JHyg 42:57-93  
 [71] de Jong et al., Haemagglutination-inhibiting antibody to influenza virus. In Brown y colaboradores (eds.) (2003) Laboratory Correlates of Immunity to Influenza- A Reassessment. Basel, Switzerland: Karger, 63-73

**LISTADO DE SECUENCIAS**

10

<110> NOVARTIS AG  
 CHILDREN'S MEDICAL CENTER CORPORATION  
 BRANDEIS UNIVERSITY

15

<120> PARTÍCULAS DE ROTAVIRUS CON PROTEÍNAS DE SUPERFICIE QUIMÉRICAS

<130> P064203WO

20

<141> 2014-07-07

<160> 83

<170> SeqWin2010, versión 1.0

25

<210> 1  
 <211> 28  
 <212> PRT  
 <213> Secuencia Artificial

30

<220>  
 <223> Secuencia HR2 de Cepa HXB2 del VIH

35

<400> 1  
 Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile His  
 1 5 10 15

Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys  
 20 25

40

<210> 2  
 <211> 34  
 <212> PRT  
 <213> Secuencia Artificial

45

<220>  
 <223> Secuencia HR1 de cepa HXB2 de VIH

50

<400> 2  
 Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala  
 1 5 10 15

Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln  
 20 25 30

55

Ala Arg

60

<210> 3  
 <211> 30  
 <212> PRT

ES 2 733 279 T3

<213> Secuencia Artificial

<220>

<223> Secuencia HR2 del virus Nipah

5

<400> 3  
 Asp Ile Ser Ser Gln Ile Ser Ser Met Asn Gln Ser Leu Gln Gln Ser  
 1 5 10 15

10 Lys Asp Tyr Ile Lys Glu Ala Gln Arg Leu Leu Asp Thr Val  
 20 25 30

<210> 4  
 <211> 32  
 <212> PRT  
 <213> Secuencia Artificial

<220>

<223> Secuencia HR1 del virus Nipah

20

<400> 4  
 Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn Glu Ala Val Val Lys  
 1 5 10 15

25 Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val Leu Thr Ala Leu Gln  
 20 25 30

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

<220>

<223> Ligador

35

<400> 5  
 Gly Gly Ser Gly Gly  
 1 5

40 <210> 6  
 <211> 8  
 <212> PRT  
 <213> Secuencia Artificial

<220>

<223> Ligador

<400> 6  
 Gly Ser Gly Ser Gly Ser Gly Thr  
 1 5

50

<210> 7  
 <211> 5  
 <212> PRT  
 <213> Secuencia Artificial

<220>

<223> Ligador

60 <400> 7

ES 2 733 279 T3

Glu Ala Ala Ala Arg  
 1 5

5 <210> 8  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 10 <223> Sitio de reconocimiento del factor XI humano

<400> 8  
 Gln Thr Ser Lys Leu Thr Arg Ala Glu Ala Val Phe  
 1 5 10

15 <210> 9  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

20 <220>  
 <223> Sitio de reconocimiento del factor XI humano

<400> 9  
 25 Ser Phe Asn Asp Phe Thr Arg Val Val Gly Gly Glu  
 1 5 10

<210> 10  
 <211> 13  
 30 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Sitio de reconocimiento de calicreína humana

35 <400> 10  
 Leu Phe Ser Ser Met Thr Arg Val Val Gly Gly Leu Val  
 1 5 10

40 <210> 11  
 <211> 10  
 <212> PRT  
 <213> Secuencia Artificial

45 <220>  
 <223> Sitio de reconocimiento del factor XII humano

<400> 11  
 Lys Ile Lys Pro Arg Ile Val Gly Gly Thr  
 50 1 5 10

<210> 12  
 <211> 8  
 <212> PRT  
 55 <213> Secuencia Artificial

<220>  
 <223> Etiqueta FLAG

60 <400> 12



ES 2 733 279 T3

Asp Tyr Lys Asp Asp Asp Asp Lys  
 1 5

5 <210> 13  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 10 <223> Etiqueta HA

<400> 13

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala  
 1 5

15 <210> 14  
 <211> 6  
 <212> PRT  
 20 <213> Secuencia Artificial

<220>  
 <223> Etiqueta His

25 <400> 14  
 His His His His His His  
 1 5

30 <210> 15  
 <211> 10  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 35 <223> Etiqueta Myc

<400> 15  
 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu  
 1 5 10

40 <210> 16  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial

45 <220>  
 <223> Etiqueta Strep I  
 <400> 16

Ala Trp Arg His Pro Gln Phe Gly Gly  
 1 5

50 <210> 17  
 <211> 9  
 <212> PRT  
 55 <213> Secuencia Artificial

<220>  
 <223> Etiqueta II Strep

60 <400> 17

ES 2 733 279 T3

Asn Trp Ser His Pro Gln Phe Glu Lys  
 1 5

5 <210> 18  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 10 <223> Etiqueta de proteína C

<400> 18  
 Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys  
 1 5 10

15 <210> 19  
 <211> 6  
 <212> ADN  
 <213> Secuencia Artificial

20 <220>  
 <223> Secuencia Kozak

<400> 19  
 gccacc 6

25 <210> 20  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Ligador

<400> 20  
 35 Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser  
 1 5 10 15

40 <210> 21  
 <211> 17  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Ligador

45 <400> 21  
 Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly  
 1 5 10 15

50 Gly

<210> 22  
 <211> 367  
 55 <212> PRT  
 <213> Secuencia Artificial

<220>  
 60 <223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
 ingeniería

ES 2 733 279 T3

<400> 22  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 5 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30  
 10 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45  
 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60  
 15 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 20 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
 25 Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
 130 135 140  
 30 Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
 145 150 155 160  
 Asn Gln Gln Glu Lys Ile Glu Gly Arg Ser Gly Gly Glu Asp Gln Val  
 165 170 175  
 35 Asp Pro Arg Leu Ile Asp Gly Lys Ser Gly Gly Thr Leu Gln Leu Asp  
 180 185 190  
 Met Ser Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro  
 195 200 205  
 40 Met Asp Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys  
 210 215 220  
 45 Trp Ile Ser Met Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu Asn  
 225 230 235 240  
 Thr Gln Thr Leu Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe  
 245 250 255  
 50 Glu Glu Val Ala Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val Asp  
 260 265 270  
 Gly Val Asn His Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile Arg  
 275 280 285  
 55 Asn Cys Lys Lys Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val  
 290 295 300  
 60 Gly Gly Ser Asp Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro  
 305 310 315 320

ES 2 733 279 T3

Gln Thr Glu Arg Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln Val  
 325 330 335

5 Phe Tyr Thr Val Val Asp Tyr Val Asn Gln Ile Ile Gln Ala Met Ser  
 340 345 350

Lys Arg Ser Arg Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 355 360 365

10 <210> 23  
 <211> 365  
 <212> PRT  
 <213> Secuencia Artificial

15 <220>  
 <223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
 ingeniería  
 <400> 23

20 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15

Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30

25 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45

30 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60

Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80

35 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95

40 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110

Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125

45 Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
 130 135 140

Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
 145 150 155 160

50 Asn Gln Gln Glu Lys Ile Glu Gly Arg Ser Gly Gly Glu Asp Gln Val  
 165 170 175

55 Asp Pro Arg Leu Ile Asp Gly Lys Gly Thr Leu Gln Leu Asp Met Ser  
 180 185 190

Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp  
 195 200 205

60 Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile

ES 2 733 279 T3

	210		215		220														
	Ser	Met	Gly	Ser	Ser	Cys	Thr	Ile	Lys	Val	Cys	Pro	Leu	Asn	Thr	Gln			
	225					230					235					240			
5	Thr	Leu	Gly	Ile	Gly	Cys	Leu	Thr	Thr	Asp	Thr	Ala	Thr	Phe	Glu	Glu			
					245					250					255				
10	Val	Ala	Thr	Ala	Glu	Lys	Leu	Val	Ile	Thr	Asp	Val	Val	Asp	Gly	Val			
				260					265					270					
	Asn	His	Lys	Leu	Asp	Val	Thr	Thr	Ala	Thr	Cys	Thr	Ile	Arg	Asn	Cys			
			275					280					285						
15	Lys	Lys	Leu	Gly	Pro	Arg	Glu	Asn	Val	Ala	Val	Ile	Gln	Val	Gly	Gly			
		290					295					300							
	Ser	Asp	Val	Leu	Asp	Ile	Thr	Ala	Asp	Pro	Thr	Thr	Ala	Pro	Gln	Thr			
20	305					310					315					320			
	Glu	Arg	Met	Met	Arg	Ile	Asn	Trp	Lys	Lys	Trp	Trp	Gln	Val	Phe	Tyr			
					325					330					335				
25	Thr	Val	Val	Asp	Tyr	Val	Asn	Gln	Ile	Ile	Gln	Ala	Met	Ser	Lys	Arg			
				340					345					350					
	Ser	Arg	Ser	Leu	Asn	Ser	Ala	Ala	Phe	Tyr	Tyr	Arg	Ile						
30			355					360					365						
	<210>	24																	
	<211>	363																	
	<212>	PRT																	
	<213>	Secuencia Artificial																	
35	<220>																		
	<223>	Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por ingeniería																	
40	<400>	24																	
	Met	Lys	Phe	Leu	Val	Asn	Val	Ala	Leu	Val	Phe	Met	Val	Val	Tyr	Ile			
	1				5					10					15				
45	Ser	Tyr	Ile	Tyr	Ala	Glu	Asp	Ser	Ala	Trp	Ser	His	Pro	Gln	Phe	Glu			
				20					25					30					
	Lys	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Gln	Asn	Tyr	Gly	Ile	Asn	Leu	Pro			
			35					40					45						
50	Ile	Thr	Gly	Ser	Met	Asp	Thr	Ala	Tyr	Ala	Asn	Ser	Thr	Gln	Glu	Glu			
		50					55					60							
	Thr	Phe	Leu	Thr	Ser	Thr	Leu	Cys	Leu	Tyr	Tyr	Pro	Thr	Glu	Ala	Ala			
55	65					70					75					80			
	Thr	Glu	Ile	Asn	Asp	Asn	Ser	Trp	Lys	Asp	Thr	Leu	Ser	Gln	Leu	Phe			
				85						90					95				
60	Leu	Thr	Lys	Gly	Trp	Pro	Thr	Gly	Ser	Val	Tyr	Phe	Lys	Glu	Tyr	Thr			
				100					105					110					

ES 2 733 279 T3

Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
115 120 125

5 Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
130 135 140

Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
145 150 155 160

10 Asn Gln Gln Glu Lys Ser Gly Gly Glu Asp Gln Val Asp Pro Arg Leu  
165 170 175

15 Ile Asp Gly Lys Ser Gly Gly Thr Leu Gln Leu Asp Met Ser Glu Leu  
180 185 190

Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr  
195 200 205

20 Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met  
210 215 220

Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu  
225 230 235 240

25 Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala  
245 250 255

30 Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His  
260 265 270

Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys  
275 280 285

35 Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp  
290 295 300

Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg  
305 310 315 320

40 Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val  
325 330 335

45 Val Asp Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg  
340 345 350

Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
355 360

50 <210> 25  
<211> 361  
<212> PRT  
<213> Secuencia Artificial

55 <220>  
<223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
ingeniería

<400> 25

60 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile

ES 2 733 279 T3

	1			5					10					15		
	Ser	Tyr	Ile	Tyr	Ala	Glu	Asp	Ser	Ala	Trp	Ser	His	Pro	Gln	Phe	Glu
				20					25					30		
5	Lys	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Gln	Asn	Tyr	Gly	Ile	Asn	Leu	Pro
			35					40					45			
	Ile	Thr	Gly	Ser	Met	Asp	Thr	Ala	Tyr	Ala	Asn	Ser	Thr	Gln	Glu	Glu
10		50					55					60				
	Thr	Phe	Leu	Thr	Ser	Thr	Leu	Cys	Leu	Tyr	Tyr	Pro	Thr	Glu	Ala	Ala
	65					70					75					80
15	Thr	Glu	Ile	Asn	Asp	Asn	Ser	Trp	Lys	Asp	Thr	Leu	Ser	Gln	Leu	Phe
				85						90					95	
	Leu	Thr	Lys	Gly	Trp	Pro	Thr	Gly	Ser	Val	Tyr	Phe	Lys	Glu	Tyr	Thr
20				100					105					110		
	Asp	Ile	Ala	Ser	Phe	Ser	Val	Asp	Pro	Gln	Leu	Tyr	Cys	Asp	Tyr	Asn
			115					120					125			
25	Val	Val	Leu	Met	Lys	Tyr	Asp	Ala	Gly	Trp	Met	Glu	Trp	Asp	Arg	Glu
		130					135					140				
	Ile	Asn	Asn	Tyr	Thr	Ser	Leu	Ile	His	Ser	Leu	Ile	Glu	Glu	Ser	Gln
30		145				150					155					160
	Asn	Gln	Gln	Glu	Lys	Ser	Gly	Gly	Glu	Asp	Gln	Val	Asp	Pro	Arg	Leu
				165						170					175	
35	Ile	Asp	Gly	Lys	Gly	Thr	Leu	Gln	Leu	Asp	Met	Ser	Glu	Leu	Ala	Asp
				180					185					190		
	Leu	Ile	Leu	Asn	Glu	Trp	Leu	Cys	Asn	Pro	Met	Asp	Ile	Thr	Leu	Tyr
			195					200					205			
40	Tyr	Tyr	Gln	Gln	Thr	Asp	Glu	Ala	Asn	Lys	Trp	Ile	Ser	Met	Gly	Ser
		210					215					220				
	Ser	Cys	Thr	Ile	Lys	Val	Cys	Pro	Leu	Asn	Thr	Gln	Thr	Leu	Gly	Ile
45		225				230					235					240
	Gly	Cys	Leu	Thr	Thr	Asp	Thr	Ala	Thr	Phe	Glu	Glu	Val	Ala	Thr	Ala
				245						250					255	
50	Glu	Lys	Leu	Val	Ile	Thr	Asp	Val	Val	Asp	Gly	Val	Asn	His	Lys	Leu
				260					265					270		
	Asp	Val	Thr	Thr	Ala	Thr	Cys	Thr	Ile	Arg	Asn	Cys	Lys	Lys	Leu	Gly
			275					280					285			
55	Pro	Arg	Glu	Asn	Val	Ala	Val	Ile	Gln	Val	Gly	Gly	Ser	Asp	Val	Leu
		290					295					300				
	Asp	Ile	Thr	Ala	Asp	Pro	Thr	Thr	Ala	Pro	Gln	Thr	Glu	Arg	Met	Met
60						310					315					320

ES 2 733 279 T3

Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val Val Asp  
 325 330 335  
 Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg Ser Leu  
 5 340 345 350  
 Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 355 360  
 10 <210> 26  
 <211> 359  
 <212> PRT  
 <213> Secuencia Artificial  
 15 <220>  
 <223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
 ingeniería  
 <400> 26  
 20 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30  
 25 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45  
 30 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60  
 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
 35 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 40 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
 45 Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
 130 135 140  
 50 Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
 145 150 155 160  
 Asn Gln Gln Glu Lys Gly Glu Asp Gln Val Asp Pro Arg Leu Ile Asp  
 165 170 175  
 55 Gly Lys Gly Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp Leu Ile  
 180 185 190  
 Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr  
 195 200 205  
 60 Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser Ser Cys



ES 2 733 279 T3

210 215 220

5 Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu Gly Ile Gly Cys  
225 230 235 240

Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala Thr Ala Glu Lys  
245 250 255

10 Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His Lys Leu Asp Val  
260 265 270

Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly Pro Arg  
275 280 285

15 Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp Val Leu Asp Ile  
290 295 300

Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg Met Met Arg Ile  
305 310 315 320

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

25 Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg Ser Leu Asn Ser  
340 345 350

30 Ala Ala Phe Tyr Tyr Arg Ile  
355  
<210> 27  
<211> 377  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
ingeniería

<400> 27  
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
1 5 10 15

40 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
20 25 30

45 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
35 40 45

Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
50 55 60

50 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
65 70 75 80

55 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
85 90 95

Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
100 105 110

60 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
115 120 125

ES 2 733 279 T3

5 Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
130 135 140

Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
145 150 155 160

10 Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp  
165 170 175

Ala Ser Leu Ser Gly Gly Glu Asp Gln Val Asp Pro Arg Leu Ile Asp  
180 185 190

15 Gly Lys Ser Gly Gly Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp  
195 200 205

20 Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr  
210 215 220

Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser  
225 230 235 240

25 Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu Gly Ile  
245 250 255

Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala Thr Ala  
260 265 270

30 Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His Lys Leu  
275 280 285

35 Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly  
290 295 300

Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp Val Leu  
305 310 315 320

40 Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg Met Met  
325 330 335

Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val Val Asp  
340 345 350

45 Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg Ser Leu  
355 360 365

50 Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
370 375

<210> 28  
<211> 373  
<212> PRT  
55 <213> Secuencia Artificial

<220>  
<223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
ingeniería

60

ES 2 733 279 T3

<400> 28  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 5 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30  
 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45  
 10 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60  
 15 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 20 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
 25 Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
 130 135 140  
 30 Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
 145 150 155 160  
 Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp  
 165 170 175  
 35 Ala Ser Leu Gly Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys  
 180 185 190  
 Gly Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp Leu Ile Leu Asn  
 195 200 205  
 Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr Gln Gln  
 210 215 220  
 45 Thr Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser Ser Cys Thr Ile  
 225 230 235 240  
 Lys Val Cys Pro Leu Asn Thr Gln Thr Leu Gly Ile Gly Cys Leu Thr  
 245 250 255  
 50 Thr Asp Thr Ala Thr Phe Glu Glu Val Ala Thr Ala Glu Lys Leu Val  
 260 265 270  
 Ile Thr Asp Val Val Asp Gly Val Asn His Lys Leu Asp Val Thr Thr  
 275 280 285  
 55 Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly Pro Arg Glu Asn  
 290 295 300  
 60 Val Ala Val Ile Gln Val Gly Gly Ser Asp Val Leu Asp Ile Thr Ala  
 305 310 315 320

ES 2 733 279 T3

Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg Met Met Arg Ile Asn Trp  
325 330 335

5 Lys Lys Trp Trp Gln Val Phe Tyr Thr Val Val Asp Tyr Val Asn Gln  
340 345 350

Ile Ile Gln Ala Met Ser Lys Arg Ser Arg Ser Leu Asn Ser Ala Ala  
355 360 365

10 Phe Tyr Tyr Arg Ile  
370

15 <210> 29  
<211> 373  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
ingeniería

<400> 29

25 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
1 5 10 15

Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
20 25 30

30 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
35 40 45

35 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
50 55 60

Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
65 70 75 80

40 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
85 90 95

45 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
100 105 110

Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
115 120 125

50 Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
130 135 140

Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
145 150 155 160

55 Asn Gln Gln Glu Lys Gly Leu Glu Leu Asp Lys Trp Ala Ser Leu Ser  
165 170 175

Gly Gly Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys Ser Gly  
180 185 190

60 Gly Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp Leu Ile Leu Asn

ES 2 733 279 T3

	195		200		205														
	Glu	Trp	Leu	Cys	Asn	Pro	Met	Asp	Ile	Thr	Leu	Tyr	Tyr	Tyr	Gln	Gln			
	210						215					220							
5	Thr	Asp	Glu	Ala	Asn	Lys	Trp	Ile	Ser	Met	Gly	Ser	Ser	Cys	Thr	Ile			
	225					230					235					240			
10	Lys	Val	Cys	Pro	Leu	Asn	Thr	Gln	Thr	Leu	Gly	Ile	Gly	Cys	Leu	Thr			
					245					250					255				
	Thr	Asp	Thr	Ala	Thr	Phe	Glu	Glu	Val	Ala	Thr	Ala	Glu	Lys	Leu	Val			
				260					265					270					
15	Ile	Thr	Asp	Val	Val	Asp	Gly	Val	Asn	His	Lys	Leu	Asp	Val	Thr	Thr			
			275					280					285						
	Ala	Thr	Cys	Thr	Ile	Arg	Asn	Cys	Lys	Lys	Leu	Gly	Pro	Arg	Glu	Asn			
	290						295					300							
20	Val	Ala	Val	Ile	Gln	Val	Gly	Gly	Ser	Asp	Val	Leu	Asp	Ile	Thr	Ala			
	305					310					315					320			
25	Asp	Pro	Thr	Thr	Ala	Pro	Gln	Thr	Glu	Arg	Met	Met	Arg	Ile	Asn	Trp			
					325					330					335				
	Lys	Lys	Trp	Trp	Gln	Val	Phe	Tyr	Thr	Val	Val	Asp	Tyr	Val	Asn	Gln			
				340					345					350					
30	Ile	Ile	Gln	Ala	Met	Ser	Lys	Arg	Ser	Arg	Ser	Leu	Asn	Ser	Ala	Ala			
			355					360					365						
	Phe	Tyr	Tyr	Arg	Ile														
	370																		
35	<210>	30																	
	<211>	369																	
	<212>	PRT																	
	<213>	Secuencia Artificial																	
40	<220>																		
	<223>	Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por ingeniería																	
	<400>	30																	
45	Met	Lys	Phe	Leu	Val	Asn	Val	Ala	Leu	Val	Phe	Met	Val	Val	Tyr	Ile			
	1				5					10					15				
	Ser	Tyr	Ile	Tyr	Ala	Glu	Asp	Ser	Ala	Trp	Ser	His	Pro	Gln	Phe	Glu			
				20					25					30					
50	Lys	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Gln	Asn	Tyr	Gly	Ile	Asn	Leu	Pro			
			35					40					45						
	Ile	Thr	Gly	Ser	Met	Asp	Thr	Ala	Tyr	Ala	Asn	Ser	Thr	Gln	Glu	Glu			
	50						55					60							
55	Thr	Phe	Leu	Thr	Ser	Thr	Leu	Cys	Leu	Tyr	Tyr	Pro	Thr	Glu	Ala	Ala			
	65					70					75					80			
60	Thr	Glu	Ile	Asn	Asp	Asn	Ser	Trp	Lys	Asp	Thr	Leu	Ser	Gln	Leu	Phe			
				85						90					95				

ES 2 733 279 T3

Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
100 105 110

5 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
115 120 125

Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
130 135 140

10 Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
145 150 155 160

15 Asn Gln Gln Glu Lys Gly Leu Glu Leu Asp Lys Trp Ala Ser Leu Gly  
165 170 175

Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys Gly Thr Leu Gln  
180 185 190

20 Leu Asp Met Ser Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys  
195 200 205

Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala  
210 215 220

25 Asn Lys Trp Ile Ser Met Gly Ser Ser Cys Thr Ile Lys Val Cys Pro  
225 230 235 240

30 Leu Asn Thr Gln Thr Leu Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala  
245 250 255

Thr Phe Glu Glu Val Ala Thr Ala Glu Lys Leu Val Ile Thr Asp Val  
260 265 270

35 Val Asp Gly Val Asn His Lys Leu Asp Val Thr Thr Ala Thr Cys Thr  
275 280 285

Ile Arg Asn Cys Lys Lys Leu Gly Pro Arg Glu Asn Val Ala Val Ile  
290 295 300

40 Gln Val Gly Gly Ser Asp Val Leu Asp Ile Thr Ala Asp Pro Thr Thr  
305 310 315 320

45 Ala Pro Gln Thr Glu Arg Met Met Arg Ile Asn Trp Lys Lys Trp Trp  
325 330 335

Gln Val Phe Tyr Thr Val Val Asp Tyr Val Asn Gln Ile Ile Gln Ala  
340 345 350

50 Met Ser Lys Arg Ser Arg Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg  
355 360 365

Ile

55 <210> 31  
<211> 379  
<212> PRT  
<213> Secuencia Artificial

60

ES 2 733 279 T3

<220>

<223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por ingeniería

5 <400> 31  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15

10 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30

Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45

15 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60

20 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80

Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95

25 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110

Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125

30 Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp Arg Glu  
 130 135 140

35 Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln  
 145 150 155 160

Asn Gln Gln Glu Lys Gly Leu Glu Leu Asp Lys Trp Ala Ser Leu Gly  
 165 170 175

40 Ser Ile Glu Gly Arg Ser Gly Gly Glu Asp Gln Val Asp Pro Arg Leu  
 180 185 190

45 Ile Asp Gly Lys Ser Gly Gly Thr Leu Gln Leu Asp Met Ser Glu Leu  
 195 200 205

Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr  
 210 215 220

50 Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met  
 225 230 235 240

Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu  
 245 250 255

55 Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala  
 260 265 270

60 Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His  
 275 280 285

ES 2 733 279 T3

Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys  
 290 295 300  
 5 Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp  
 305 310 315 320  
 Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg  
 325 330 335  
 10 Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val  
 340 345 350  
 Val Asp Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg  
 355 360 365  
 15 Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 370 375  
 20 <210> 32  
 <211> 370  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 25 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por  
 ingeniería  
 <400> 32  
 30 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30  
 35 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45  
 40 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60  
 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
 45 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
 50 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
 55 Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140  
 Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160  
 60 Gln Arg Leu Leu Asp Thr Val Lys Ile Glu Gly Arg Ser Gly Gly Glu



ES 2 733 279 T3

				165					170					175			
	Asp	Gln	Val	Asp	Pro	Arg	Leu	Ile	Asp	Gly	Lys	Ser	Gly	Gly	Thr	Leu	
				180					185					190			
5	Gln	Leu	Asp	Met	Ser	Glu	Leu	Ala	Asp	Leu	Ile	Leu	Asn	Glu	Trp	Leu	
			195					200					205				
10	Cys	Asn	Pro	Met	Asp	Ile	Thr	Leu	Tyr	Tyr	Tyr	Gln	Gln	Thr	Asp	Glu	
		210					215					220					
15	Ala	Asn	Lys	Trp	Ile	Ser	Met	Gly	Ser	Ser	Cys	Thr	Ile	Lys	Val	Cys	
		225				230					235					240	
20	Pro	Leu	Asn	Thr	Gln	Thr	Leu	Gly	Ile	Gly	Cys	Leu	Thr	Thr	Asp	Thr	
					245					250					255		
25	Ala	Thr	Phe	Glu	Glu	Val	Ala	Thr	Ala	Glu	Lys	Leu	Val	Ile	Thr	Asp	
				260					265					270			
30	Val	Val	Asp	Gly	Val	Asn	His	Lys	Leu	Asp	Val	Thr	Thr	Ala	Thr	Cys	
			275					280					285				
35	Thr	Ile	Arg	Asn	Cys	Lys	Lys	Leu	Gly	Pro	Arg	Glu	Asn	Val	Ala	Val	
		290					295					300					
40	Ile	Gln	Val	Gly	Gly	Ser	Asp	Val	Leu	Asp	Ile	Thr	Ala	Asp	Pro	Thr	
						310					315					320	
45	Thr	Ala	Pro	Gln	Thr	Glu	Arg	Met	Met	Arg	Ile	Asn	Trp	Lys	Lys	Trp	
					325					330					335		
50	Trp	Gln	Val	Phe	Tyr	Thr	Val	Val	Asp	Tyr	Val	Asn	Gln	Ile	Ile	Gln	
				340					345				350				
55	Ala	Met	Ser	Lys	Arg	Ser	Arg	Ser	Leu	Asn	Ser	Ala	Ala	Phe	Tyr	Tyr	
			355					360					365				
60	Arg	Ile															
		370															
65	<210>	33															
	<211>	368															
	<212>	PRT															
	<213>	Secuencia Artificial															
70	<220>																
	<223>	Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por ingeniería															
75	<400>	33															
80	Met	Lys	Phe	Leu	Val	Asn	Val	Ala	Leu	Val	Phe	Met	Val	Val	Tyr	Ile	
	1				5					10					15		
85	Ser	Tyr	Ile	Tyr	Ala	Glu	Asp	Ser	Ala	Trp	Ser	His	Pro	Gln	Phe	Glu	
				20					25					30			
90	Lys	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Gln	Asn	Tyr	Gly	Ile	Asn	Leu	Pro	
			35					40					45				
95	Ile	Thr	Gly	Ser	Met	Asp	Thr	Ala	Tyr	Ala	Asn	Ser	Thr	Gln	Glu	Glu	
		50					55					60					

ES 2 733 279 T3

Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
**5** Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
**10** Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
**15** Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140  
 Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160  
**20** Gln Arg Leu Leu Asp Thr Val Lys Ile Glu Gly Arg Ser Gly Gly Glu  
 165 170 175  
**25** Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys Gly Thr Leu Gln Leu  
 180 185 190  
 Asp Met Ser Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn  
 195 200 205  
**30** Pro Met Asp Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn  
 210 215 220  
 Lys Trp Ile Ser Met Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu  
 225 230 235 240  
**35** Asn Thr Gln Thr Leu Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr  
 245 250 255  
**40** Phe Glu Glu Val Ala Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val  
 260 265 270  
 Asp Gly Val Asn His Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile  
 275 280 285  
**45** Arg Asn Cys Lys Lys Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln  
 290 295 300  
**50** Val Gly Gly Ser Asp Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala  
 305 310 315 320  
 Pro Gln Thr Glu Arg Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln  
 325 330 335  
**55** Val Phe Tyr Thr Val Val Asp Tyr Val Asn Gln Ile Ile Gln Ala Met  
 340 345 350  
 Ser Lys Arg Ser Arg Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 355 360 365  
**60**

ES 2 733 279 T3

<210> 34  
 <211> 365  
 <212> PRT  
 <213> Secuencia Artificial

5  
 <220>  
 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por ingeniería

10  
 <400> 34  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30

15  
 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45

20  
 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60

25  
 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80

30  
 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95

35  
 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110

40  
 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125

45  
 Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140

50  
 Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160

55  
 Gln Arg Leu Leu Asp Thr Val Ser Gly Gly Glu Asp Gln Val Asp Pro  
 165 170 175

60  
 Arg Leu Ile Asp Gly Lys Ser Gly Gly Thr Leu Gln Leu Asp Met Ser  
 180 185 190

65  
 Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp  
 195 200 205

70  
 Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile  
 210 215 220

75  
 Ser Met Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln  
 225 230 235 240

80  
 Thr Leu Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu  
 245 250 255

85  
 Val Ala Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val  
 260 265 270

ES 2 733 279 T3

Asn His Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys  
 275 280 285

5 Lys Lys Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly  
 290 295 300

10 Ser Asp Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr  
 305 310 315 320

Glu Arg Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr  
 325 330 335

15 Thr Val Val Asp Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg  
 340 345 350

Ser Arg Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 355 360 365

20 <210> 35  
 <211> 363  
 <212> PRT  
 <213> Secuencia Artificial

25 <220>  
 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por  
 ingeniería

<400> 35

30 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15

Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30

35 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45

40 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60

Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80

45 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95

50 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110

Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125

55 Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140

Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160

60

ES 2 733 279 T3

Gln Arg Leu Leu Asp Thr Val Ser Gly Gly Glu Asp Gln Val Asp Pro  
 165 170 175  
 Arg Leu Ile Asp Gly Lys Gly Thr Leu Gln Leu Asp Met Ser Glu Leu  
 180 185 190  
**5**  
 Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr  
 195 200 205  
**10**  
 Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met  
 210 215 220  
 Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu  
 225 230 235 240  
**15**  
 Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala  
 245 250 255  
**20**  
 Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His  
 260 265 270  
 Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys  
 275 280 285  
**25**  
 Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp  
 290 295 300  
 Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg  
 305 310 315 320  
**30**  
 Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val  
 325 330 335  
 Val Asp Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg  
 340 345 350  
**35**  
 Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 355 360  
**40**  
 <210> 36  
 <211> 361  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
**45**  
 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por  
 ingeniería  
 <400> 36  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
**50**  
 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30  
**55**  
 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45  
**60**  
 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60

ES 2 733 279 T3

Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
 5 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
 10 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
 Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140  
 15 Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160  
 20 Gln Arg Leu Leu Asp Thr Val Gly Glu Asp Gln Val Asp Pro Arg Leu  
 165 170 175  
 Ile Asp Gly Lys Gly Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp  
 180 185 190  
 25 Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr  
 195 200 205  
 Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser  
 210 215 220  
 30 Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu Gly Ile  
 225 230 235 240  
 35 Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala Thr Ala  
 245 250 255  
 Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His Lys Leu  
 260 265 270  
 40 Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly  
 275 280 285  
 Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp Val Leu  
 290 295 300  
 45 Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg Met Met  
 305 310 315 320  
 50 Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val Val Asp  
 325 330 335  
 Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg Ser Leu  
 340 345 350  
 55 Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 355 360  
 60 <210> 37  
 <211> 379

ES 2 733 279 T3

<212> PRT  
 <213> Secuencia Artificial

<220>

5 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por ingeniería

<400> 37

10 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15

Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30

15 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45

Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60

20 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80

25 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95

Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110

30 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125

Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140

35 Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160

40 Gln Arg Leu Leu Asp Thr Val Asn Glu Gln Glu Leu Leu Glu Leu Asp  
 165 170 175

Lys Trp Ala Ser Leu Ser Gly Gly Glu Asp Gln Val Asp Pro Arg Leu  
 180 185 190

45 Ile Asp Gly Lys Ser Gly Gly Thr Leu Gln Leu Asp Met Ser Glu Leu  
 195 200 205

Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr  
 210 215 220

50 Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met  
 225 230 235 240

55 Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu  
 245 250 255

Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala  
 260 265 270

60

ES 2 733 279 T3

Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His  
 275 280 285  
 5 Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys  
 290 295 300  
 Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp  
 305 310 315 320  
 10 Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg  
 325 330 335  
 Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val  
 340 345 350  
 15 Val Asp Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg  
 355 360 365  
 20 Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 370 375  
 <210> 38  
 <211> 376  
 <212> PRT  
 25 <213> Secuencia Artificial  
 <220>  
 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por  
 ingeniería  
 30 <400> 38  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 35 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30  
 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45  
 40 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60  
 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
 45 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 50 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
 55 Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140  
 60 Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160



ES 2 733 279 T3

Gln Arg Leu Leu Asp Thr Val Lys Asn Glu Gln Glu Leu Leu Glu Leu  
165 170 175

5 Asp Lys Trp Ala Ser Leu Gly Glu Asp Gln Val Asp Pro Arg Leu Ile  
180 185 190

Asp Gly Lys Gly Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp Leu  
195 200 205

10 Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr Tyr  
210 215 220

15 Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser Ser  
225 230 235 240

Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu Gly Ile Gly  
245 250 255

20 Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala Thr Ala Glu  
260 265 270

Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His Lys Leu Asp  
275 280 285

25 Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly Pro  
290 295 300

30 Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp Val Leu Asp  
305 310 315 320

Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg Met Met Arg  
325 330 335

35 Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val Val Asp Tyr  
340 345 350

Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg Ser Leu Asn  
355 360 365

40 Ser Ala Ala Phe Tyr Tyr Arg Ile  
370 375

45 <210> 39  
<211> 375  
<212> PRT  
<213> Secuencia Artificial

<220>  
50 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por  
ingeniería

<400> 39

55 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
1 5 10 15

Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
20 25 30

60

ES 2 733 279 T3

Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45  
 5 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60  
 Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
 10 Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
 15 Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
 Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140  
 20 Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160  
 25 Gln Arg Leu Leu Asp Thr Val Gly Leu Glu Leu Asp Lys Trp Ala Ser  
 165 170 175  
 Leu Ser Gly Gly Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys  
 180 185 190  
 30 Ser Gly Gly Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp Leu Ile  
 195 200 205  
 Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr  
 210 215 220  
 35 Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser Ser Cys  
 225 230 235 240  
 Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu Gly Ile Gly Cys  
 245 250 255  
 40 Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala Thr Ala Glu Lys  
 260 265 270  
 45 Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His Lys Leu Asp Val  
 275 280 285  
 Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly Pro Arg  
 290 295 300  
 50 Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp Val Leu Asp Ile  
 305 310 315 320  
 Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg Met Met Arg Ile  
 325 330 335  
 55 Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val Val Asp Tyr Val  
 340 345 350  
 Asn Gln Ile Ile Gln Ala Met Ser Lys Arg Ser Arg Ser Leu Asn Ser  
 355 360 365  
 60

ES 2 733 279 T3

Ala Ala Phe Tyr Tyr Arg Ile  
 370 375

5 <210> 40  
 <211> 371  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por ingeniería

<400> 40

15 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15

Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 20 25 30

20 Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45

Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 25 50 55 60

Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 30 65 70 75 80

Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95

Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 35 100 105 110

Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125

40 Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140

Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 45 145 150 155 160

Gln Arg Leu Leu Asp Thr Val Gly Leu Glu Leu Asp Lys Trp Ala Ser  
 165 170 175

Leu Gly Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys Gly Thr  
 50 180 185 190

Leu Gln Leu Asp Met Ser Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp  
 195 200 205

55 Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr Asp  
 210 215 220

Glu Ala Asn Lys Trp Ile Ser Met Gly Ser Ser Cys Thr Ile Lys Val  
 60 225 230 235 240

ES 2 733 279 T3

Cys Pro Leu Asn Thr Gln Thr Leu Gly Ile Gly Cys Leu Thr Thr Asp  
 245 250 255  
**5** Thr Ala Thr Phe Glu Glu Val Ala Thr Ala Glu Lys Leu Val Ile Thr  
 260 265 270  
 Asp Val Val Asp Gly Val Asn His Lys Leu Asp Val Thr Thr Ala Thr  
 275 280 285  
**10** Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly Pro Arg Glu Asn Val Ala  
 290 295 300  
**15** Val Ile Gln Val Gly Gly Ser Asp Val Leu Asp Ile Thr Ala Asp Pro  
 305 310 315 320  
 Thr Thr Ala Pro Gln Thr Glu Arg Met Met Arg Ile Asn Trp Lys Lys  
 325 330 335  
**20** Trp Trp Gln Val Phe Tyr Thr Val Val Asp Tyr Val Asn Gln Ile Ile  
 340 345 350  
 Gln Ala Met Ser Lys Arg Ser Arg Ser Leu Asn Ser Ala Ala Phe Tyr  
 355 360 365  
**25** Tyr Arg Ile  
 370  
**30** <210> 41  
 <211> 381  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
**35** <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por  
 ingeniería  
 <400> 41  
**40** Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 Ser Tyr Ile Tyr Ala Glu Asp Ser Ala Trp Ser His Pro Gln Phe Glu  
 20 25 30  
**45** Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn Leu Pro  
 35 40 45  
 Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln Glu Glu  
 50 55 60  
**50** Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ala  
 65 70 75 80  
**55** Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln Leu Phe  
 85 90 95  
 Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr  
 100 105 110  
**60**

ES 2 733 279 T3

Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn  
 115 120 125  
 5 Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln Ile Ser  
 130 135 140  
 Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala  
 145 150 155 160  
 10 Gln Arg Leu Leu Asp Thr Val Gly Leu Glu Leu Asp Lys Trp Ala Ser  
 165 170 175  
 Leu Gly Ser Ile Glu Gly Arg Ser Gly Gly Glu Asp Gln Val Asp Pro  
 180 185 190  
 15 Arg Leu Ile Asp Gly Lys Ser Gly Gly Thr Leu Gln Leu Asp Met Ser  
 195 200 205  
 Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp  
 210 215 220  
 20 Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile  
 225 230 235 240  
 Ser Met Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln  
 245 250 255  
 25 Thr Leu Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu  
 260 265 270  
 30 Val Ala Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val  
 275 280 285  
 Asn His Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys  
 290 295 300  
 35 Lys Lys Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly  
 305 310 315 320  
 Ser Asp Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr  
 325 330 335  
 40 Glu Arg Met Met Arg Ile Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr  
 340 345 350  
 45 Thr Val Val Asp Tyr Val Asn Gln Ile Ile Gln Ala Met Ser Lys Arg  
 355 360 365  
 Ser Arg Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Ile  
 370 375 380  
 50 <210> 42  
 <211> 369  
 <212> PRT  
 <213> Secuencia Artificial  
 55 <220>  
 <223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
 ingeniería  
 60 <400> 42

ES 2 733 279 T3

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly  
 1 5 10 15

5 Ala Val Phe Val Ser Pro Ser Glu Asp Ser Ala Trp Ser His Pro Gln  
 20 25 30

Phe Glu Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn  
 35 40 45

10 Leu Pro Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln  
 50 55 60

Glu Glu Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu  
 65 70 75 80

15 Ala Ala Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln  
 85 90 95

20 Leu Phe Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu  
 100 105 110

Tyr Thr Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp  
 115 120 125

25 Tyr Asn Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp  
 130 135 140

Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu  
 145 150 155 160

30 Ser Gln Asn Gln Gln Glu Lys Ile Glu Gly Arg Ser Gly Gly Glu Asp  
 165 170 175

Gln Val Asp Pro Arg Leu Ile Asp Gly Lys Ser Gly Gly Thr Leu Gln  
 180 185 190

35 Leu Asp Met Ser Glu Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys  
 195 200 205

40 Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala  
 210 215 220

Asn Lys Trp Ile Ser Met Gly Ser Ser Cys Thr Ile Lys Val Cys Pro  
 225 230 235 240

45 Leu Asn Thr Gln Thr Leu Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala  
 245 250 255

Thr Phe Glu Glu Val Ala Thr Ala Glu Lys Leu Val Ile Thr Asp Val  
 260 265 270

50 Val Asp Gly Val Asn His Lys Leu Asp Val Thr Thr Ala Thr Cys Thr  
 275 280 285

55 Ile Arg Asn Cys Lys Lys Leu Gly Pro Arg Glu Asn Val Ala Val Ile  
 290 295 300

Gln Val Gly Gly Ser Asp Val Leu Asp Ile Thr Ala Asp Pro Thr Thr  
 305 310 315 320

60

ES 2 733 279 T3

Ala Pro Gln Thr Glu Arg Met Met Arg Ile Asn Trp Lys Lys Trp Trp  
 325 330 335  
 Gln Val Phe Tyr Thr Val Val Asp Tyr Val Asn Gln Ile Ile Gln Ala  
 340 345 350

5 Met Ser Lys Arg Ser Arg Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg  
 355 360 365  
 Ile

10 <210> 43  
 <211> 361  
 <212> PRT  
 <213> Secuencia Artificial

15 <220>  
 <223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por  
 ingeniería

20 <400> 43  
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly  
 1 5 10 15

25 Ala Val Phe Val Ser Pro Ser Glu Asp Ser Ala Trp Ser His Pro Gln  
 20 25 30

Phe Glu Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn  
 35 40 45

30 Leu Pro Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln  
 50 55 60

Glu Glu Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu  
 65 70 75 80

35 Ala Ala Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln  
 85 90 95

40 Leu Phe Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu  
 100 105 110

Tyr Thr Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp  
 115 120 125

45 Tyr Asn Val Val Leu Met Lys Tyr Asp Ala Gly Trp Met Glu Trp Asp  
 130 135 140

Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu  
 145 150 155 160

50 Ser Gln Asn Gln Gln Glu Lys Gly Glu Asp Gln Val Asp Pro Arg Leu  
 165 170 175

Ile Asp Gly Lys Gly Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp  
 180 185 190

55 Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr  
 195 200 205

60 Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser

ES 2 733 279 T3

	210		215		220														
	Ser	Cys	Thr	Ile	Lys	Val	Cys	Pro	Leu	Asn	Thr	Gln	Thr	Leu	Gly	Ile			
	225					230					235					240			
5	Gly	Cys	Leu	Thr	Thr	Asp	Thr	Ala	Thr	Phe	Glu	Glu	Val	Ala	Thr	Ala			
					245					250						255			
10	Glu	Lys	Leu	Val	Ile	Thr	Asp	Val	Val	Asp	Gly	Val	Asn	His	Lys	Leu			
				260					265					270					
	Asp	Val	Thr	Thr	Ala	Thr	Cys	Thr	Ile	Arg	Asn	Cys	Lys	Lys	Leu	Gly			
			275					280					285						
15	Pro	Arg	Glu	Asn	Val	Ala	Val	Ile	Gln	Val	Gly	Gly	Ser	Asp	Val	Leu			
		290					295					300							
	Asp	Ile	Thr	Ala	Asp	Pro	Thr	Thr	Ala	Pro	Gln	Thr	Glu	Arg	Met	Met			
	305					310					315					320			
20	Arg	Ile	Asn	Trp	Lys	Lys	Trp	Trp	Gln	Val	Phe	Tyr	Thr	Val	Val	Asp			
					325					330						335			
25	Tyr	Val	Asn	Gln	Ile	Ile	Gln	Ala	Met	Ser	Lys	Arg	Ser	Arg	Ser	Leu			
				340					345					350					
	Asn	Ser	Ala	Ala	Phe	Tyr	Tyr	Arg	Ile										
			355					360											
30	<210>	44																	
	<211>	372																	
	<212>	PRT																	
	<213>	Secuencia Artificial																	
35	<220>																		
	<223>	Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por ingeniería																	
	<400>	44																	
40	Met	Asp	Ala	Met	Lys	Arg	Gly	Leu	Cys	Cys	Val	Leu	Leu	Leu	Cys	Gly			
	1				5					10					15				
	Ala	Val	Phe	Val	Ser	Pro	Ser	Glu	Asp	Ser	Ala	Trp	Ser	His	Pro	Gln			
				20					25					30					
45	Phe	Glu	Lys	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Gln	Asn	Tyr	Gly	Ile	Asn			
			35					40					45						
50	Leu	Pro	Ile	Thr	Gly	Ser	Met	Asp	Thr	Ala	Tyr	Ala	Asn	Ser	Thr	Gln			
		50					55					60							
	Glu	Glu	Thr	Phe	Leu	Thr	Ser	Thr	Leu	Cys	Leu	Tyr	Tyr	Pro	Thr	Glu			
	65					70					75					80			
55	Ala	Ala	Thr	Glu	Ile	Asn	Asp	Asn	Ser	Trp	Lys	Asp	Thr	Leu	Ser	Gln			
					85					90					95				
	Leu	Phe	Leu	Thr	Lys	Gly	Trp	Pro	Thr	Gly	Ser	Val	Tyr	Phe	Lys	Glu			
				100					105					110					
60																			



ES 2 733 279 T3

Tyr Thr Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp  
 115 120 125

**5** Tyr Asn Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln  
 130 135 140

Ile Ser Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys  
 145 150 155 160

**10** Glu Ala Gln Arg Leu Leu Asp Thr Val Lys Ile Glu Gly Arg Ser Gly  
 165 170 175

**15** Gly Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys Ser Gly Gly  
 180 185 190

Thr Leu Gln Leu Asp Met Ser Glu Leu Ala Asp Leu Ile Leu Asn Glu  
 195 200 205

**20** Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr Gln Gln Thr  
 210 215 220

Asp Glu Ala Asn Lys Trp Ile Ser Met Gly Ser Ser Cys Thr Ile Lys  
 225 230 235 240

**25** Val Cys Pro Leu Asn Thr Gln Thr Leu Gly Ile Gly Cys Leu Thr Thr  
 245 250 255

**30** Asp Thr Ala Thr Phe Glu Glu Val Ala Thr Ala Glu Lys Leu Val Ile  
 260 265 270

Thr Asp Val Val Asp Gly Val Asn His Lys Leu Asp Val Thr Thr Ala  
 275 280 285

**35** Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly Pro Arg Glu Asn Val  
 290 295 300

Ala Val Ile Gln Val Gly Gly Ser Asp Val Leu Asp Ile Thr Ala Asp  
 305 310 315 320

**40** Pro Thr Thr Ala Pro Gln Thr Glu Arg Met Met Arg Ile Asn Trp Lys  
 325 330 335

**45** Lys Trp Trp Gln Val Phe Tyr Thr Val Val Asp Tyr Val Asn Gln Ile  
 340 345 350

Ile Gln Ala Met Ser Lys Arg Ser Arg Ser Leu Asn Ser Ala Ala Phe  
 355 360 365

**50** Tyr Tyr Arg Ile  
 370

**55** <210> 45  
 <211> 363  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Proteína de fusión VP7-HR2Nipah del Rotavirus Rhesus producido por  
 ingeniería

**60**

ES 2 733 279 T3

<400> 45  
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly  
 1 5 10 15  
 5 Ala Val Phe Val Ser Pro Ser Glu Asp Ser Ala Trp Ser His Pro Gln  
 20 25 30  
 10 Phe Glu Lys Glu Asn Leu Tyr Phe Gln Gly Gln Asn Tyr Gly Ile Asn  
 35 40 45  
 15 Leu Pro Ile Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Thr Gln  
 50 55 60  
 20 Glu Glu Thr Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu  
 65 70 75 80  
 Ala Ala Thr Glu Ile Asn Asp Asn Ser Trp Lys Asp Thr Leu Ser Gln  
 85 90 95  
 25 Leu Phe Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu  
 100 105 110  
 Tyr Thr Asp Ile Ala Ser Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp  
 115 120 125  
 30 Tyr Asn Val Val Leu Met Lys Tyr Asp Ala Gly Asp Ile Ser Ser Gln  
 130 135 140  
 Ile Ser Ser Met Asn Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys  
 145 150 155 160  
 35 Glu Ala Gln Arg Leu Leu Asp Thr Val Gly Glu Asp Gln Val Asp Pro  
 165 170 175  
 Arg Leu Ile Asp Gly Lys Gly Thr Leu Gln Leu Asp Met Ser Glu Leu  
 180 185 190  
 40 Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr  
 195 200 205  
 Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met  
 210 215 220  
 45 Gly Ser Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu  
 225 230 235 240  
 Gly Ile Gly Cys Leu Thr Thr Asp Thr Ala Thr Phe Glu Glu Val Ala  
 245 250 255  
 50 Thr Ala Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His  
 260 265 270  
 55 Lys Leu Asp Val Thr Thr Ala Thr Cys Thr Ile Arg Asn Cys Lys Lys  
 275 280 285  
 Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asp  
 290 295 300  
 60 Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg



ES 2 733 279 T3

Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
 210 215 220  
**5** Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
 225 230 235 240  
 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
 245 250 255  
**10** Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
 260 265 270  
**15** Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
 275 280 285  
 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
 290 295 300  
**20** Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
 305 310 315 320  
 Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
 325 330 335  
**25** Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350  
**30** Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365  
 Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Asp Asp Asp  
 370 375 380  
**35** Asp Lys Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp  
 385 390 395 400  
 Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln  
 405 410 415  
**40** Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn  
 420 425 430  
**45** Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln  
 435 440 445  
 Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu  
 450 455 460  
**50** Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr  
 465 470 475 480  
 Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe  
 485 490 495  
**55** His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu  
 500 505 510  
**60** Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His  
 515 520 525

ES 2 733 279 T3

Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp  
530 535 540

5 Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp  
545 550 555 560

Gly Val Arg Ser Leu Gln Glu Ala Gly Tyr Ile Pro Glu Ala Pro Arg  
565 570 575

10 Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser  
580 585 590

Thr Phe Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala  
595 600 605

15 Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys  
610 615 620

Gln Leu Gln Ala Arg  
20 625

<210> 47  
<211> 627  
<212> PRT  
25 <213> Secuencia Artificial

<220>  
<223> Proteína fusión ectodominio-HR1Nipah de hemaglutinina del virus de  
influenza (H1 Islas Salomón 2006) producido por ingeniería

30 <400> 47

Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
1 5 10 15

35 Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
20 25 30

40 Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
35 40 45

Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
50 55 60

45 Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
65 70 75 80

Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
85 90 95

50 Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
100 105 110

55 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
115 120 125

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

60 Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu

ES 2 733 279 T3

	145				150					155				160		
	Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe	Glu	Arg	Phe	Glu	Ile	Phe
					165					170				175		
5	Pro	Lys	Glu	Ser	Ser	Trp	Pro	Asn	His	Thr	Thr	Thr	Gly	Val	Ser	Ala
				180					185					190		
10	Ser	Cys	Ser	His	Asn	Gly	Glu	Ser	Ser	Phe	Tyr	Lys	Asn	Leu	Leu	Trp
			195					200					205			
	Leu	Thr	Gly	Lys	Asn	Gly	Leu	Tyr	Pro	Asn	Leu	Ser	Lys	Ser	Tyr	Ala
		210					215					220				
15	Asn	Asn	Lys	Glu	Lys	Glu	Val	Leu	Val	Leu	Trp	Gly	Val	His	His	Pro
	225					230					235					240
20	Pro	Asn	Ile	Gly	Asp	Gln	Arg	Ala	Leu	Tyr	His	Thr	Glu	Asn	Ala	Tyr
					245					250					255	
	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg	Lys	Phe	Thr	Pro	Glu	Ile
				260					265					270		
25	Ala	Lys	Arg	Pro	Lys	Val	Arg	Asp	Arg	Glu	Gly	Arg	Ile	Asn	Tyr	Tyr
			275					280					285			
30	Trp	Thr	Leu	Leu	Glu	Pro	Gly	Asp	Thr	Ile	Ile	Phe	Glu	Ala	Asn	Gly
	290						295					300				
	Asn	Leu	Ile	Ala	Pro	Arg	Tyr	Ala	Phe	Ala	Leu	Ser	Arg	Gly	Phe	Gly
	305					310					315					320
35	Ser	Gly	Ile	Ile	Asn	Ser	Asn	Ala	Pro	Met	Asp	Glu	Cys	Asp	Ala	Lys
					325					330					335	
	Cys	Gln	Thr	Pro	Gln	Gly	Ala	Ile	Asn	Ser	Ser	Leu	Pro	Phe	Gln	Asn
				340					345					350		
40	Val	His	Pro	Val	Thr	Ile	Gly	Glu	Cys	Pro	Lys	Tyr	Val	Arg	Ser	Ala
			355					360					365			
45	Lys	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn	Ile	Pro	Ser	Asp	Asp	Asp
	370						375					380				
	Asp	Lys	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Gly	Gly	Trp
	385					390					395					400
50	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Gln	Asn	Glu	Gln
					405					410					415	
	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Gln	Lys	Ser	Thr	Gln	Asn	Ala	Ile	Asn
				420					425					430		
55	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	Met	Asn	Thr	Gln
			435					440					445			
60	Phe	Thr	Ala	Val	Gly	Lys	Glu	Phe	Asn	Lys	Leu	Glu	Arg	Arg	Met	Glu
		450					455					460				

ES 2 733 279 T3

Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr  
 465 470 475 480  
**5** Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe  
 485 490 495  
 His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu  
 500 505 510  
**10** Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His  
 515 520 525  
 Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp  
 530 535 540  
**15** Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp  
 545 550 555 560  
 Gly Val Arg Ser Leu Gln Glu Ala Gly Tyr Ile Pro Glu Ala Pro Arg  
 565 570 575  
**20** Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser  
 580 585 590  
 Thr Asn Ile Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn Glu Ala  
 595 600 605  
**25** Val Val Lys Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val Leu Thr  
 610 615 620  
**30** Ala Leu Gln  
 625  
 <210> 48  
 <211> 596  
 <212> PRT  
**35** <213> Secuencia Artificial  
 <220>  
 <223> Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de  
 influenza (H1 Islas Salomón 2006) producido por ingeniería  
**40** <400> 48  
 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
 1 5 10 15  
**45** Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
 20 25 30  
 Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
 35 40 45  
**50** Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
 50 55 60  
**55** Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
 65 70 75 80  
 Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
 85 90 95  
**60** Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly

ES 2 733 279 T3

				100					105					110			
	Asn	Cys	Ser	Val	Ala	Gly	Trp	Ile	Leu	Gly	Asn	Pro	Glu	Cys	Glu	Leu	
			115					120					125				
5	Leu	Ile	Ser	Arg	Glu	Ser	Trp	Ser	Tyr	Ile	Val	Glu	Lys	Pro	Asn	Pro	
		130					135					140					
	Glu	Asn	Gly	Thr	Cys	Tyr	Pro	Gly	His	Phe	Ala	Asp	Tyr	Glu	Glu	Leu	
10	145					150					155					160	
	Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe	Glu	Arg	Phe	Glu	Ile	Phe	
					165					170					175		
15	Pro	Lys	Glu	Ser	Ser	Trp	Pro	Asn	His	Thr	Thr	Thr	Gly	Val	Ser	Ala	
					180				185					190			
	Ser	Cys	Ser	His	Asn	Gly	Glu	Ser	Ser	Phe	Tyr	Lys	Asn	Leu	Leu	Trp	
20			195					200					205				
	Leu	Thr	Gly	Lys	Asn	Gly	Leu	Tyr	Pro	Asn	Leu	Ser	Lys	Ser	Tyr	Ala	
		210					215					220					
25	Asn	Asn	Lys	Glu	Lys	Glu	Val	Leu	Val	Leu	Trp	Gly	Val	His	His	Pro	
	225					230					235					240	
	Pro	Asn	Ile	Gly	Asp	Gln	Arg	Ala	Leu	Tyr	His	Thr	Glu	Asn	Ala	Tyr	
					245					250					255		
30	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg	Lys	Phe	Thr	Pro	Glu	Ile	
				260					265					270			
	Ala	Lys	Arg	Pro	Lys	Val	Arg	Asp	Arg	Glu	Gly	Arg	Ile	Asn	Tyr	Tyr	
			275					280					285				
35	Trp	Thr	Leu	Leu	Glu	Pro	Gly	Asp	Thr	Ile	Ile	Phe	Glu	Ala	Asn	Gly	
		290					295					300					
	Asn	Leu	Ile	Ala	Pro	Arg	Tyr	Ala	Phe	Ala	Leu	Ser	Arg	Gly	Phe	Gly	
40	305					310					315					320	
	Ser	Gly	Ile	Ile	Asn	Ser	Asn	Ala	Pro	Met	Asp	Glu	Cys	Asp	Ala	Lys	
					325					330					335		
45	Cys	Gln	Thr	Pro	Gln	Gly	Ala	Ile	Asn	Ser	Ser	Leu	Pro	Phe	Gln	Asn	
				340					345					350			
	Val	His	Pro	Val	Thr	Ile	Gly	Glu	Cys	Pro	Lys	Tyr	Val	Arg	Ser	Ala	
50			355					360					365				
	Lys	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn	Ile	Pro	Ser	Asp	Asp	Asp	
		370					375					380					
55	Asp	Lys	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Gly	Gly	Trp	
	385					390					395					400	
	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Gln	Asn	Glu	Gln	
					405					410					415		
60	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Gln	Lys	Ser	Thr	Gln	Asn	Ala	Ile	Asn	



ES 2 733 279 T3

				420					425					430			
	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	Met	Asn	Thr	Gln	
			435					440					445				
5	Phe	Thr	Ala	Val	Gly	Lys	Glu	Phe	Asn	Lys	Leu	Glu	Arg	Arg	Met	Glu	
		450					455					460					
10	Asn	Leu	Asn	Lys	Lys	Val	Asp	Asp	Gly	Phe	Ile	Asp	Ile	Trp	Thr	Tyr	
	465					470					475					480	
15	Asn	Ala	Glu	Leu	Leu	Val	Leu	Leu	Glu	Asn	Glu	Arg	Thr	Leu	Asp	Phe	
					485					490					495		
	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	Glu	Lys	Val	Lys	Ser	Gln	Leu	
				500					505					510			
20	Lys	Asn	Asn	Ala	Lys	Glu	Ile	Gly	Asn	Gly	Cys	Phe	Glu	Phe	Tyr	His	
			515					520					525				
	Lys	Cys	Asn	Asp	Glu	Cys	Met	Glu	Ser	Val	Lys	Asn	Gly	Thr	Tyr	Asp	
		530					535					540					
25	Tyr	Pro	Lys	Tyr	Ser	Glu	Glu	Ser	Lys	Leu	Asn	Arg	Glu	Lys	Ile	Asp	
	545					550					555					560	
30	Gly	Val	Ser	Gly	Ile	Val	Gln	Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	
				565						570					575		
	Glu	Ala	Gln	Gln	His	Leu	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	
				580					585					590			
35	Leu	Gln	Ala	Arg													
			595														
	<210>	49															
	<211>	598															
	<212>	PRT															
	<213>	Secuencia Artificial															
40	<220>																
	<223>	Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de influenza (H1 Islas Salomón 2006) producido por ingeniería															
45	<400>	49															
	Met	Leu	Leu	Val	Asn	Gln	Ser	His	Gln	Gly	Phe	Asn	Lys	Glu	His	Thr	
	1			5					10					15			
50	Ser	Lys	Met	Val	Ser	Ala	Ile	Val	Leu	Tyr	Val	Leu	Leu	Ala	Ala	Ala	
				20					25					30			
	Ala	His	Ser	Ala	Phe	Ala	Ala	Asp	Pro	His	His	His	His	His	His	Glu	
			35					40					45				
55	Asn	Leu	Tyr	Phe	Gln	Gly	Gly	Tyr	Leu	Leu	Glu	Asp	Thr	Ile	Cys	Ile	
	50					55						60					
60	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr	Val	Asp	Thr	Val	Leu	Glu	
	65					70					75					80	

ES 2 733 279 T3

Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
 85 90 95  
 5 Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
 100 105 110  
 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
 115 120 125  
 10 Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val Glu Lys Pro Asn Pro  
 130 135 140  
 15 Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
 145 150 155 160  
 Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
 165 170 175  
 20 Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
 180 185 190  
 Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
 195 200 205  
 Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
 210 215 220  
 30 Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
 225 230 235 240  
 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
 245 250 255  
 35 Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
 260 265 270  
 Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
 275 280 285  
 40 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
 290 295 300  
 45 Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
 305 310 315 320  
 Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
 325 330 335  
 50 Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350  
 Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365  
 55 Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Asp Asp Asp  
 370 375 380  
 60 Asp Lys Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp



ES 2 733 279 T3

Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
 50 55 60  
**5** Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
 65 70 75 80  
 Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
 85 90 95  
**10** Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
 100 105 110  
 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
 115 120 125  
**15** Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val Glu Lys Pro Asn Pro  
 130 135 140  
**20** Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
 145 150 155 160  
 Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
 165 170 175  
**25** Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
 180 185 190  
**30** Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
 195 200 205  
 Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
 210 215 220  
**35** Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
 225 230 235 240  
 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
 245 250 255  
**40** Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
 260 265 270  
**45** Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
 275 280 285  
 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
 290 295 300  
**50** Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
 305 310 315 320  
 Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
 325 330 335  
**55** Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350  
**60** Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365

ES 2 733 279 T3

Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser  
 370 375 380

**5** Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr  
 385 390 395 400

Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly  
 405 410 415

**10** Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly  
 420 425 430

**15** Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe  
 435 440 445

Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn  
 450 455 460

**20** Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr Asn  
 465 470 475 480

Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His  
 485 490 495

**25** Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys  
 500 505 510

Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys  
 515 520 525

**30** Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr  
 530 535 540

Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly  
 545 550 555 560

Val Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu  
 565 570 575

**40** Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu  
 580 585 590

Gln Ala Arg  
 595

**45** <210> 51  
 <211> 597  
 <212> PRT  
 <213> Secuencia Artificial

**50** <220>  
 <223> Proteína fusión ectodominio-HR1Nipah de hemaglutinina del virus de  
 influenza (H1 Islas Salomón 2006) producido por ingeniería

**55** <400> 51  
 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
 1 5 10 15

**60** Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
 20 25 30

ES 2 733 279 T3

Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
35 40 45

5 Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
50 55 60

Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
65 70 75 80

10 Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
85 90 95

15 Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
100 105 110

Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
115 120 125

20 Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val Glu Lys Pro Asn Pro  
130 135 140

Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
145 150 155 160

25 Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
165 170 175

30 Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
180 185 190

Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
195 200 205

35 Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
210 215 220

40 Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
225 230 235 240

Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
245 250 255

45 Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
260 265 270

Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
275 280 285

50 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
290 295 300

55 Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
305 310 315 320

Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
325 330 335

60

ES 2 733 279 T3

Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350  
 5 Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365  
 Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser  
 370 375 380  
 10 Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr  
 385 390 395 400  
 Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly  
 405 410 415  
 15 Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly  
 420 425 430  
 20 Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe  
 435 440 445  
 Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn  
 450 455 460  
 25 Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr Asn  
 465 470 475 480  
 Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His  
 485 490 495  
 30 Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys  
 500 505 510  
 35 Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys  
 515 520 525  
 Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr  
 530 535 540  
 Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly  
 545 550 555 560  
 40 Val Arg Ser Asn Ile Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn  
 565 570 575  
 45 Glu Ala Val Val Lys Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val  
 580 585 590  
 Leu Thr Ala Leu Gln  
 595  
 50 <210> 52  
 <211> 595  
 <212> PRT  
 <213> Secuencia Artificial  
 55 <220>  
 <223> Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de  
 influenza (H1 Islas Salomón 2006) producido por ingeniería  
 <400> 52  
 60 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr

ES 2 733 279 T3

	1			5					10					15			
	Ser	Lys	Met	Val	Ser	Ala	Ile	Val	Leu	Tyr	Val	Leu	Leu	Ala	Ala	Ala	
				20					25					30			
5	Ala	His	Ser	Ala	Phe	Ala	Ala	Asp	Pro	His	His	His	His	His	His	Glu	
			35					40					45				
10	Asn	Leu	Tyr	Phe	Gln	Gly	Gly	Tyr	Leu	Leu	Glu	Asp	Thr	Ile	Cys	Ile	
		50					55					60					
15	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr	Val	Asp	Thr	Val	Leu	Glu	
		65				70					75					80	
20	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn	Leu	Leu	Glu	Asp	Ser	His	
					85					90					95		
25	Asn	Gly	Lys	Leu	Cys	Leu	Leu	Lys	Gly	Ile	Ala	Pro	Leu	Gln	Leu	Gly	
				100					105					110			
30	Asn	Cys	Ser	Val	Ala	Gly	Trp	Ile	Leu	Gly	Asn	Pro	Glu	Cys	Glu	Leu	
			115					120					125				
35	Leu	Ile	Ser	Arg	Glu	Ser	Trp	Ser	Tyr	Ile	Val	Glu	Lys	Pro	Asn	Pro	
		130					135					140					
40	Glu	Asn	Gly	Thr	Cys	Tyr	Pro	Gly	His	Phe	Ala	Asp	Tyr	Glu	Glu	Leu	
		145				150					155					160	
45	Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe	Glu	Arg	Phe	Glu	Ile	Phe	
					165					170					175		
50	Pro	Lys	Glu	Ser	Ser	Trp	Pro	Asn	His	Thr	Thr	Thr	Gly	Val	Ser	Ala	
					180				185					190			
55	Ser	Cys	Ser	His	Asn	Gly	Glu	Ser	Ser	Phe	Tyr	Lys	Asn	Leu	Leu	Trp	
			195					200					205				
60	Leu	Thr	Gly	Lys	Asn	Gly	Leu	Tyr	Pro	Asn	Leu	Ser	Lys	Ser	Tyr	Ala	
			210				215					220					
65	Asn	Asn	Lys	Glu	Lys	Glu	Val	Leu	Val	Leu	Trp	Gly	Val	His	His	Pro	
						230					235					240	
70	Pro	Asn	Ile	Gly	Asp	Gln	Arg	Ala	Leu	Tyr	His	Thr	Glu	Asn	Ala	Tyr	
					245					250					255		
75	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg	Lys	Phe	Thr	Pro	Glu	Ile	
					260				265					270			
80	Ala	Lys	Arg	Pro	Lys	Val	Arg	Asp	Arg	Glu	Gly	Arg	Ile	Asn	Tyr	Tyr	
			275					280					285				
85	Trp	Thr	Leu	Leu	Glu	Pro	Gly	Asp	Thr	Ile	Ile	Phe	Glu	Ala	Asn	Gly	
							295					300					
90	Asn	Leu	Ile	Ala	Pro	Arg	Tyr	Ala	Phe	Ala	Leu	Ser	Arg	Gly	Phe	Gly	
						310					315					320	



ES 2 733 279 T3

Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
 325 330 335  
**5** Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350  
 Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365  
**10** Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Glu Gly  
 370 375 380  
**15** Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr  
 385 390 395 400  
 Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly  
 405 410 415  
**20** Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly  
 420 425 430  
 Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe  
 435 440 445  
**25** Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn  
 450 455 460  
**30** Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr Asn  
 465 470 475 480  
 Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His  
 485 490 495  
**35** Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys  
 500 505 510  
**40** Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys  
 515 520 525  
 Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr  
 530 535 540  
**45** Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly  
 545 550 555 560  
 Val Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu  
 565 570 575  
**50** Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu  
 580 585 590  
 Gln Ala Arg  
 595  
**55** <210> 53  
 <211> 597  
 <212> PRT  
 <213> Secuencia Artificial  
**60** <220>  
 <223> Proteína fusión ectodominio-HR1Nipah de hemaglutinina del virus de

ES 2 733 279 T3

influenza (H1 Islas Salomón 2006) producido por ingeniería

<400> 53

5 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
1 5 10 15

Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
20 25 30

10 Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
35 40 45

Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
50 55 60

15 Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
65 70 75 80

20 Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
85 90 95

Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
100 105 110

25 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
115 120 125

30 Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val Glu Lys Pro Asn Pro  
130 135 140

Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
145 150 155 160

35 Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
165 170 175

Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
180 185 190

40 Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
195 200 205

45 Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
210 215 220

Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
225 230 235 240

50 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
245 250 255

55 Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
260 265 270

Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
275 280 285

60 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly

ES 2 733 279 T3

	290		295		300											
	Asn 305	Leu	Ile	Ala	Pro	Arg 310	Tyr	Ala	Phe	Ala	Leu 315	Ser	Arg	Gly	Phe	Gly 320
5	Ser	Gly	Ile	Ile	Asn 325	Ser	Asn	Ala	Pro	Met 330	Asp	Glu	Cys	Asp	Ala	Lys 335
	Cys	Gln	Thr	Pro 340	Gln	Gly	Ala	Ile	Asn 345	Ser	Ser	Leu	Pro	Phe	Gln	Asn 350
10	Val	His	Pro 355	Val	Thr	Ile	Gly	Glu 360	Cys	Pro	Lys	Tyr	Val 365	Arg	Ser	Ala
15	Lys	Leu	Arg	Met	Val	Thr	Gly 375	Leu	Arg	Asn	Ile	Pro 380	Ser	Ile	Glu	Gly
	Arg 385	Gly	Leu	Phe	Gly	Ala 390	Ile	Ala	Gly	Phe	Ile 395	Glu	Gly	Gly	Trp	Thr 400
20	Gly	Met	Val	Asp	Gly 405	Trp	Tyr	Gly	Tyr	His 410	His	Gln	Asn	Glu	Gln	Gly 415
	Ser	Gly	Tyr	Ala 420	Ala	Asp	Gln	Lys	Ser 425	Thr	Gln	Asn	Ala	Ile	Asn	Gly 430
25	Ile	Thr	Asn 435	Lys	Val	Asn	Ser	Val 440	Ile	Glu	Lys	Met	Asn 445	Thr	Gln	Phe
30	Thr	Ala 450	Val	Gly	Lys	Glu	Phe 455	Asn	Lys	Leu	Glu	Arg 460	Arg	Met	Glu	Asn
	Leu 465	Asn	Lys	Lys	Val	Asp 470	Asp	Gly	Phe	Ile	Asp 475	Ile	Trp	Thr	Tyr	Asn 480
35	Ala	Glu	Leu	Leu 485	Val	Leu	Leu	Glu	Asn 490	Glu	Arg	Thr	Leu	Asp	Phe	His 495
	Asp	Ser	Asn 500	Val	Lys	Asn	Leu	Tyr	Glu 505	Lys	Val	Lys	Ser	Gln	Leu	Lys 510
40	Asn	Asn	Ala 515	Lys	Glu	Ile	Gly	Asn 520	Gly	Cys	Phe	Glu	Phe 525	Tyr	His	Lys
45	Cys 530	Asn	Asp	Glu	Cys	Met	Glu 535	Ser	Val	Lys	Asn	Gly 540	Thr	Tyr	Asp	Tyr
	Pro 545	Lys	Tyr	Ser	Glu	Glu 550	Ser	Lys	Leu	Asn	Arg 555	Glu	Lys	Ile	Asp	Gly 560
50	Val	Arg	Ser	Asn 565	Ile	Asn	Lys	Leu	Lys	Ser 570	Ser	Ile	Glu	Ser	Thr	Asn 575
	Glu	Ala	Val	Val 580	Lys	Leu	Gln	Glu	Thr 585	Ala	Glu	Lys	Thr	Val	Tyr	Val 590
55	Leu	Thr	Ala 595	Leu	Gln											
60	<210>	54														
	<211>	596														

ES 2 733 279 T3

<212> PRT  
 <213> Secuencia Artificial

<220>

5 <223> Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de influenza (H1 Islas Salomón 2006) producido por ingeniería

<400> 54

10 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
 1 5 10 15

15 Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
 20 25 30

20 Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
 35 40 45

25 Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
 50 55 60

30 Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
 65 70 75 80

35 Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
 85 90 95

40 Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
 100 105 110

45 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
 115 120 125

50 Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val Glu Lys Pro Asn Pro  
 130 135 140

55 Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
 145 150 155 160

60 Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
 165 170 175

65 Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
 180 185 190

70 Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
 195 200 205

75 Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
 210 215 220

80 Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
 225 230 235 240

85 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
 245 250 255

90 Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
 260 265 270

ES 2 733 279 T3

Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
 275 280 285

5 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
 290 295 300

Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
 305 310 315 320

10 Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
 325 330 335

15 Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350

Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365

20 Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Asp Asp Asp  
 370 375 380

Asp Lys Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp  
 385 390 395 400

25 Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln  
 405 410 415

30 Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn  
 420 425 430

Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln  
 435 440 445

35 Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu  
 450 455 460

Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr  
 465 470 475 480

40 Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe  
 485 490 495

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

Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His  
 515 520 525

50 Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp  
 530 535 540

Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Ile Gly Glu Ala Arg Gln  
 545 550 555 560

55 Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile  
 565 570 575

Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln  
 580 585 590

60

ES 2 733 279 T3

Leu Gln Ala Arg  
595  
<210> 55  
<211> 594  
5 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Proteína fusión ectodominio-HR1Nipah de hemaglutinina del virus de  
10 influenza (H1 Islas Salomón 2006) producido por ingeniería

<400> 55  
Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
1 5 10 15  
15 Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
20 20 25 30  
Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
20 35 40 45  
Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
50 55 60  
25 Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
65 70 75 80  
Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
30 85 90 95  
Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
100 105 110  
35 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
115 120 125  
Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val Glu Lys Pro Asn Pro  
130 135 140  
40 Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
145 150 155 160  
Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
45 165 170 175  
Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
180 185 190  
50 Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
195 200 205  
Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
55 210 215 220  
Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
225 230 235 240  
60 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
245 250 255

ES 2 733 279 T3

Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
260 265 270

5 Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
275 280 285

Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
290 295 300

10 Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
305 310 315 320

15 Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
325 330 335

Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
340 345 350

20 Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
355 360 365

Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Asp Asp Asp  
370 375 380

25 Asp Lys Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp  
385 390 395 400

30 Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln  
405 410 415

Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn  
420 425 430

35 Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln  
435 440 445

Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu  
450 455 460

40 Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr  
465 470 475 480

45 Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe  
485 490 495

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

50 Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His  
515 520 525

55 Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp  
530 535 540

Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Ile Gly Glu Ala Arg Gln  
545 550 555 560

60 Asn Ile Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn Glu Ala Val  
565 570 575

ES 2 733 279 T3

Val Lys Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val Leu Thr Ala  
580 585 590

5 Leu Gln

<210> 56  
<211> 603  
10 <212> PRT  
<213> Secuencia Artificial

<220>  
15 <223> Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de  
influenza (H1 Islas Salomón 2006) producido por ingeniería

<400> 56  
Met Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
1 5 10 15

20 Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
20 25 30

25 Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
35 40 45

Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
50 55 60

30 Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
65 70 75 80

Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
85 90 95

35 Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
100 105 110

40 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
115 120 125

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

45 Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
145 150 155 160

Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
165 170 175

50 Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
180 185 190

55 Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
195 200 205

Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
210 215 220

60 Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro



ES 2 733 279 T3

	225				230					235					240	
	Pro	Asn	Ile	Gly	Asp	Gln	Arg	Ala	Leu	Tyr	His	Thr	Glu	Asn	Ala	Tyr
					245					250					255	
5	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg	Lys	Phe	Thr	Pro	Glu	Ile
				260					265					270		
	Ala	Lys	Arg	Pro	Lys	Val	Arg	Asp	Arg	Glu	Gly	Arg	Ile	Asn	Tyr	Tyr
10			275					280					285			
	Trp	Thr	Leu	Leu	Glu	Pro	Gly	Asp	Thr	Ile	Ile	Phe	Glu	Ala	Asn	Gly
		290					295					300				
15	Asn	Leu	Ile	Ala	Pro	Arg	Tyr	Ala	Phe	Ala	Leu	Ser	Arg	Gly	Phe	Gly
	305					310					315					320
	Ser	Gly	Ile	Ile	Asn	Ser	Asn	Ala	Pro	Met	Asp	Glu	Cys	Asp	Ala	Lys
20					325					330					335	
	Cys	Gln	Thr	Pro	Gln	Gly	Ala	Ile	Asn	Ser	Ser	Leu	Pro	Phe	Gln	Asn
				340					345					350		
25	Val	His	Pro	Val	Thr	Ile	Gly	Glu	Cys	Pro	Lys	Tyr	Val	Arg	Ser	Ala
			355					360					365			
	Lys	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn	Ile	Pro	Ser	Asp	Asp	Asp
		370					375					380				
30	Asp	Lys	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Gly	Gly	Trp
	385					390					395					400
	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Gln	Asn	Glu	Gln
				405						410					415	
35	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Gln	Lys	Ser	Thr	Gln	Asn	Ala	Ile	Asn
				420					425					430		
40	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	Met	Asn	Thr	Gln
			435					440					445			
	Phe	Thr	Ala	Val	Gly	Lys	Glu	Phe	Asn	Lys	Leu	Glu	Arg	Arg	Met	Glu
45		450					455					460				
	Asn	Leu	Asn	Lys	Lys	Val	Asp	Asp	Gly	Phe	Ile	Asp	Ile	Trp	Thr	Tyr
	465					470					475					480
50	Asn	Ala	Glu	Leu	Leu	Val	Leu	Leu	Glu	Asn	Glu	Arg	Thr	Leu	Asp	Phe
				485						490				495		
	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	Glu	Lys	Val	Lys	Ser	Gln	Leu
				500					505					510		
55	Lys	Asn	Asn	Ala	Lys	Glu	Ile	Gly	Asn	Gly	Cys	Phe	Glu	Phe	Tyr	His
		515						520					525			
60	Lys	Cys	Asn	Asp	Glu	Cys	Met	Glu	Ser	Val	Lys	Asn	Gly	Thr	Tyr	Asp
		530					535					540				

ES 2 733 279 T3

Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp  
 545 550 555 560  
 5 Gly Val Arg Ser Ile Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln  
 565 570 575  
 Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu  
 580 585 590  
 10 Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg  
 595 600  
 <210> 57  
 <211> 601  
 15 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Proteína fusión ectodominio-HR1Nipah de hemaglutinina del virus de  
 20 influenza (H1 Islas Salomón 2006) producido por ingeniería  
 <400> 57  
 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
 1 5 10 15  
 25 Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
 20 25 30  
 Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
 30 35 40 45  
 Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
 50 55 60  
 35 Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
 65 70 75 80  
 Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
 85 90 95  
 40 Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
 100 105 110  
 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
 115 120 125  
 45 Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val Glu Lys Pro Asn Pro  
 130 135 140  
 50 Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
 145 150 155 160  
 Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
 165 170 175  
 55 Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
 180 185 190  
 Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
 195 200 205  
 60

ES 2 733 279 T3

Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
 210 215 220  
**5** Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
 225 230 235 240  
 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
 245 250 255  
**10** Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
 260 265 270  
**15** Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
 275 280 285  
 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
 290 295 300  
**20** Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
 305 310 315 320  
 Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
 325 330 335  
**25** Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350  
**30** Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365  
 Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Asp Asp Asp  
 370 375 380  
**35** Asp Lys Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp  
 385 390 395 400  
**40** Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln  
 405 410 415  
 Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn  
 420 425 430  
**45** Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln  
 435 440 445  
 Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu  
 450 455 460  
**50** Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr  
 465 470 475 480  
**55** Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe  
 485 490 495  
 His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu  
 500 505 510  
**60** Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His

ES 2 733 279 T3

		515					520						525						
		Lys	Cys	Asn	Asp	Glu	Cys	Met	Glu	Ser	Val	Lys	Asn	Gly	Thr	Tyr	Asp		
			530					535					540						
<b>5</b>		Tyr	Pro	Lys	Tyr	Ser	Glu	Glu	Ser	Lys	Leu	Asn	Arg	Glu	Lys	Ile	Asp		
		545					550					555					560		
		Gly	Val	Arg	Ser	Ile	Arg	Gln	Asn	Ile	Asn	Lys	Leu	Lys	Ser	Ser	Ile		
<b>10</b>						565					570					575			
		Glu	Ser	Thr	Asn	Glu	Ala	Val	Val	Lys	Leu	Gln	Glu	Thr	Ala	Glu	Lys		
					580					585					590				
<b>15</b>		Thr	Val	Tyr	Val	Leu	Thr	Ala	Leu	Gln									
				595					600										
		<210>		58															
		<211>		610															
<b>20</b>		<212>		PRT															
		<213>		Secuencia Artificial															
		<220>																	
<b>25</b>		<223>		Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de															
				influenza (H1 Islas Salomón 2006) producido por ingeniería															
		<400>		58															
		Met	Leu	Leu	Val	Asn	Gln	Ser	His	Gln	Gly	Phe	Asn	Lys	Glu	His	Thr		
<b>30</b>		1				5					10					15			
		Ser	Lys	Met	Val	Ser	Ala	Ile	Val	Leu	Tyr	Val	Leu	Leu	Ala	Ala	Ala		
					20					25					30				
		Ala	His	Ser	Ala	Phe	Ala	Ala	Asp	Pro	His	His	His	His	His	His	Glu		
<b>35</b>					35				40						45				
		Asn	Leu	Tyr	Phe	Gln	Gly	Gly	Tyr	Leu	Leu	Glu	Asp	Thr	Ile	Cys	Ile		
			50				55						60						
<b>40</b>		Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr	Val	Asp	Thr	Val	Leu	Glu		
		65					70					75					80		
		Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn	Leu	Leu	Glu	Asp	Ser	His		
					85						90					95			
<b>45</b>		Asn	Gly	Lys	Leu	Cys	Leu	Leu	Lys	Gly	Ile	Ala	Pro	Leu	Gln	Leu	Gly		
					100					105					110				
		Asn	Cys	Ser	Val	Ala	Gly	Trp	Ile	Leu	Gly	Asn	Pro	Glu	Cys	Glu	Leu		
<b>50</b>					115				120					125					
		Leu	Ile	Ser	Arg	Glu	Ser	Trp	Ser	Tyr	Ile	Val	Glu	Lys	Pro	Asn	Pro		
					130			135					140						
<b>55</b>		Glu	Asn	Gly	Thr	Cys	Tyr	Pro	Gly	His	Phe	Ala	Asp	Tyr	Glu	Glu	Leu		
		145					150					155					160		
		Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe	Glu	Arg	Phe	Glu	Ile	Phe		
					165						170					175			
<b>60</b>																			

ES 2 733 279 T3

Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
 180 185 190  
 5 Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
 195 200 205  
 Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
 210 215 220  
 10 Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
 225 230 235 240  
 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
 245 250 255  
 15 Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
 260 265 270  
 20 Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
 275 280 285  
 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
 290 295 300  
 25 Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
 305 310 315 320  
 30 Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
 325 330 335  
 Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350  
 35 Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365  
 Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Asp Asp Asp  
 370 375 380  
 40 Asp Lys Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp  
 385 390 395 400  
 45 Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln  
 405 410 415  
 Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn  
 420 425 430  
 50 Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln  
 435 440 445  
 Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu  
 450 455 460  
 55 Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr  
 465 470 475 480  
 60 Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe  
 485 490 495

ES 2 733 279 T3

His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu  
 500 505 510  
**5** Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His  
 515 520 525  
 Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp  
 530 535 540  
**10** Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp  
 545 550 555 560  
**15** Gly Val Arg Ser Ile Lys Lys Leu Ile Gly Glu Ala Arg Gln Leu Leu  
 565 570 575  
 Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala  
 580 585 590  
**20** Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln  
 595 600 605  
**25** Ala Arg  
 610  
**30** <210> 59  
 <211> 608  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Proteína fusión ectodominio-HR1Nipah de hemaglutinina del virus de  
 influenza (H1 Islas Salomón 2006) producido por ingeniería  
**35** <400> 59  
 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
 1 5 10 15  
**40** Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
 20 25 30  
 Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
 35 40 45  
**45** Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Thr Ile Cys Ile  
 50 55 60  
**50** Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu  
 65 70 75 80  
 Lys Asn Val Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His  
 85 90 95  
**55** Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly  
 100 105 110  
 Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu  
 115 120 125  
**60**

ES 2 733 279 T3

Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val Glu Lys Pro Asn Pro  
 130 135 140  
 5 Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala Asp Tyr Glu Glu Leu  
 145 150 155 160  
 Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe  
 165 170 175  
 10 Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr Thr Gly Val Ser Ala  
 180 185 190  
 Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr Lys Asn Leu Leu Trp  
 195 200 205  
 15 Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Ala  
 210 215 220  
 Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro  
 225 230 235 240  
 20 Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr  
 245 250 255  
 25 Val Ser Val Val Ser Ser His Tyr Ser Arg Lys Phe Thr Pro Glu Ile  
 260 265 270  
 Ala Lys Arg Pro Lys Val Arg Asp Arg Glu Gly Arg Ile Asn Tyr Tyr  
 275 280 285  
 30 Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly  
 290 295 300  
 Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly  
 305 310 315 320  
 35 Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys  
 325 330 335  
 40 Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn  
 340 345 350  
 Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala  
 355 360 365  
 45 Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser Asp Asp Asp  
 370 375 380  
 Asp Lys Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp  
 385 390 395 400  
 Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln  
 405 410 415  
 55 Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn  
 420 425 430  
 Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln  
 435 440 445  
 60

ES 2 733 279 T3

Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu  
 450 455 460  
 5 Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp Ile Trp Thr Tyr  
 465 470 475 480  
 10 Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe  
 485 490 495  
 His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu  
 500 505 510  
 15 Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His  
 515 520 525  
 Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp  
 530 535 540  
 20 Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp  
 545 550 555 560  
 Gly Val Arg Ser Ile Lys Lys Leu Ile Gly Glu Ala Arg Gln Asn Ile  
 565 570 575  
 25 Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn Glu Ala Val Val Lys  
 580 585 590  
 30 Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val Leu Thr Ala Leu Gln  
 595 600 605  
 <210> 60  
 <211> 606  
 <212> PRT  
 35 <213> Secuencia Artificial  
 <220>  
 <223> Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de  
 influenza (H5 Vietnam 2006) producido por ingeniería  
 40 <400> 60  
 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
 1 5 10 15  
 45 Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
 20 25 30  
 Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
 35 40 45  
 50 Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Gln Ile Cys Ile  
 50 55 60  
 55 Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val Asp Thr Ile Met Glu  
 65 70 75 80  
 Lys Asn Val Thr Val Thr His Ala Gln Asp Ile Leu Glu Lys Lys His  
 85 90 95  
 60 Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys Pro Leu Ile Leu Arg



ES 2 733 279 T3

				100					105					110			
	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn	Pro	Met	Cys	Asp	Glu	
			115					120					125				
5	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val	Glu	Lys	Ala	Asn	Pro	
	130						135					140					
	Val	Asn	Asp	Leu	Cys	Tyr	Pro	Gly	Asp	Phe	Asn	Asp	Tyr	Glu	Glu	Leu	
10	145					150					155					160	
	Lys	His	Leu	Leu	Ser	Arg	Ile	Asn	His	Phe	Glu	Lys	Ile	Gln	Ile	Ile	
					165					170					175		
15	Pro	Lys	Ser	Ser	Trp	Ser	Ser	His	Glu	Ala	Ser	Leu	Gly	Val	Ser	Ser	
					180				185					190			
	Ala	Cys	Pro	Tyr	Gln	Gly	Lys	Ser	Ser	Phe	Phe	Arg	Asn	Val	Val	Trp	
20			195					200					205				
	Leu	Ile	Lys	Lys	Asn	Ser	Thr	Tyr	Pro	Thr	Ile	Lys	Arg	Ser	Tyr	Asn	
	210						215					220					
25	Asn	Thr	Asn	Gln	Glu	Asp	Leu	Leu	Val	Leu	Trp	Gly	Ile	His	His	Pro	
	225					230					235					240	
	Asn	Asp	Ala	Ala	Glu	Gln	Thr	Lys	Leu	Tyr	Gln	Asn	Pro	Thr	Thr	Tyr	
					245					250					255		
30	Ile	Ser	Val	Gly	Thr	Ser	Thr	Leu	Asn	Gln	Arg	Leu	Val	Pro	Arg	Ile	
				260					265					270			
	Ala	Thr	Arg	Ser	Lys	Val	Asn	Gly	Gln	Ser	Gly	Arg	Met	Glu	Phe	Phe	
35			275					280					285				
	Trp	Thr	Ile	Leu	Lys	Pro	Asn	Asp	Ala	Ile	Asn	Phe	Glu	Ser	Asn	Gly	
	290						295					300					
40	Asn	Phe	Ile	Ala	Pro	Glu	Tyr	Ala	Tyr	Lys	Ile	Val	Lys	Lys	Gly	Asp	
	305					310					315					320	
	Ser	Thr	Ile	Met	Lys	Ser	Glu	Leu	Glu	Tyr	Gly	Asn	Cys	Asn	Thr	Lys	
				325						330					335		
45	Cys	Gln	Thr	Pro	Met	Gly	Ala	Ile	Asn	Ser	Ser	Met	Pro	Phe	His	Asn	
				340					345					350			
	Ile	His	Pro	Leu	Thr	Ile	Gly	Glu	Cys	Pro	Lys	Tyr	Val	Lys	Ser	Asn	
50			355					360					365				
	Arg	Leu	Val	Leu	Ala	Thr	Gly	Leu	Arg	Asn	Ser	Pro	Gln	Arg	Glu	Arg	
	370						375					380					
55	Asp	Asp	Asp	Asp	Lys	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	
	385					390					395					400	
	Gly	Gly	Trp	Gln	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Ser	
60				405						410					415		

ES 2 733 279 T3

Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln Lys  
420 425 430

5 Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp Lys Met  
435 440 445

Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu Arg  
450 455 460

10 Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp Val  
465 470 475 480

Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg Thr  
485 490 495

15 Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val Arg  
500 505 510

20 Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe Glu  
515 520 525

Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg Asn Gly  
530 535 540

25 Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg Glu  
545 550 555 560

Glu Ile Ser Gly Val Arg Ser Leu Arg Gln Leu Leu Ser Gly Ile Val  
565 570 575

30 Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu  
580 585 590

35 Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg  
595 600 605

<210> 61  
<211> 632  
<212> PRT  
40 <213> Secuencia Artificial

<220>  
<223> Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de  
influenza (H5 Vietnam 2006) producido por ingeniería

45 <400> 61  
Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr  
1 5 10 15

50 Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
20 25 30

Ala His Ser Ala Phe Ala Ala Asp Pro His His His His His His Glu  
35 40 45

55 Asn Leu Tyr Phe Gln Gly Gly Tyr Leu Leu Glu Asp Gln Ile Cys Ile  
50 55 60

60 Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val Asp Thr Ile Met Glu  
65 70 75 80

ES 2 733 279 T3

Lys Asn Val Thr Val Thr His Ala Gln Asp Ile Leu Glu Lys Lys His  
 85 90 95  
**5** Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys Pro Leu Ile Leu Arg  
 100 105 110  
 Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn Pro Met Cys Asp Glu  
 115 120 125  
**10** Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val Glu Lys Ala Asn Pro  
 130 135 140  
**15** Val Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn Asp Tyr Glu Glu Leu  
 145 150 155 160  
 Lys His Leu Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln Ile Ile  
 165 170 175  
**20** Pro Lys Ser Ser Trp Ser Ser His Glu Ala Ser Leu Gly Val Ser Ser  
 180 185 190  
 Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe Arg Asn Val Val Trp  
 195 200 205  
**25** Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys Arg Ser Tyr Asn  
 210 215 220  
**30** Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly Ile His His Pro  
 225 230 235 240  
 Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln Asn Pro Thr Thr Tyr  
 245 250 255  
**35** Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg Leu Val Pro Arg Ile  
 260 265 270  
 Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg Met Glu Phe Phe  
 275 280 285  
**40** Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe Glu Ser Asn Gly  
 290 295 300  
**45** Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Val Lys Lys Gly Asp  
 305 310 315 320  
 Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly Asn Cys Asn Thr Lys  
 325 330 335  
**50** Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser Met Pro Phe His Asn  
 340 345 350  
**55** Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys Tyr Val Lys Ser Asn  
 355 360 365  
 Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln Arg Glu Arg  
 370 375 380  
**60** Asp Asp Asp Asp Lys Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu



ES 2 733 279 T3

Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala  
 20 25 30  
**5** Ala His Ser Ala Phe Ala Gln Lys Leu Pro His His His His His His  
 35 40 45  
 Glu Asn Leu Tyr Phe Gln Gly Gly Asn Asp Asn Ser Thr Ala Thr Leu  
 50 55 60  
**10** Cys Leu Gly His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile  
 65 70 75 80  
**15** Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser  
 85 90 95  
 Ser Ser Thr Gly Gly Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly  
 100 105 110  
**20** Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp  
 115 120 125  
 Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala  
 130 135 140  
**25** Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg  
 145 150 155 160  
**30** Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe Asn Asp Glu Ser Phe  
 165 170 175  
 Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser Ser Ser Cys Lys Arg  
 180 185 190  
**35** Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr Gln Leu  
 195 200 205  
 Lys Phe Lys Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Lys  
 210 215 220  
**40** Phe Asp Lys Leu Tyr Ile Trp Gly Val His His Pro Val Thr Asp Asn  
 225 230 235 240  
**45** Asp Gln Ile Phe Leu Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser  
 245 250 255  
 Thr Lys Arg Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser Arg Pro  
 260 265 270  
**50** Arg Ile Arg Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val  
 275 280 285  
 Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala  
 290 295 300  
**55** Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg  
 305 310 315 320  
**60** Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn  
 325 330 335

ES 2 733 279 T3

Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr  
 340 345 350  
**5** Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala  
 355 360 365  
 Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr Arg Asp Asp Asp Asp  
 370 375 380  
**10** Lys Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu  
 385 390 395 400  
**15** Gly Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly  
 405 410 415  
 Ile Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asn Gln  
 420 425 430  
**20** Ile Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys Phe  
 435 440 445  
 His Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln Asp  
 450 455 460  
**25** Leu Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr Asn  
 465 470 475 480  
**30** Ala Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr  
 485 490 495  
**35** Asp Ser Glu Met Asn Lys Leu Phe Glu Arg Thr Lys Lys Gln Leu Arg  
 500 505 510  
 Glu Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys  
 515 520 525  
**40** Cys Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His  
 530 535 540  
 Asp Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly  
 545 550 555 560  
**45** Val Arg Ser Leu Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn  
 565 570 575  
**50** Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr  
 580 585 590  
 Val Trp Gly Ile Lys Gln Leu Gln Ala Arg  
 595 600  
**55** <210> 63  
 <211> 628  
 <212> PRT  
 <213> Secuencia Artificial  
**60** <220>  
 <223> Proteína fusión ectodominio-HR1HXB2 de hemaglutinina del virus de

# ES 2 733 279 T3

influenza (H3 Wisconsin 2005) producido por ingeniería

```

<400> 63
5 Met Leu Leu Val Asn Gln Ser His Gln Gly Phe Asn Lys Glu His Thr
  1          5          10          15

  Ser Lys Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala
    20          25          30

10 Ala His Ser Ala Phe Ala Gln Lys Leu Pro His His His His His His
    35          40          45

  Glu Asn Leu Tyr Phe Gln Gly Gly Asn Asp Asn Ser Thr Ala Thr Leu
    50          55          60

15 Cys Leu Gly His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile
    65          70          75          80

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

  Ser Ser Thr Gly Gly Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly
   100          105          110

25 Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp
   115          120          125

  Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala
   130          135          140

30 Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg
   145          150          155          160

  Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe Asn Asp Glu Ser Phe
   165          170          175

  Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser Ser Ser Cys Lys Arg
   180          185          190

40 Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr Gln Leu
   195          200          205

  Lys Phe Lys Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Lys
   210          215          220

45 Phe Asp Lys Leu Tyr Ile Trp Gly Val His His Pro Val Thr Asp Asn
   225          230          235          240

  Asp Gln Ile Phe Leu Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser
   245          250          255

  Thr Lys Arg Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser Arg Pro
   260          265          270

55 Arg Ile Arg Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val
   275          280          285

  Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala
   290          295          300

60

```

ES 2 733 279 T3

Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg  
 305 310 315 320  
 5 Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn  
 325 330 335  
 Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr  
 340 345 350  
 10 Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala  
 355 360 365  
 Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr Arg Asp Asp Asp Asp  
 370 375 380  
 15 Lys Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu  
 385 390 395 400  
 20 Gly Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly  
 405 410 415  
 Ile Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asn Gln  
 420 425 430  
 25 Ile Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys Phe  
 435 440 445  
 30 His Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln Asp  
 450 455 460  
 Leu Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr Asn  
 465 470 475 480  
 35 Ala Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr  
 485 490 495  
 Asp Ser Glu Met Asn Lys Leu Phe Glu Arg Thr Lys Lys Gln Leu Arg  
 500 505 510  
 40 Glu Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys  
 515 520 525  
 45 Cys Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His  
 530 535 540  
 Asp Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly  
 545 550 555 560  
 50 Val Arg Ser Leu Gln Glu Ala Gly Tyr Ile Pro Glu Ala Pro Arg Asp  
 565 570 575  
 Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr  
 580 585 590  
 55 Phe Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile  
 595 600 605  
 60 Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln  
 610 615 620



ES 2 733 279 T3

Leu Gln Ala Arg  
625

5 <210> 64  
<211> 755  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> Proteína fusión ectodominio (gp140)-HR1HXB2 de glicoproteína que envuelve el VIH-1 (1992 Uganda, serotipo 037.8, clado A) producido por ingeniería

15 <400> 64  
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
1 5 10 15  
Ser Tyr Ile Tyr Ala Glu Asp His His His His His Glu Asn Leu  
20 25 30

20 Tyr Phe Gln Gly Ala Glu Asn Leu Trp Val Thr Val Tyr Tyr Gly Val  
35 40 45

25 Pro Val Trp Lys Asp Ala Glu Thr Thr Leu Phe Cys Ala Ser Asp Ala  
50 55 60

Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys  
65 70 75 80

30 Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Tyr Met Glu Asn Val Thr  
85 90 95

Glu Glu Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His Thr  
100 105 110

35 Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Gln Leu  
115 120 125

40 Thr Pro Leu Cys Val Thr Leu Asp Cys Ser Tyr Asn Ile Thr Asn Asn  
130 135 140

Ile Thr Asn Ser Ile Thr Asn Ser Ser Val Asn Met Arg Glu Glu Ile  
145 150 155 160

45 Lys Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp Lys Asn Arg  
165 170 175

Lys Val Tyr Ser Leu Phe Tyr Lys Leu Asp Val Val Gln Ile Asn Asn  
180 185 190

50 Gly Asn Asn Ser Ser Asn Leu Tyr Arg Leu Ile Asn Cys Asn Thr Ser  
195 200 205

55 Ala Leu Thr Gln Ala Cys Pro Lys Val Thr Phe Glu Pro Ile Pro Ile  
210 215 220

Arg Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asp Lys  
225 230 235 240

60 Glu Phe Asn Gly Thr Gly Leu Cys Lys Asn Val Ser Thr Val Gln Cys

ES 2 733 279 T3

				245					250					255			
	Thr	His	Gly	Ile	Arg	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	
				260					265					270			
5	Ser	Leu	Ala	Glu	Gly	Lys	Val	Met	Ile	Arg	Ser	Glu	Asn	Ile	Thr	Asn	
			275					280					285				
10	Asn	Val	Lys	Asn	Ile	Ile	Val	Gln	Leu	Asn	Glu	Thr	Val	Thr	Ile	Asn	
		290					295					300					
15	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Val	Arg	Ile	Gly	Pro	
	305					310					315					320	
	Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln	
				325						330					335		
20	Ala	His	Cys	Asn	Val	Ser	Gly	Ser	Gln	Trp	Asn	Arg	Ala	Leu	His	Gln	
				340					345					350			
	Val	Val	Gly	Gln	Leu	Arg	Glu	Tyr	Trp	Asn	Thr	Thr	Ile	Ile	Phe	Lys	
			355					360					365				
25	Asn	Ser	Ser	Gly	Gly	Asp	Leu	Glu	Ile	Thr	Thr	His	Ser	Phe	Asn	Cys	
		370					375					380					
30	Gly	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	Gly	Leu	Phe	Asn	Ser	Asn	
	385					390					395					400	
	Trp	Thr	His	Asn	Asp	Thr	Ala	Ser	Met	Lys	Pro	Asn	Asp	Thr	Ile	Thr	
				405						410					415		
35	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	Arg	Val	Gly	
				420					425					430			
	Gln	Ala	Ile	Tyr	Ala	Pro	Pro	Ile	Gln	Gly	Val	Ile	Arg	Cys	Glu	Ser	
			435					440					445				
40	Asn	Ile	Thr	Gly	Leu	Ile	Leu	Thr	Arg	Asp	Gly	Gly	Gly	Asn	Ile	Asn	
		450					455					460					
45	Glu	Ser	Gln	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	
	465					470					475				480		
	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Arg	Ile	Glu	Pro	Leu	Gly	
				485						490					495		
50	Val	Ala	Pro	Thr	Lys	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg	Glu	Lys	Arg	
				500					505					510			
	Ala	Val	Val	Glu	Leu	Gly	Ala	Val	Phe	Ile	Gly	Phe	Leu	Gly	Thr	Ala	
			515					520					525				
55	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Ile	Thr	Leu	Thr	Val	Gln	Val	Arg	
		530					535					540					
60	Lys	Leu	Leu	Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu	Arg	Ala	
	545					550					555					560	

ES 2 733 279 T3

Ile Glu Ala Gln Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys  
565 570 575

5 Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Arg Asp Gln  
580 585 590

10 Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr  
595 600 605

Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Glu Arg Glu Ile  
610 615 620

15 Trp Glu Asn Met Thr Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr  
625 630 635 640

Thr His Ile Ile Tyr Glu Leu Ile Glu Glu Ser Gln Lys Gln Gln Glu  
645 650 655

20 Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Asn Leu Trp  
660 665 670

Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ser Arg Ile  
675 680 685

25 Glu Gly Arg Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly  
690 695 700

30 Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe  
705 710 715 720

Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu  
725 730 735

35 Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu  
740 745 750

40 Gln Ala Arg  
755

<210> 65  
<211> 763  
<212> PRT  
<213> Secuencia Artificial

45 <220>  
<223> Proteína fusión ectodominio (gp140)-HR1HXB2 de glicoproteína que  
envuelve el VIH-1 (1992 Uganda, serotipo 037.8, clado A) producido por  
ingeniería

50 <400> 65  
Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
1 5 10 15

55 Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Glu Asp His  
20 25 30

His His His His His Glu Asn Leu Tyr Phe Gln Gly Ala Glu Asn Leu  
35 40 45

60

ES 2 733 279 T3

Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Glu Thr  
 50 55 60  
 Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His  
 65 70 75 80  
**5**  
 Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln  
 85 90 95  
 Glu Ile Tyr Met Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys Asn  
 100 105 110  
**10**  
 Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp Gln  
 115 120 125  
**15**  
 Ser Leu Lys Pro Cys Val Gln Leu Thr Pro Leu Cys Val Thr Leu Asp  
 130 135 140  
 Cys Ser Tyr Asn Ile Thr Asn Asn Ile Thr Asn Ser Ile Thr Asn Ser  
 145 150 155 160  
**20**  
 Ser Val Asn Met Arg Glu Glu Ile Lys Asn Cys Ser Phe Asn Met Thr  
 165 170 175  
**25**  
 Thr Glu Leu Arg Asp Lys Asn Arg Lys Val Tyr Ser Leu Phe Tyr Lys  
 180 185 190  
 Leu Asp Val Val Gln Ile Asn Asn Gly Asn Asn Ser Ser Asn Leu Tyr  
 195 200 205  
**30**  
 Arg Leu Ile Asn Cys Asn Thr Ser Ala Leu Thr Gln Ala Cys Pro Lys  
 210 215 220  
 Val Thr Phe Glu Pro Ile Pro Ile Arg Tyr Cys Ala Pro Ala Gly Tyr  
 225 230 235 240  
**35**  
 Ala Ile Leu Lys Cys Asn Asp Lys Glu Phe Asn Gly Thr Gly Leu Cys  
 245 250 255  
**40**  
 Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val  
 260 265 270  
 Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Gly Lys Val Met  
 275 280 285  
**45**  
 Ile Arg Ser Glu Asn Ile Thr Asn Asn Val Lys Asn Ile Ile Val Gln  
 290 295 300  
 Leu Asn Glu Thr Val Thr Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr  
 305 310 315 320  
**50**  
 Arg Lys Ser Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly  
 325 330 335  
**55**  
 Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Val Ser Gly Ser  
 340 345 350  
 Gln Trp Asn Arg Ala Leu His Gln Val Val Gly Gln Leu Arg Glu Tyr  
 355 360 365  
**60**

ES 2 733 279 T3

Trp Asn Thr Thr Ile Ile Phe Lys Asn Ser Ser Gly Gly Asp Leu Glu  
 370 375 380  
 5 Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn  
 385 390 395 400  
 Thr Ser Gly Leu Phe Asn Ser Asn Trp Thr His Asn Asp Thr Ala Ser  
 405 410 415  
 10 Met Lys Pro Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile  
 420 425 430  
 15 Ile Asn Met Trp Gln Arg Val Gly Gln Ala Ile Tyr Ala Pro Pro Ile  
 435 440 445  
 Gln Gly Val Ile Arg Cys Glu Ser Asn Ile Thr Gly Leu Ile Leu Thr  
 450 455 460  
 20 Arg Asp Gly Gly Gly Asn Ile Asn Glu Ser Gln Ile Phe Arg Pro Gly  
 465 470 475 480  
 Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys  
 485 490 495  
 25 Val Val Arg Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg  
 500 505 510  
 30 Arg Val Val Glu Arg Glu Lys Arg Ala Val Val Glu Leu Gly Ala Val  
 515 520 525  
 Phe Ile Gly Phe Leu Gly Thr Ala Gly Ser Thr Met Gly Ala Ala Ser  
 530 535 540  
 35 Ile Thr Leu Thr Val Gln Val Arg Lys Leu Leu Ser Gly Ile Val Gln  
 545 550 555 560  
 Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu  
 565 570 575  
 40 Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala  
 580 585 590  
 Val Glu Arg Tyr Leu Arg Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys  
 595 600 605  
 45 Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp  
 610 615 620  
 50 Ser Asn Lys Ser Glu Arg Glu Ile Trp Glu Asn Met Thr Trp Leu Gln  
 625 630 635 640  
 55 Trp Asp Lys Glu Ile Ser Asn Tyr Thr His Ile Ile Tyr Glu Leu Ile  
 645 650 655  
 Glu Glu Ser Gln Lys Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu  
 660 665 670  
 60 Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp  
 675 680 685

ES 2 733 279 T3

Leu Trp Tyr Ile Lys Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr  
 690 695 700  
**5** Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly  
 705 710 715 720  
 Glu Trp Val Leu Leu Ser Thr Phe Leu Ser Gly Ile Val Gln Gln Gln  
 725 730 735  
**10** Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu  
 740 745 750  
**15** Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg  
 755 760  
 <210> 66  
 <211> 722  
 <212> PRT  
**20** <213> Secuencia Artificial  
 <220>  
 <223> Proteína fusión ectodominio (gp140)-HR1HXB2 de glicoproteína que  
 envuelve el VIH-1 (1992 Uganda, serotipo 037.8, clado A) producido por  
**25** ingeniería  
 <400> 66  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
**30** Ser Tyr Ile Tyr Ala Glu Asp His His His His His Glu Asn Leu  
 20 25 30  
**35** Tyr Phe Gln Gly Ala Glu Asn Leu Trp Val Thr Val Tyr Tyr Gly Val  
 35 40 45  
 Pro Val Trp Lys Asp Ala Glu Thr Thr Leu Phe Cys Ala Ser Asp Ala  
 50 55 60  
**40** Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys  
 65 70 75 80  
 Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Tyr Met Glu Asn Val Thr  
 85 90 95  
**45** Glu Glu Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His Thr  
 100 105 110  
 Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Gln Leu  
 115 120 125  
**50** Thr Pro Leu Cys Val Thr Leu Asp Cys Ser Tyr Asn Ile Thr Asn Asn  
 130 135 140  
**55** Ile Thr Asn Ser Ile Thr Asn Ser Ser Val Asn Met Arg Glu Glu Ile  
 145 150 155 160  
 Lys Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp Lys Asn Arg  
 165 170 175  
**60** Lys Val Tyr Ser Leu Phe Tyr Lys Leu Asp Val Val Gln Ile Asn Asn

ES 2 733 279 T3

				180					185					190			
	Gly	Asn	Asn	Ser	Ser	Asn	Leu	Tyr	Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser	
			195					200					205				
5	Ala	Leu	Thr	Gln	Ala	Cys	Pro	Lys	Val	Thr	Phe	Glu	Pro	Ile	Pro	Ile	
		210					215					220					
10	Arg	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr	Ala	Ile	Leu	Lys	Cys	Asn	Asp	Lys	
	225					230					235					240	
	Glu	Phe	Asn	Gly	Thr	Gly	Leu	Cys	Lys	Asn	Val	Ser	Thr	Val	Gln	Cys	
					245					250					255		
15	Thr	His	Gly	Ile	Arg	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	
				260					265						270		
	Ser	Leu	Ala	Glu	Gly	Lys	Val	Met	Ile	Arg	Ser	Glu	Asn	Ile	Thr	Asn	
			275					280					285				
20	Asn	Val	Lys	Asn	Ile	Ile	Val	Gln	Leu	Asn	Glu	Thr	Val	Thr	Ile	Asn	
		290					295					300					
25	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Val	Arg	Ile	Gly	Pro	
	305					310					315				320		
	Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln	
					325					330					335		
30	Ala	His	Cys	Asn	Val	Ser	Gly	Ser	Gln	Trp	Asn	Arg	Ala	Leu	His	Gln	
				340					345					350			
	Val	Val	Gly	Gln	Leu	Arg	Glu	Tyr	Trp	Asn	Thr	Thr	Ile	Ile	Phe	Lys	
			355					360					365				
35	Asn	Ser	Ser	Gly	Gly	Asp	Leu	Glu	Ile	Thr	Thr	His	Ser	Phe	Asn	Cys	
		370					375					380					
40	Gly	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	Gly	Leu	Phe	Asn	Ser	Asn	
	385					390					395					400	
	Trp	Thr	His	Asn	Asp	Thr	Ala	Ser	Met	Lys	Pro	Asn	Asp	Thr	Ile	Thr	
				405						410					415		
45	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	Arg	Val	Gly	
				420					425					430			
	Gln	Ala	Ile	Tyr	Ala	Pro	Pro	Ile	Gln	Gly	Val	Ile	Arg	Cys	Glu	Ser	
			435				440						445				
50	Asn	Ile	Thr	Gly	Leu	Ile	Leu	Thr	Arg	Asp	Gly	Gly	Gly	Asn	Ile	Asn	
		450					455					460					
55	Glu	Ser	Gln	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	
	465					470					475					480	
	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Arg	Ile	Glu	Pro	Leu	Gly	
					485					490					495		
60																	

ES 2 733 279 T3

Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg  
500 505 510

5 Ala Val Val Glu Leu Gly Ala Val Phe Ile Gly Phe Leu Gly Thr Ala  
515 520 525

Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Val Arg  
530 535 540

10 Lys Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala  
545 550 555 560

Ile Glu Ala Gln Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys  
15 565 570 575

Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Arg Asp Gln  
580 585 590

20 Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr  
595 600 605

Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Glu Arg Glu Ile  
610 615 620

25 Trp Glu Asn Met Thr Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr  
625 630 635 640

Thr His Ile Ile Tyr Glu Leu Ile Glu Glu Ser Gln Lys Gln Gln Glu  
30 645 650 655

Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Asn Leu Trp  
660 665 670

35 Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys Gly Ser Gly  
675 680 685

Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala  
40 690 695 700

Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln  
705 710 715 720

45 Ala Arg

<210> 67  
<211> 730  
<212> PRT  
50 <213> Secuencia Artificial

<220>  
<223> Proteína fusión ectodominio (gp140)-HR1HXB2 de glicoproteína que  
55 envuelve el VIH-1 (1992 Uganda, serotipo 037.8, clado A) producido por  
ingeniería

<400> 67

60 Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
1 5 10 15



ES 2 733 279 T3

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Glu Asp His  
 20 25 30  
**5** His His His His His Glu Asn Leu Tyr Phe Gln Gly Ala Glu Asn Leu  
 35 40 45  
 Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Glu Thr  
 50 55 60  
**10** Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His  
 65 70 75 80  
**15** Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln  
 85 90 95  
 Glu Ile Tyr Met Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys Asn  
 100 105 110  
**20** Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp Gln  
 115 120 125  
 Ser Leu Lys Pro Cys Val Gln Leu Thr Pro Leu Cys Val Thr Leu Asp  
 130 135 140  
**25** Cys Ser Tyr Asn Ile Thr Asn Asn Ile Thr Asn Ser Ile Thr Asn Ser  
 145 150 155 160  
**30** Ser Val Asn Met Arg Glu Glu Ile Lys Asn Cys Ser Phe Asn Met Thr  
 165 170 175  
 Thr Glu Leu Arg Asp Lys Asn Arg Lys Val Tyr Ser Leu Phe Tyr Lys  
 180 185 190  
**35** Leu Asp Val Val Gln Ile Asn Asn Gly Asn Asn Ser Ser Asn Leu Tyr  
 195 200 205  
**40** Arg Leu Ile Asn Cys Asn Thr Ser Ala Leu Thr Gln Ala Cys Pro Lys  
 210 215 220  
 Val Thr Phe Glu Pro Ile Pro Ile Arg Tyr Cys Ala Pro Ala Gly Tyr  
 225 230 235 240  
**45** Ala Ile Leu Lys Cys Asn Asp Lys Glu Phe Asn Gly Thr Gly Leu Cys  
 245 250 255  
 Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val  
 260 265 270  
**50** Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Gly Lys Val Met  
 275 280 285  
**55** Ile Arg Ser Glu Asn Ile Thr Asn Asn Val Lys Asn Ile Ile Val Gln  
 290 295 300  
 Leu Asn Glu Thr Val Thr Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr  
 305 310 315 320  
**60** Arg Lys Ser Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly

ES 2 733 279 T3

				325					330					335			
	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn	Val	Ser	Gly	Ser	
				340					345					350			
5	Gln	Trp	Asn	Arg	Ala	Leu	His	Gln	Val	Val	Gly	Gln	Leu	Arg	Glu	Tyr	
			355					360					365				
10	Trp	Asn	Thr	Thr	Ile	Ile	Phe	Lys	Asn	Ser	Ser	Gly	Gly	Asp	Leu	Glu	
		370					375					380					
15	Ile	Thr	Thr	His	Ser	Phe	Asn	Cys	Gly	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	
		385				390					395					400	
20	Thr	Ser	Gly	Leu	Phe	Asn	Ser	Asn	Trp	Thr	His	Asn	Asp	Thr	Ala	Ser	
					405					410					415		
25	Met	Lys	Pro	Asn	Asp	Thr	Ile	Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	
				420					425					430			
30	Ile	Asn	Met	Trp	Gln	Arg	Val	Gly	Gln	Ala	Ile	Tyr	Ala	Pro	Pro	Ile	
			435					440					445				
35	Gln	Gly	Val	Ile	Arg	Cys	Glu	Ser	Asn	Ile	Thr	Gly	Leu	Ile	Leu	Thr	
		450					455					460					
40	Arg	Asp	Gly	Gly	Gly	Asn	Ile	Asn	Glu	Ser	Gln	Ile	Phe	Arg	Pro	Gly	
		465				470					475					480	
45	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	
					485					490					495		
50	Val	Val	Arg	Ile	Glu	Pro	Leu	Gly	Val	Ala	Pro	Thr	Lys	Ala	Lys	Arg	
				500					505					510			
55	Arg	Val	Val	Glu	Arg	Glu	Lys	Arg	Ala	Val	Val	Glu	Leu	Gly	Ala	Val	
			515					520					525				
60	Phe	Ile	Gly	Phe	Leu	Gly	Thr	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	
		530					535					540					
65	Ile	Thr	Leu	Thr	Val	Gln	Val	Arg	Lys	Leu	Leu	Ser	Gly	Ile	Val	Gln	
		545				550						555				560	
70	Gln	Gln	Ser	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	
					565					570					575		
75	Lys	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Ala	Arg	Val	Leu	Ala	
				580					585					590			
80	Val	Glu	Arg	Tyr	Leu	Arg	Asp	Gln	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	
			595					600					605				
85	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Asn	Val	Pro	Trp	Asn	Ser	Ser	Trp	
		610					615					620					
90	Ser	Asn	Lys	Ser	Glu	Arg	Glu	Ile	Trp	Glu	Asn	Met	Thr	Trp	Leu	Gln	
		625				630						635				640	

ES 2 733 279 T3

Trp Asp Lys Glu Ile Ser Asn Tyr Thr His Ile Ile Tyr Glu Leu Ile  
645 650 655

5 Glu Glu Ser Gln Lys Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu  
660 665 670

Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp  
675 680 685

10 Leu Trp Tyr Ile Lys Gly Ser Gly Ser Gly Ile Val Gln Gln Gln Asn  
690 695 700

Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr  
705 710 715 720

15 Val Trp Gly Ile Lys Gln Leu Gln Ala Arg  
725 730

20 <210> 68  
<211> 753  
<212> PRT  
<213> Secuencia Artificial  
<220>  
<223> Proteína fusión ectodominio (gp140)-HR1Nipah de glicoproteína que  
25 envuelve el VIH-1 (1992 Uganda, serotipo 037.8, clado A) producido por  
ingeniería

<400> 68

30 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
1 5 10 15

Ser Tyr Ile Tyr Ala Glu Asp His His His His His Glu Asn Leu  
20 25 30

35 Tyr Phe Gln Gly Ala Glu Asn Leu Trp Val Thr Val Tyr Tyr Gly Val  
35 40 45

Pro Val Trp Lys Asp Ala Glu Thr Thr Leu Phe Cys Ala Ser Asp Ala  
50 55 60

40 Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys  
65 70 75 80

Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Tyr Met Glu Asn Val Thr  
85 90 95

45 Glu Glu Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His Thr  
100 105 110

50 Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Gln Leu  
115 120 125

Thr Pro Leu Cys Val Thr Leu Asp Cys Ser Tyr Asn Ile Thr Asn Asn  
130 135 140

55 Ile Thr Asn Ser Ile Thr Asn Ser Ser Val Asn Met Arg Glu Glu Ile  
145 150 155 160

60 Lys Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp Lys Asn Arg  
165 170 175

ES 2 733 279 T3

Lys Val Tyr Ser Leu Phe Tyr Lys Leu Asp Val Val Gln Ile Asn Asn  
 180 185 190  
**5** Gly Asn Asn Ser Ser Asn Leu Tyr Arg Leu Ile Asn Cys Asn Thr Ser  
 195 200 205  
 Ala Leu Thr Gln Ala Cys Pro Lys Val Thr Phe Glu Pro Ile Pro Ile  
 210 215 220  
**10** Arg Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asp Lys  
 225 230 235 240  
**15** Glu Phe Asn Gly Thr Gly Leu Cys Lys Asn Val Ser Thr Val Gln Cys  
 245 250 255  
 Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly  
 260 265 270  
**20** Ser Leu Ala Glu Gly Lys Val Met Ile Arg Ser Glu Asn Ile Thr Asn  
 275 280 285  
**25** Asn Val Lys Asn Ile Ile Val Gln Leu Asn Glu Thr Val Thr Ile Asn  
 290 295 300  
 Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Val Arg Ile Gly Pro  
 305 310 315 320  
**30** Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln  
 325 330 335  
 Ala His Cys Asn Val Ser Gly Ser Gln Trp Asn Arg Ala Leu His Gln  
 340 345 350  
**35** Val Val Gly Gln Leu Arg Glu Tyr Trp Asn Thr Thr Ile Ile Phe Lys  
 355 360 365  
**40** Asn Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys  
 370 375 380  
 Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Ser Asn  
 385 390 395 400  
**45** Trp Thr His Asn Asp Thr Ala Ser Met Lys Pro Asn Asp Thr Ile Thr  
 405 410 415  
 Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Val Gly  
 420 425 430  
**50** Gln Ala Ile Tyr Ala Pro Pro Ile Gln Gly Val Ile Arg Cys Glu Ser  
 435 440 445  
**55** Asn Ile Thr Gly Leu Ile Leu Thr Arg Asp Gly Gly Gly Asn Ile Asn  
 450 455 460  
 Glu Ser Gln Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp  
 465 470 475 480  
**60** Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Arg Ile Glu Pro Leu Gly

ES 2 733 279 T3

					485					490						495
	Val	Ala	Pro	Thr	Lys	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg	Glu	Lys	Arg
				500					505					510		
5	Ala	Val	Val	Glu	Leu	Gly	Ala	Val	Phe	Ile	Gly	Phe	Leu	Gly	Thr	Ala
			515					520					525			
10	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Ile	Thr	Leu	Thr	Val	Gln	Val	Arg
		530					535					540				
15	Lys	Leu	Leu	Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu	Arg	Ala
	545					550					555					560
	Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	Lys	Leu	Thr	Val	Trp	Gly	Ile	Lys
				565						570					575	
20	Gln	Leu	Gln	Ala	Arg	Val	Leu	Ala	Val	Glu	Arg	Tyr	Leu	Arg	Asp	Gln
				580					585					590		
	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr
			595					600					605			
25	Asn	Val	Pro	Trp	Asn	Ser	Ser	Trp	Ser	Asn	Lys	Ser	Glu	Arg	Glu	Ile
		610					615					620				
30	Trp	Glu	Asn	Met	Thr	Trp	Leu	Gln	Trp	Asp	Lys	Glu	Ile	Ser	Asn	Tyr
	625					630					635					640
	Thr	His	Ile	Ile	Tyr	Glu	Leu	Ile	Glu	Glu	Ser	Gln	Lys	Gln	Gln	Glu
				645					650						655	
35	Lys	Asn	Glu	Gln	Glu	Leu	Leu	Glu	Leu	Asp	Lys	Trp	Ala	Asn	Leu	Trp
				660					665					670		
	Asn	Trp	Phe	Asp	Ile	Ser	Asn	Trp	Leu	Trp	Tyr	Ile	Lys	Ser	Arg	Ile
			675					680					685			
40	Glu	Gly	Arg	Gly	Ser	Gly	Gly	Tyr	Ile	Pro	Glu	Ala	Pro	Arg	Asp	Gly
		690					695					700				
45	Gln	Ala	Tyr	Val	Arg	Lys	Asp	Gly	Glu	Trp	Val	Leu	Leu	Ser	Thr	Phe
	705					710					715					720
	Leu	Asn	Lys	Leu	Lys	Ser	Ser	Ile	Glu	Ser	Thr	Asn	Glu	Ala	Val	Val
				725					730						735	
50	Lys	Leu	Gln	Glu	Thr	Ala	Glu	Lys	Thr	Val	Tyr	Val	Leu	Thr	Ala	Leu
				740					745					750		
	Gln															
55	<210>	69														
	<211>	761														
	<212>	PRT														
	<213>	Secuencia Artificial														
60	<220>															

ES 2 733 279 T3

<223> Proteína fusión ectodominio (gp140)-HR1Nipah de glicoproteína que envuelve el VIH-1 (1992 Uganda, serotipo 037.8, clado A) producido por ingeniería

5 <400> 69  
 Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
 1 5 10 15  
 Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Glu Asp His  
 20 25 30

10 His His His His His Glu Asn Leu Tyr Phe Gln Gly Ala Glu Asn Leu  
 35 40 45

15 Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Glu Thr  
 50 55 60

Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His  
 65 70 75 80

20 Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln  
 85 90 95

25 Glu Ile Tyr Met Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys Asn  
 100 105 110

Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp Gln  
 115 120 125

30 Ser Leu Lys Pro Cys Val Gln Leu Thr Pro Leu Cys Val Thr Leu Asp  
 130 135 140

Cys Ser Tyr Asn Ile Thr Asn Asn Ile Thr Asn Ser Ile Thr Asn Ser  
 145 150 155 160

35 Ser Val Asn Met Arg Glu Glu Ile Lys Asn Cys Ser Phe Asn Met Thr  
 165 170 175

40 Thr Glu Leu Arg Asp Lys Asn Arg Lys Val Tyr Ser Leu Phe Tyr Lys  
 180 185 190

Leu Asp Val Val Gln Ile Asn Asn Gly Asn Asn Ser Ser Asn Leu Tyr  
 195 200 205

45 Arg Leu Ile Asn Cys Asn Thr Ser Ala Leu Thr Gln Ala Cys Pro Lys  
 210 215 220

Val Thr Phe Glu Pro Ile Pro Ile Arg Tyr Cys Ala Pro Ala Gly Tyr  
 225 230 235 240

50 Ala Ile Leu Lys Cys Asn Asp Lys Glu Phe Asn Gly Thr Gly Leu Cys  
 245 250 255

55 Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val  
 260 265 270

Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Gly Lys Val Met  
 275 280 285

60 Ile Arg Ser Glu Asn Ile Thr Asn Asn Val Lys Asn Ile Ile Val Gln  
 290 295 300

ES 2 733 279 T3

5 Leu Asn Glu Thr Val Thr Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr  
 305 310 315 320  
 Arg Lys Ser Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly  
 325 330 335  
 10 Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Val Ser Gly Ser  
 340 345 350  
 Gln Trp Asn Arg Ala Leu His Gln Val Val Gly Gln Leu Arg Glu Tyr  
 355 360 365  
 15 Trp Asn Thr Thr Ile Ile Phe Lys Asn Ser Ser Gly Gly Asp Leu Glu  
 370 375 380  
 20 Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn  
 385 390 395 400  
 Thr Ser Gly Leu Phe Asn Ser Asn Trp Thr His Asn Asp Thr Ala Ser  
 405 410 415  
 25 Met Lys Pro Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile  
 420 425 430  
 Ile Asn Met Trp Gln Arg Val Gly Gln Ala Ile Tyr Ala Pro Pro Ile  
 435 440 445  
 30 Gln Gly Val Ile Arg Cys Glu Ser Asn Ile Thr Gly Leu Ile Leu Thr  
 450 455 460  
 Arg Asp Gly Gly Gly Asn Ile Asn Glu Ser Gln Ile Phe Arg Pro Gly  
 465 470 475 480  
 35 Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys  
 485 490 495  
 40 Val Val Arg Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg  
 500 505 510  
 Arg Val Val Glu Arg Glu Lys Arg Ala Val Val Glu Leu Gly Ala Val  
 515 520 525  
 45 Phe Ile Gly Phe Leu Gly Thr Ala Gly Ser Thr Met Gly Ala Ala Ser  
 530 535 540  
 50 Ile Thr Leu Thr Val Gln Val Arg Lys Leu Leu Ser Gly Ile Val Gln  
 545 550 555 560  
 Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu  
 565 570 575  
 55 Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala  
 580 585 590  
 60 Val Glu Arg Tyr Leu Arg Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys  
 595 600 605

ES 2 733 279 T3

Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp  
610 615 620

5 Ser Asn Lys Ser Glu Arg Glu Ile Trp Glu Asn Met Thr Trp Leu Gln  
625 630 635 640

Trp Asp Lys Glu Ile Ser Asn Tyr Thr His Ile Ile Tyr Glu Leu Ile  
645 650 655

10 Glu Glu Ser Gln Lys Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu  
660 665 670

Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp  
675 680 685

15 Leu Trp Tyr Ile Lys Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr  
690 695 700

20 Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly  
705 710 715 720

Glu Trp Val Leu Leu Ser Thr Phe Leu Asn Lys Leu Lys Ser Ser Ile  
725 730 735

25 Glu Ser Thr Asn Glu Ala Val Val Lys Leu Gln Glu Thr Ala Glu Lys  
740 745 750

Thr Val Tyr Val Leu Thr Ala Leu Gln  
755 760

30 <210> 70  
<211> 720  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> Proteína fusión ectodominio (gp140)-HR1Nipah de glicoproteína que  
envuelve el VIH-1 (1992 Uganda, serotipo 037.8, clado A) producido por  
ingeniería

40 <400> 70  
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
1 5 10 15

45 Ser Tyr Ile Tyr Ala Glu Asp His His His His His His Glu Asn Leu  
20 25 30

Tyr Phe Gln Gly Ala Glu Asn Leu Trp Val Thr Val Tyr Tyr Gly Val  
35 40 45

50 Pro Val Trp Lys Asp Ala Glu Thr Thr Leu Phe Cys Ala Ser Asp Ala  
50 55 60  
Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys  
65 70 75 80

55 Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Tyr Met Glu Asn Val Thr  
85 90 95

60 Glu Glu Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His Thr  
100 105 110



ES 2 733 279 T3

Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Gln Leu  
 115 120 125  
**5** Thr Pro Leu Cys Val Thr Leu Asp Cys Ser Tyr Asn Ile Thr Asn Asn  
 130 135 140  
 Ile Thr Asn Ser Ile Thr Asn Ser Ser Val Asn Met Arg Glu Glu Ile  
 145 150 155 160  
**10** Lys Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp Lys Asn Arg  
 165 170 175  
**15** Lys Val Tyr Ser Leu Phe Tyr Lys Leu Asp Val Val Gln Ile Asn Asn  
 180 185 190  
 Gly Asn Asn Ser Ser Asn Leu Tyr Arg Leu Ile Asn Cys Asn Thr Ser  
 195 200 205  
**20** Ala Leu Thr Gln Ala Cys Pro Lys Val Thr Phe Glu Pro Ile Pro Ile  
 210 215 220  
 Arg Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asp Lys  
 225 230 235 240  
**25** Glu Phe Asn Gly Thr Gly Leu Cys Lys Asn Val Ser Thr Val Gln Cys  
 245 250 255  
**30** Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly  
 260 265 270  
 Ser Leu Ala Glu Gly Lys Val Met Ile Arg Ser Glu Asn Ile Thr Asn  
 275 280 285  
**35** Asn Val Lys Asn Ile Ile Val Gln Leu Asn Glu Thr Val Thr Ile Asn  
 290 295 300  
 Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Val Arg Ile Gly Pro  
 305 310 315 320  
**40** Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln  
 325 330 335  
**45** Ala His Cys Asn Val Ser Gly Ser Gln Trp Asn Arg Ala Leu His Gln  
 340 345 350  
 Val Val Gly Gln Leu Arg Glu Tyr Trp Asn Thr Thr Ile Ile Phe Lys  
 355 360 365  
**50** Asn Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys  
 370 375 380  
**55** Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Ser Asn  
 385 390 395 400  
 Trp Thr His Asn Asp Thr Ala Ser Met Lys Pro Asn Asp Thr Ile Thr  
 405 410 415  
**60** Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Val Gly

ES 2 733 279 T3

				420					425					430			
	Gln	Ala	Ile	Tyr	Ala	Pro	Pro	Ile	Gln	Gly	Val	Ile	Arg	Cys	Glu	Ser	
			435					440					445				
5	Asn	Ile	Thr	Gly	Leu	Ile	Leu	Thr	Arg	Asp	Gly	Gly	Gly	Asn	Ile	Asn	
		450					455					460					
10	Glu	Ser	Gln	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	
	465					470					475					480	
	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Arg	Ile	Glu	Pro	Leu	Gly	
					485					490					495		
15	Val	Ala	Pro	Thr	Lys	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg	Glu	Lys	Arg	
				500					505					510			
	Ala	Val	Val	Glu	Leu	Gly	Ala	Val	Phe	Ile	Gly	Phe	Leu	Gly	Thr	Ala	
20			515					520					525				
	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Ile	Thr	Leu	Thr	Val	Gln	Val	Arg	
		530					535					540					
25	Lys	Leu	Leu	Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu	Arg	Ala	
	545					550					555					560	
	Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	Lys	Leu	Thr	Val	Trp	Gly	Ile	Lys	
					565					570					575		
30	Gln	Leu	Gln	Ala	Arg	Val	Leu	Ala	Val	Glu	Arg	Tyr	Leu	Arg	Asp	Gln	
				580					585					590			
	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	
			595					600					605				
35	Asn	Val	Pro	Trp	Asn	Ser	Ser	Trp	Ser	Asn	Lys	Ser	Glu	Arg	Glu	Ile	
		610					615					620					
40	Trp	Glu	Asn	Met	Thr	Trp	Leu	Gln	Trp	Asp	Lys	Glu	Ile	Ser	Asn	Tyr	
	625					630					635					640	
	Thr	His	Ile	Ile	Tyr	Glu	Leu	Ile	Glu	Glu	Ser	Gln	Lys	Gln	Gln	Glu	
					645					650					655		
45	Lys	Asn	Glu	Gln	Glu	Leu	Leu	Glu	Leu	Asp	Lys	Trp	Ala	Asn	Leu	Trp	
				660					665					670			
50	Asn	Trp	Phe	Asp	Ile	Ser	Asn	Trp	Leu	Trp	Tyr	Ile	Lys	Ser	Asn	Ile	
			675					680					685				
	Asn	Lys	Leu	Lys	Ser	Ser	Ile	Glu	Ser	Thr	Asn	Glu	Ala	Val	Val	Lys	
		690					695					700					
55	Leu	Gln	Glu	Thr	Ala	Glu	Lys	Thr	Val	Tyr	Val	Leu	Thr	Ala	Leu	Gln	
	705					710					715					720	
	<210>		71														
	<211>		728														
60	<212>		PRT														

ES 2 733 279 T3

<213> Secuencia Artificial

<220>

5 <223> Proteína fusión ectodominio (gp140)-HR1Nipah de glicoproteína que envuelve el VIH-1 (1992 Uganda, serotipo 037.8, clado A) producido por ingeniería

<400> 71

10	Met	Arg	Val	Arg	Gly	Ile	Gln	Arg	Asn	Cys	Gln	His	Leu	Trp	Arg	Trp
	1				5					10					15	
	Gly	Thr	Leu	Ile	Leu	Gly	Met	Leu	Met	Ile	Cys	Ser	Ala	Glu	Asp	His
				20					25					30		
15	His	His	His	His	His	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Ala	Glu	Asn	Leu
			35					40					45			
	Trp	Val	Thr	Val	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Lys	Asp	Ala	Glu	Thr
20		50					55					60				
	Thr	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys	Ala	Tyr	Asp	Thr	Glu	Val	His
	65					70					75					80
	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro	Gln
25					85					90					95	
	Glu	Ile	Tyr	Met	Glu	Asn	Val	Thr	Glu	Glu	Phe	Asn	Met	Trp	Lys	Asn
				100					105					110		
30	Asn	Met	Val	Glu	Gln	Met	His	Thr	Asp	Ile	Ile	Ser	Leu	Trp	Asp	Gln
			115					120					125			
	Ser	Leu	Lys	Pro	Cys	Val	Gln	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu	Asp
35		130					135					140				
	Cys	Ser	Tyr	Asn	Ile	Thr	Asn	Asn	Ile	Thr	Asn	Ser	Ile	Thr	Asn	Ser
40		145				150					155					160
	Ser	Val	Asn	Met	Arg	Glu	Glu	Ile	Lys	Asn	Cys	Ser	Phe	Asn	Met	Thr
				165						170					175	
	Thr	Glu	Leu	Arg	Asp	Lys	Asn	Arg	Lys	Val	Tyr	Ser	Leu	Phe	Tyr	Lys
45				180					185					190		
	Leu	Asp	Val	Val	Gln	Ile	Asn	Asn	Gly	Asn	Asn	Ser	Ser	Asn	Leu	Tyr
			195					200					205			
50	Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser	Ala	Leu	Thr	Gln	Ala	Cys	Pro	Lys
		210					215					220				
	Val	Thr	Phe	Glu	Pro	Ile	Pro	Ile	Arg	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr
55						230					235					240
	Ala	Ile	Leu	Lys	Cys	Asn	Asp	Lys	Glu	Phe	Asn	Gly	Thr	Gly	Leu	Cys
				245						250					255	
	Lys	Asn	Val	Ser	Thr	Val	Gln	Cys	Thr	His	Gly	Ile	Arg	Pro	Val	Val
60				260					265						270	

ES 2 733 279 T3

Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Gly Lys Val Met  
 275 280 285

5 Ile Arg Ser Glu Asn Ile Thr Asn Asn Val Lys Asn Ile Ile Val Gln  
 290 295 300

Leu Asn Glu Thr Val Thr Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr  
 305 310 315 320

10 Arg Lys Ser Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly  
 325 330 335

15 Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Val Ser Gly Ser  
 340 345 350

Gln Trp Asn Arg Ala Leu His Gln Val Val Gly Gln Leu Arg Glu Tyr  
 355 360 365

20 Trp Asn Thr Thr Ile Ile Phe Lys Asn Ser Ser Gly Gly Asp Leu Glu  
 370 375 380

Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn  
 385 390 395 400

25 Thr Ser Gly Leu Phe Asn Ser Asn Trp Thr His Asn Asp Thr Ala Ser  
 405 410 415

30 Met Lys Pro Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile  
 420 425 430

Ile Asn Met Trp Gln Arg Val Gly Gln Ala Ile Tyr Ala Pro Pro Ile  
 435 440 445

35 Gln Gly Val Ile Arg Cys Glu Ser Asn Ile Thr Gly Leu Ile Leu Thr  
 450 455 460

40 Arg Asp Gly Gly Gly Asn Ile Asn Glu Ser Gln Ile Phe Arg Pro Gly  
 465 470 475 480

Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys  
 485 490 495

45 Val Val Arg Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg  
 500 505 510

Arg Val Val Glu Arg Glu Lys Arg Ala Val Val Glu Leu Gly Ala Val  
 515 520 525

50 Phe Ile Gly Phe Leu Gly Thr Ala Gly Ser Thr Met Gly Ala Ala Ser  
 530 535 540

Ile Thr Leu Thr Val Gln Val Arg Lys Leu Leu Ser Gly Ile Val Gln  
 545 550 555 560

Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu  
 565 570 575

60 Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala

ES 2 733 279 T3

				580					585					590			
	Val	Glu	Arg	Tyr	Leu	Arg	Asp	Gln	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	
			595					600					605				
5	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Asn	Val	Pro	Trp	Asn	Ser	Ser	Trp	
		610					615					620					
10	Ser	Asn	Lys	Ser	Glu	Arg	Glu	Ile	Trp	Glu	Asn	Met	Thr	Trp	Leu	Gln	
	625					630					635					640	
	Trp	Asp	Lys	Glu	Ile	Ser	Asn	Tyr	Thr	His	Ile	Ile	Tyr	Glu	Leu	Ile	
				645						650					655		
15	Glu	Glu	Ser	Gln	Lys	Gln	Gln	Glu	Lys	Asn	Glu	Gln	Glu	Leu	Leu	Glu	
				660					665					670			
	Leu	Asp	Lys	Trp	Ala	Asn	Leu	Trp	Asn	Trp	Phe	Asp	Ile	Ser	Asn	Trp	
20			675					680					685				
	Leu	Trp	Tyr	Ile	Lys	Ser	Asn	Ile	Asn	Lys	Leu	Lys	Ser	Ser	Ile	Glu	
		690					695					700					
25	Ser	Thr	Asn	Glu	Ala	Val	Val	Lys	Leu	Gln	Glu	Thr	Ala	Glu	Lys	Thr	
	705					710					715					720	
	Val	Tyr	Val	Leu	Thr	Ala	Leu	Gln									
30					725												
	<210>	72															
	<211>	741															
	<212>	PRT															
	<213>	Secuencia Artificial															
35	<220>																
	<223>	Proteína fusión ectodominio (gp140)-HR1HXB2 de glicoproteína que envuelve el VIH-1 clado C 1997 producido por ingeniería															
	<400>	72															
40	Met	Lys	Phe	Leu	Val	Asn	Val	Ala	Leu	Val	Phe	Met	Val	Val	Tyr	Ile	
	1				5					10					15		
	Ser	Tyr	Ile	Tyr	Ala	Glu	Asp	His	His	His	His	His	His	Glu	Asn	Leu	
45				20				25						30			
	Tyr	Phe	Gln	Gly	Ala	Glu	Asn	Leu	Trp	Val	Gly	Asn	Met	Trp	Val	Thr	
			35					40					45				
50	Val	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Thr	Asp	Ala	Lys	Thr	Thr	Leu	Phe	
		50					55					60					
	Cys	Ala	Ser	Asp	Thr	Lys	Ala	Tyr	Asp	Arg	Glu	Val	His	Asn	Val	Trp	
	65					70					75					80	
55	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro	Gln	Glu	Ile	Val	
					85					90					95		
	Leu	Glu	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp	Lys	Asn	Asp	Met	Val	
60				100					105					110			

ES 2 733 279 T3

Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys  
 115 120 125  
 5 Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn  
 130 135 140  
 Ala Thr Phe Lys Asn Asn Val Thr Asn Asp Met Asn Lys Glu Ile Arg  
 145 150 155 160  
 10 Asn Cys Ser Phe Asn Thr Thr Thr Glu Ile Arg Asp Lys Lys Gln Gln  
 165 170 175  
 Gly Tyr Ala Leu Phe Tyr Arg Pro Asp Ile Val Leu Leu Lys Glu Asn  
 180 185 190  
 15 Arg Asn Asn Ser Asn Asn Ser Glu Tyr Ile Leu Ile Asn Cys Asn Ala  
 195 200 205  
 20 Ser Thr Ile Thr Gln Ala Cys Pro Lys Val Asn Phe Asp Pro Ile Pro  
 210 215 220  
 Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn  
 225 230 235 240  
 25 Lys Thr Phe Ser Gly Lys Gly Pro Cys Asn Asn Val Ser Thr Val Gln  
 245 250 255  
 30 Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn  
 260 265 270  
 Gly Ser Leu Ala Glu Lys Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr  
 275 280 285  
 35 Asp Asn Val Lys Thr Ile Ile Val His Leu Asn Lys Ser Val Glu Ile  
 290 295 300  
 Val Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Met Arg Ile Gly  
 305 310 315 320  
 40 Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg  
 325 330 335  
 45 Gln Ala Tyr Cys Asn Ile Ser Gly Ser Lys Trp Asn Glu Thr Leu Lys  
 340 345 350  
 Arg Val Lys Glu Lys Leu Gln Glu Asn Tyr Asn Asn Asn Lys Thr Ile  
 355 360 365  
 50 Lys Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser  
 370 375 380  
 Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Thr Arg Leu Phe  
 385 390 395 400  
 55 Asn Asn Asn Ala Thr Glu Asp Glu Thr Ile Thr Leu Pro Cys Arg Ile  
 405 410 415  
 60 Lys Gln Ile Ile Asn Met Trp Gln Gly Val Gly Arg Ala Met Tyr Ala  
 420 425 430

ES 2 733 279 T3

Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu  
 435 440 445  
**5** Leu Leu Val Arg Asp Gly Gly Glu Asp Asn Lys Thr Glu Glu Ile Phe  
 450 455 460  
 Arg Pro Gly Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr  
 465 470 475 480  
**10** Lys Tyr Lys Val Ile Glu Leu Lys Pro Leu Gly Ile Ala Pro Thr Gly  
 485 490 495  
**15** Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly  
 500 505 510  
 Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala  
 515 520 525  
**20** Ala Ser Leu Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Ser Ile  
 530 535 540  
**25** Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His  
 545 550 555 560  
 Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val  
 565 570 575  
**30** Leu Ala Ile Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp  
 580 585 590  
 Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser  
 595 600 605  
**35** Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asn Asn Met Thr Trp  
 610 615 620  
**40** Met Glu Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg  
 625 630 635 640  
 Leu Leu Glu Asp Ser Gln Thr Gln Gln Glu Lys Asn Glu Lys Asp Leu  
 645 650 655  
**45** Leu Ala Leu Asp Ser Trp Lys Asn Leu Trp Ser Trp Phe Asp Ile Ser  
 660 665 670  
 Asn Trp Leu Trp Tyr Ile Lys Ser Gly Tyr Ile Pro Glu Ala Pro Arg  
 675 680 685  
**50** Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser  
 690 695 700  
**55** Thr Phe Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala  
 705 710 715 720  
 Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys  
 725 730 735  
**60** Gln Leu Gln Ala Arg

ES 2 733 279 T3

740

<210> 73  
 <211> 749  
 5 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Proteína fusión ectodominio (gp140)-HR1HXB2 de glicoproteína que  
 10 envuelve el VIH-1 clado C 1997 producido por ingeniería  
  
 <400> 73  
 Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
 1 5 10 15  
 15 Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Glu Asp His  
 20 20 25 30  
 His His His His His Glu Asn Leu Tyr Phe Gln Gly Ala Glu Asn Leu  
 20 35 40 45  
 Trp Val Gly Asn Met Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp  
 50 55 60  
 25 Thr Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Thr Lys Ala Tyr  
 65 70 75 80  
 Asp Arg Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr  
 30 85 90 95  
 Asp Pro Asn Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe  
 100 105 110  
 35 Asn Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile  
 115 120 125  
 Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu  
 130 135 140  
 40 Cys Val Thr Leu His Cys Thr Asn Ala Thr Phe Lys Asn Asn Val Thr  
 145 150 155 160  
 Asn Asp Met Asn Lys Glu Ile Arg Asn Cys Ser Phe Asn Thr Thr Thr  
 45 165 170 175  
 Glu Ile Arg Asp Lys Lys Gln Gln Gly Tyr Ala Leu Phe Tyr Arg Pro  
 180 185 190  
 50 Asp Ile Val Leu Leu Lys Glu Asn Arg Asn Asn Ser Asn Asn Ser Glu  
 195 200 205  
 Tyr Ile Leu Ile Asn Cys Asn Ala Ser Thr Ile Thr Gln Ala Cys Pro  
 210 215 220  
 55 Lys Val Asn Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly  
 225 230 235 240  
 Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Ser Gly Lys Gly Pro  
 245 250 255  
 60



ES 2 733 279 T3

Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val  
 260 265 270  
 5 Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Glu Ile  
 275 280 285  
 10 Ile Ile Arg Ser Glu Asn Leu Thr Asp Asn Val Lys Thr Ile Ile Val  
 290 295 300  
 His Leu Asn Lys Ser Val Glu Ile Val Cys Thr Arg Pro Asn Asn Asn  
 305 310 315 320  
 15 Thr Arg Lys Ser Met Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr  
 325 330 335  
 Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala Tyr Cys Asn Ile Ser Gly  
 340 345 350  
 20 Ser Lys Trp Asn Glu Thr Leu Lys Arg Val Lys Glu Lys Leu Gln Glu  
 355 360 365  
 Asn Tyr Asn Asn Asn Lys Thr Ile Lys Phe Ala Pro Ser Ser Gly Gly  
 370 375 380  
 25 Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe  
 385 390 395 400  
 30 Tyr Cys Asn Thr Thr Arg Leu Phe Asn Asn Asn Ala Thr Glu Asp Glu  
 405 410 415  
 Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln  
 420 425 430  
 35 Gly Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr  
 435 440 445  
 Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Val Arg Asp Gly Gly Glu  
 450 455 460  
 40 Asp Asn Lys Thr Glu Glu Ile Phe Arg Pro Gly Gly Gly Asn Met Lys  
 465 470 475 480  
 Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Ile Glu Leu Lys  
 485 490 495  
 Pro Leu Gly Ile Ala Pro Thr Gly Ala Lys Arg Arg Val Val Glu Arg  
 500 505 510  
 50 Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly  
 515 520 525  
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Leu Thr Leu Thr Val Gln  
 530 535 540  
 55 Ala Arg Gln Leu Leu Ser Ser Ile Val Gln Gln Gln Ser Asn Leu Leu  
 545 550 555 560  
 Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly  
 565 570 575  
 60

ES 2 733 279 T3

5 Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys  
 580 585 590  
 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys  
 595 600 605  
 10 Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Thr  
 610 615 620  
 Asp Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile Ser  
 625 630 635 640  
 15 Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Thr Gln  
 645 650 655  
 20 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn  
 660 665 670  
 Leu Trp Ser Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ser  
 675 680 685  
 25 Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys  
 690 695 700  
 Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Ser Gly Ile Val Gln  
 705 710 715 720  
 30 Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu  
 725 730 735  
 Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg  
 740 745  
 35 <210> 74  
 <211> 716  
 <212> PRT  
 <213> Secuencia Artificial  
 40 <220>  
 <223> Proteína fusión ectodominio (gp140)-HR1HXB2 de glicoproteína que  
 envuelve el VIH-1 clado C 1997 producido por ingeniería  
 45 <400> 74  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 50 Ser Tyr Ile Tyr Ala Glu Asp His His His His His His Glu Asn Leu  
 20 25 30  
 Tyr Phe Gln Gly Ala Glu Asn Leu Trp Val Gly Asn Met Trp Val Thr  
 35 40 45  
 55 Val Tyr Tyr Gly Val Pro Val Trp Thr Asp Ala Lys Thr Thr Leu Phe  
 50 55 60  
 Cys Ala Ser Asp Thr Lys Ala Tyr Asp Arg Glu Val His Asn Val Trp  
 65 70 75 80  
 60

ES 2 733 279 T3

Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val  
85 90 95

5 Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met Val  
100 105 110

Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys  
115 120 125

10 Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn  
130 135 140

Ala Thr Phe Lys Asn Asn Val Thr Asn Asp Met Asn Lys Glu Ile Arg  
145 150 155 160

15 Asn Cys Ser Phe Asn Thr Thr Thr Glu Ile Arg Asp Lys Lys Gln Gln  
165 170 175

20 Gly Tyr Ala Leu Phe Tyr Arg Pro Asp Ile Val Leu Leu Lys Glu Asn  
180 185 190

Arg Asn Asn Ser Asn Asn Ser Glu Tyr Ile Leu Ile Asn Cys Asn Ala  
195 200 205

25 Ser Thr Ile Thr Gln Ala Cys Pro Lys Val Asn Phe Asp Pro Ile Pro  
210 215 220

Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn  
225 230 235 240

30 Lys Thr Phe Ser Gly Lys Gly Pro Cys Asn Asn Val Ser Thr Val Gln  
245 250 255

35 Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn  
260 265 270

Gly Ser Leu Ala Glu Lys Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr  
275 280 285

40 Asp Asn Val Lys Thr Ile Ile Val His Leu Asn Lys Ser Val Glu Ile  
290 295 300

Val Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Met Arg Ile Gly  
305 310 315 320

45 Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg  
325 330 335

50 Gln Ala Tyr Cys Asn Ile Ser Gly Ser Lys Trp Asn Glu Thr Leu Lys  
340 345 350

Arg Val Lys Glu Lys Leu Gln Glu Asn Tyr Asn Asn Asn Lys Thr Ile  
355 360 365

55 Lys Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser  
370 375 380

60 Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Thr Arg Leu Phe  
385 390 395 400

ES 2 733 279 T3

Asn Asn Asn Ala Thr Glu Asp Glu Thr Ile Thr Leu Pro Cys Arg Ile  
405 410 415

5 Lys Gln Ile Ile Asn Met Trp Gln Gly Val Gly Arg Ala Met Tyr Ala  
420 425 430

Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu  
435 440 445

10 Leu Leu Val Arg Asp Gly Gly Glu Asp Asn Lys Thr Glu Glu Ile Phe  
450 455 460

15 Arg Pro Gly Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr  
465 470 475 480

Lys Tyr Lys Val Ile Glu Leu Lys Pro Leu Gly Ile Ala Pro Thr Gly  
485 490 495

20 Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly  
500 505 510

Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala  
515 520 525

25 Ala Ser Leu Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Ser Ile  
530 535 540

30 Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His  
545 550 555 560

Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val  
565 570 575

35 Leu Ala Ile Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp  
580 585 590

Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser  
595 600 605

40 Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asn Asn Met Thr Trp  
610 615 620

45 Met Glu Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg  
625 630 635 640

Leu Leu Glu Asp Ser Gln Thr Gln Gln Glu Lys Asn Glu Lys Asp Leu  
645 650 655

50 Leu Ala Leu Asp Ser Trp Lys Asn Leu Trp Ser Trp Phe Asp Ile Ser  
660 665 670

Asn Trp Leu Trp Tyr Ile Lys Ser Gly Ile Ser Gly Ile Val Gln Gln  
675 680 685

Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln  
690 695 700

60 Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg

ES 2 733 279 T3

705                      710                      715

<210>    75  
<211>    724  
**5**    <212>    PRT  
     <213>    Secuencia Artificial

<220>  
<223>    Proteína fusión ectodominio (gp140)-HR1HXB2 de glicoproteína que  
**10**    envuelve el VIH-1 clado C 1997 producido por ingeniería

<400>    75

**15**    Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
     1                      5                      10                      15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Glu Asp His  
                                20                      25                      30

**20**    His His His His His Glu Asn Leu Tyr Phe Gln Gly Ala Glu Asn Leu  
                                35                      40                      45

Trp Val Gly Asn Met Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp  
                                50                      55                      60

**25**    Thr Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Thr Lys Ala Tyr  
     65                      70                      75                      80

Asp Arg Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr  
                                85                      90                      95

**30**    Asp Pro Asn Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe  
                                100                      105                      110

**35**    Asn Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile  
                                115                      120                      125

Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu  
                                130                      135                      140

**40**    Cys Val Thr Leu His Cys Thr Asn Ala Thr Phe Lys Asn Asn Val Thr  
     145                      150                      155                      160

Asn Asp Met Asn Lys Glu Ile Arg Asn Cys Ser Phe Asn Thr Thr Thr  
                                165                      170                      175

**45**    Glu Ile Arg Asp Lys Lys Gln Gln Gly Tyr Ala Leu Phe Tyr Arg Pro  
                                180                      185                      190

**50**    Asp Ile Val Leu Leu Lys Glu Asn Arg Asn Asn Ser Asn Asn Ser Glu  
                                195                      200                      205

Tyr Ile Leu Ile Asn Cys Asn Ala Ser Thr Ile Thr Gln Ala Cys Pro  
     210                      215                      220

**55**    Lys Val Asn Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly  
     225                      230                      235                      240

Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Ser Gly Lys Gly Pro  
                                245                      250                      255

**60**

ES 2 733 279 T3

Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val  
 260 265 270  
 5 Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Glu Ile  
 275 280 285  
 Ile Ile Arg Ser Glu Asn Leu Thr Asp Asn Val Lys Thr Ile Ile Val  
 290 295 300  
 10 His Leu Asn Lys Ser Val Glu Ile Val Cys Thr Arg Pro Asn Asn Asn  
 305 310 315 320  
 Thr Arg Lys Ser Met Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr  
 325 330 335  
 15 Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala Tyr Cys Asn Ile Ser Gly  
 340 345 350  
 20 Ser Lys Trp Asn Glu Thr Leu Lys Arg Val Lys Glu Lys Leu Gln Glu  
 355 360 365  
 Asn Tyr Asn Asn Asn Lys Thr Ile Lys Phe Ala Pro Ser Ser Gly Gly  
 370 375 380  
 25 Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe  
 385 390 395 400  
 Tyr Cys Asn Thr Thr Arg Leu Phe Asn Asn Asn Ala Thr Glu Asp Glu  
 405 410 415  
 30 Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln  
 420 425 430  
 35 Gly Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr  
 435 440 445  
 Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Val Arg Asp Gly Gly Glu  
 450 455 460  
 40 Asp Asn Lys Thr Glu Glu Ile Phe Arg Pro Gly Gly Gly Asn Met Lys  
 465 470 475 480  
 45 Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Ile Glu Leu Lys  
 485 490 495  
 Pro Leu Gly Ile Ala Pro Thr Gly Ala Lys Arg Arg Val Val Glu Arg  
 500 505 510  
 50 Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly  
 515 520 525  
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Leu Thr Leu Thr Val Gln  
 530 535 540  
 55 Ala Arg Gln Leu Leu Ser Ser Ile Val Gln Gln Gln Ser Asn Leu Leu  
 545 550 555 560  
 60 Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly  
 565 570 575

ES 2 733 279 T3

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys  
580 585 590

5 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys  
595 600 605

Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Thr  
610 615 620

10 Asp Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile Ser  
625 630 635 640

15 Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Thr Gln  
645 650 655

Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn  
660 665 670

20 Leu Trp Ser Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ser  
675 680 685

Gly Ile Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile  
690 695 700

25 Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln  
705 710 715 720

30 Leu Gln Ala Arg

<210> 76  
<211> 739  
<212> PRT  
35 <213> Secuencia Artificial  
<220>

40 <223> Proteína fusión ectodominio (gp140)-HR1Nipah de glicoproteína que  
envuelve el VIH-1 clado C 1997 producido por ingeniería

<400> 76  
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
1 5 10 15

45 Ser Tyr Ile Tyr Ala Glu Asp His His His His His His Glu Asn Leu  
20 25 30

50 Tyr Phe Gln Gly Ala Glu Asn Leu Trp Val Gly Asn Met Trp Val Thr  
35 40 45

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

55 Cys Ala Ser Asp Thr Lys Ala Tyr Asp Arg Glu Val His Asn Val Trp  
65 70 75 80

60 Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val  
85 90 95

ES 2 733 279 T3

Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met Val  
 100 105 110  
 5 Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys  
 115 120 125  
 Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn  
 130 135 140  
 10 Ala Thr Phe Lys Asn Asn Val Thr Asn Asp Met Asn Lys Glu Ile Arg  
 145 150 155 160  
 Asn Cys Ser Phe Asn Thr Thr Thr Glu Ile Arg Asp Lys Lys Gln Gln  
 165 170 175  
 15 Gly Tyr Ala Leu Phe Tyr Arg Pro Asp Ile Val Leu Leu Lys Glu Asn  
 180 185 190  
 Arg Asn Asn Ser Asn Asn Ser Glu Tyr Ile Leu Ile Asn Cys Asn Ala  
 195 200 205  
 20 Ser Thr Ile Thr Gln Ala Cys Pro Lys Val Asn Phe Asp Pro Ile Pro  
 210 215 220  
 25 Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn  
 225 230 235 240  
 Lys Thr Phe Ser Gly Lys Gly Pro Cys Asn Asn Val Ser Thr Val Gln  
 245 250 255  
 30 Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn  
 260 265 270  
 Gly Ser Leu Ala Glu Lys Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr  
 275 280 285  
 35 Asp Asn Val Lys Thr Ile Ile Val His Leu Asn Lys Ser Val Glu Ile  
 290 295 300  
 40 Val Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Met Arg Ile Gly  
 305 310 315 320  
 Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg  
 325 330 335  
 45 Gln Ala Tyr Cys Asn Ile Ser Gly Ser Lys Trp Asn Glu Thr Leu Lys  
 340 345 350  
 Arg Val Lys Glu Lys Leu Gln Glu Asn Tyr Asn Asn Asn Lys Thr Ile  
 355 360 365  
 Lys Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser  
 370 375 380  
 55 Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Thr Arg Leu Phe  
 385 390 395 400  
 Asn Asn Asn Ala Thr Glu Asp Glu Thr Ile Thr Leu Pro Cys Arg Ile  
 405 410 415  
 60



ES 2 733 279 T3

Lys Gln Ile Ile Asn Met Trp Gln Gly Val Gly Arg Ala Met Tyr Ala  
 420 425 430  
 5 Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu  
 435 440 445  
 Leu Leu Val Arg Asp Gly Gly Glu Asp Asn Lys Thr Glu Glu Ile Phe  
 450 455 460  
 10 Arg Pro Gly Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr  
 465 470 475 480  
 Lys Tyr Lys Val Ile Glu Leu Lys Pro Leu Gly Ile Ala Pro Thr Gly  
 485 490 495  
 15 Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly  
 500 505 510  
 20 Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala  
 515 520 525  
 Ala Ser Leu Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Ser Ile  
 530 535 540  
 25 Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His  
 545 550 555 560  
 30 Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val  
 565 570 575  
 Leu Ala Ile Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp  
 580 585 590  
 35 Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser  
 595 600 605  
 Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asn Asn Met Thr Trp  
 610 615 620  
 40 Met Glu Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg  
 625 630 635 640  
 45 Leu Leu Glu Asp Ser Gln Thr Gln Gln Glu Lys Asn Glu Lys Asp Leu  
 645 650 655  
 Leu Ala Leu Asp Ser Trp Lys Asn Leu Trp Ser Trp Phe Asp Ile Ser  
 660 665 670  
 50 Asn Trp Leu Trp Tyr Ile Lys Ser Gly Tyr Ile Pro Glu Ala Pro Arg  
 675 680 685  
 Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser  
 690 695 700  
 55 Thr Phe Leu Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn Glu Ala  
 705 710 715 720  
 60 Val Val Lys Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val Leu Thr  
 725 730 735

ES 2 733 279 T3

Ala Leu Gln

5 <210> 77  
 <211> 747  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Proteína fusión ectodominio (gp140)-HR1Nipah de glicoproteína que  
 envuelve el VIH-1 clado C 1997 producido por ingeniería  
 <400> 77

15 Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
 1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Glu Asp His  
 20 20 25 30

His His His His His Glu Asn Leu Tyr Phe Gln Gly Ala Glu Asn Leu  
 35 40 45

25 Trp Val Gly Asn Met Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp  
 50 55 60

Thr Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Thr Lys Ala Tyr  
 65 70 75 80

30 Asp Arg Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr  
 85 90 95

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

35 Asn Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile  
 115 120 125

40 Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu  
 130 135 140

Cys Val Thr Leu His Cys Thr Asn Ala Thr Phe Lys Asn Asn Val Thr  
 145 150 155 160

45 Asn Asp Met Asn Lys Glu Ile Arg Asn Cys Ser Phe Asn Thr Thr Thr  
 165 170 175

Glu Ile Arg Asp Lys Lys Gln Gln Gly Tyr Ala Leu Phe Tyr Arg Pro  
 180 185 190

50 Asp Ile Val Leu Leu Lys Glu Asn Arg Asn Asn Ser Asn Asn Ser Glu  
 195 200 205

Tyr Ile Leu Ile Asn Cys Asn Ala Ser Thr Ile Thr Gln Ala Cys Pro  
 210 215 220

Lys Val Asn Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly  
 225 230 235 240

60 Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Ser Gly Lys Gly Pro

ES 2 733 279 T3

					245					250					255			
	Cys	Asn	Asn	Val	Ser	Thr	Val	Gln	Cys	Thr	His	Gly	Ile	Lys	Pro	Val		
				260					265					270				
5	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Lys	Glu	Ile		
			275					280					285					
10	Ile	Ile	Arg	Ser	Glu	Asn	Leu	Thr	Asp	Asn	Val	Lys	Thr	Ile	Ile	Val		
		290					295					300						
15	His	Leu	Asn	Lys	Ser	Val	Glu	Ile	Val	Cys	Thr	Arg	Pro	Asn	Asn	Asn		
	305					310					315					320		
	Thr	Arg	Lys	Ser	Met	Arg	Ile	Gly	Pro	Gly	Gln	Thr	Phe	Tyr	Ala	Thr		
					325					330					335			
20	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln	Ala	Tyr	Cys	Asn	Ile	Ser	Gly		
				340					345					350				
	Ser	Lys	Trp	Asn	Glu	Thr	Leu	Lys	Arg	Val	Lys	Glu	Lys	Leu	Gln	Glu		
			355					360					365					
25	Asn	Tyr	Asn	Asn	Asn	Lys	Thr	Ile	Lys	Phe	Ala	Pro	Ser	Ser	Gly	Gly		
		370					375					380						
30	Asp	Leu	Glu	Ile	Thr	Thr	His	Ser	Phe	Asn	Cys	Arg	Gly	Glu	Phe	Phe		
	385					390					395					400		
	Tyr	Cys	Asn	Thr	Thr	Arg	Leu	Phe	Asn	Asn	Asn	Ala	Thr	Glu	Asp	Glu		
					405					410					415			
35	Thr	Ile	Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln		
				420					425					430				
	Gly	Val	Gly	Arg	Ala	Met	Tyr	Ala	Pro	Pro	Ile	Ala	Gly	Asn	Ile	Thr		
			435					440					445					
40	Cys	Lys	Ser	Asn	Ile	Thr	Gly	Leu	Leu	Leu	Val	Arg	Asp	Gly	Gly	Glu		
		450					455					460						
45	Asp	Asn	Lys	Thr	Glu	Glu	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asn	Met	Lys		
	465					470					475					480		
	Asp	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Ile	Glu	Leu	Lys		
					485					490					495			
50	Pro	Leu	Gly	Ile	Ala	Pro	Thr	Gly	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg		
				500					505					510				
	Glu	Lys	Arg	Ala	Val	Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu	Gly		
			515					520					525					
55	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Leu	Thr	Leu	Thr	Val	Gln		
		530					535					540						
60	Ala	Arg	Gln	Leu	Leu	Ser	Ser	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu		
	545					550					555					560		

ES 2 733 279 T3

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly  
 565 570 575  
 5 Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys  
 580 585 590  
 10 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys  
 595 600 605  
 Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Thr  
 610 615 620  
 15 Asp Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile Ser  
 625 630 635 640  
 20 Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Thr Gln  
 645 650 655  
 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn  
 660 665 670  
 Leu Trp Ser Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ser  
 675 680 685  
 25 Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys  
 690 695 700  
 30 Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Asn Lys Leu Lys Ser  
 705 710 715 720  
 Ser Ile Glu Ser Thr Asn Glu Ala Val Val Lys Leu Gln Glu Thr Ala  
 725 730 735  
 35 Glu Lys Thr Val Tyr Val Leu Thr Ala Leu Gln  
 740 745  
 <210> 78  
 <211> 714  
 <212> PRT  
 40 <213> Secuencia Artificial  
 <220>  
 <223> Proteína fusión ectodominio (gp140)-HR1Nipah de glicoproteína que  
 envuelve el VIH-1 clado C 1997 producido por ingeniería  
 45 <400> 78  
 Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
 1 5 10 15  
 50 Ser Tyr Ile Tyr Ala Glu Asp His His His His His Glu Asn Leu  
 20 25 30  
 Tyr Phe Gln Gly Ala Glu Asn Leu Trp Val Gly Asn Met Trp Val Thr  
 35 40 45  
 55 Val Tyr Tyr Gly Val Pro Val Trp Thr Asp Ala Lys Thr Thr Leu Phe  
 50 55 60  
 60 Cys Ala Ser Asp Thr Lys Ala Tyr Asp Arg Glu Val His Asn Val Trp  
 65 70 75 80

ES 2 733 279 T3

5 Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val  
 85 90 95  
 Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met Val  
 100 105 110  
 10 Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys  
 115 120 125  
 Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn  
 130 135 140  
 15 Ala Thr Phe Lys Asn Asn Val Thr Asn Asp Met Asn Lys Glu Ile Arg  
 145 150 155 160  
 20 Asn Cys Ser Phe Asn Thr Thr Thr Glu Ile Arg Asp Lys Lys Gln Gln  
 165 170 175  
 Gly Tyr Ala Leu Phe Tyr Arg Pro Asp Ile Val Leu Leu Lys Glu Asn  
 180 185 190  
 25 Arg Asn Asn Ser Asn Asn Ser Glu Tyr Ile Leu Ile Asn Cys Asn Ala  
 195 200 205  
 Ser Thr Ile Thr Gln Ala Cys Pro Lys Val Asn Phe Asp Pro Ile Pro  
 210 215 220  
 30 Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn  
 225 230 235 240  
 35 Lys Thr Phe Ser Gly Lys Gly Pro Cys Asn Asn Val Ser Thr Val Gln  
 245 250 255  
 Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn  
 260 265 270  
 40 Gly Ser Leu Ala Glu Lys Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr  
 275 280 285  
 Asp Asn Val Lys Thr Ile Ile Val His Leu Asn Lys Ser Val Glu Ile  
 290 295 300  
 45 Val Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Met Arg Ile Gly  
 305 310 315 320  
 50 Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg  
 325 330 335  
 Gln Ala Tyr Cys Asn Ile Ser Gly Ser Lys Trp Asn Glu Thr Leu Lys  
 340 345 350  
 55 Arg Val Lys Glu Lys Leu Gln Glu Asn Tyr Asn Asn Asn Lys Thr Ile  
 355 360 365  
 60 Lys Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser  
 370 375 380

ES 2 733 279 T3

Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Thr Arg Leu Phe  
 385 390 395 400  
 5 Asn Asn Asn Ala Thr Glu Asp Glu Thr Ile Thr Leu Pro Cys Arg Ile  
 405 410 415  
 Lys Gln Ile Ile Asn Met Trp Gln Gly Val Gly Arg Ala Met Tyr Ala  
 420 425 430  
 10 Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu  
 435 440 445  
 15 Leu Leu Val Arg Asp Gly Gly Glu Asp Asn Lys Thr Glu Glu Ile Phe  
 450 455 460  
 Arg Pro Gly Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr  
 465 470 475 480  
 20 Lys Tyr Lys Val Ile Glu Leu Lys Pro Leu Gly Ile Ala Pro Thr Gly  
 485 490 495  
 25 Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly  
 500 505 510  
 Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala  
 515 520 525  
 30 Ala Ser Leu Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Ser Ile  
 530 535 540  
 Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His  
 545 550 555 560  
 35 Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val  
 565 570 575  
 Leu Ala Ile Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp  
 580 585 590  
 40 Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser  
 595 600 605  
 Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asn Asn Met Thr Trp  
 610 615 620  
 45 Met Glu Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg  
 625 630 635 640  
 50 Leu Leu Glu Asp Ser Gln Thr Gln Gln Glu Lys Asn Glu Lys Asp Leu  
 645 650 655  
 Leu Ala Leu Asp Ser Trp Lys Asn Leu Trp Ser Trp Phe Asp Ile Ser  
 660 665 670  
 Asn Trp Leu Trp Tyr Ile Lys Ser Asn Ile Asn Lys Leu Lys Ser Ser  
 675 680 685  
 55 Ile Glu Ser Thr Asn Glu Ala Val Val Lys Leu Gln Glu Thr Ala Glu  
 690 695 700  
 60 Lys Thr Val Tyr Val Leu Thr Ala Leu Gln  
 705 710

ES 2 733 279 T3

<210> 79  
 <211> 722  
 <212> PRT  
 5 <213> Secuencia Artificial

<220>  
 <223> Proteína de fusión VP7-HR2HXB2 del Rotavirus Rhesus producido por ingeniería

10 <400> 79  
 Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
 1 5 10 15

15 Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Glu Asp His  
 20 25 30

20 His His His His His Glu Asn Leu Tyr Phe Gln Gly Ala Glu Asn Leu  
 35 40 45

25 Trp Val Gly Asn Met Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp  
 50 55 60

30 Thr Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Thr Lys Ala Tyr  
 65 70 75 80

35 Asp Arg Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr  
 85 90 95

40 Asp Pro Asn Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe  
 100 105 110

45 Asn Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile  
 115 120 125

50 Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu  
 130 135 140

55 Cys Val Thr Leu His Cys Thr Asn Ala Thr Phe Lys Asn Asn Val Thr  
 145 150 155 160

60 Asn Asp Met Asn Lys Glu Ile Arg Asn Cys Ser Phe Asn Thr Thr Thr  
 165 170 175

65 Glu Ile Arg Asp Lys Lys Gln Gln Gly Tyr Ala Leu Phe Tyr Arg Pro  
 180 185 190

70 Asp Ile Val Leu Leu Lys Glu Asn Arg Asn Asn Ser Asn Asn Ser Glu  
 195 200 205

75 Tyr Ile Leu Ile Asn Cys Asn Ala Ser Thr Ile Thr Gln Ala Cys Pro  
 210 215 220

80 Lys Val Asn Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly  
 225 230 235 240

85 Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Ser Gly Lys Gly Pro  
 245 250 255

90 Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val

ES 2 733 279 T3

				260					265					270			
	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Lys	Glu	Ile	
			275					280					285				
5	Ile	Ile	Arg	Ser	Glu	Asn	Leu	Thr	Asp	Asn	Val	Lys	Thr	Ile	Ile	Val	
		290					295					300					
10	His	Leu	Asn	Lys	Ser	Val	Glu	Ile	Val	Cys	Thr	Arg	Pro	Asn	Asn	Asn	
	305					310					315					320	
	Thr	Arg	Lys	Ser	Met	Arg	Ile	Gly	Pro	Gly	Gln	Thr	Phe	Tyr	Ala	Thr	
					325					330					335		
15	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln	Ala	Tyr	Cys	Asn	Ile	Ser	Gly	
				340					345					350			
	Ser	Lys	Trp	Asn	Glu	Thr	Leu	Lys	Arg	Val	Lys	Glu	Lys	Leu	Gln	Glu	
20			355					360					365				
	Asn	Tyr	Asn	Asn	Asn	Lys	Thr	Ile	Lys	Phe	Ala	Pro	Ser	Ser	Gly	Gly	
		370					375					380					
25	Asp	Leu	Glu	Ile	Thr	Thr	His	Ser	Phe	Asn	Cys	Arg	Gly	Glu	Phe	Phe	
	385					390					395					400	
	Tyr	Cys	Asn	Thr	Thr	Arg	Leu	Phe	Asn	Asn	Asn	Ala	Thr	Glu	Asp	Glu	
				405					410						415		
30	Thr	Ile	Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	
				420					425					430			
	Gly	Val	Gly	Arg	Ala	Met	Tyr	Ala	Pro	Pro	Ile	Ala	Gly	Asn	Ile	Thr	
			435					440					445				
35	Cys	Lys	Ser	Asn	Ile	Thr	Gly	Leu	Leu	Leu	Val	Arg	Asp	Gly	Gly	Glu	
		450					455					460					
40	Asp	Asn	Lys	Thr	Glu	Glu	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asn	Met	Lys	
	465					470					475					480	
	Asp	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Ile	Glu	Leu	Lys	
45				485					490						495		
	Pro	Leu	Gly	Ile	Ala	Pro	Thr	Gly	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg	
				500					505					510			
50	Glu	Lys	Arg	Ala	Val	Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu	Gly	
			515					520					525				
	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Leu	Thr	Leu	Thr	Val	Gln	
			530				535					540					
55	Ala	Arg	Gln	Leu	Leu	Ser	Ser	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu	
	545					550					555					560	
	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly	
60				565						570					575		



ES 2 733 279 T3

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys  
580 585 590

5 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys  
595 600 605

Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Thr  
610 615 620

10 Asp Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile Ser  
625 630 635 640

Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Thr Gln  
645 650 655

15 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn  
660 665 670

20 Leu Trp Ser Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ser  
675 680 685

Asn Ile Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn Glu Ala Val  
690 695 700

25 Val Lys Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val Leu Thr Ala  
705 710 715 720

Leu Gln

30 <210> 80  
<211> 282  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> CR6261 ScFv (para expresión bacteriana)

<400> 80

40 Met Glu Val Gln Leu Val Glu Ser Gly Ala Glu Val Lys Lys Pro Gly  
1 5 10 15

Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Pro Phe Arg Ser  
20 25 30

45 Tyr Ala Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Pro Glu Trp  
35 40 45

Met Gly Gly Ile Ile Pro Ile Phe Gly Thr Thr Lys Tyr Ala Pro Lys  
50 55 60

Phe Gln Gly Arg Val Thr Ile Thr Ala Asp Asp Phe Ala Gly Thr Val  
65 70 75 80

55 Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Met Tyr Tyr  
85 90 95

Cys Ala Lys His Met Gly Tyr Gln Val Arg Glu Thr Met Asp Val Trp  
100 105 110

60 Gly Lys Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly

ES 2 733 279 T3

		115				120				125							
		Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gln	Ser	Val	Leu	Thr	Gln	Pro
		130				135						140					
5		Pro	Ser	Val	Ser	Ala	Ala	Pro	Gly	Gln	Lys	Val	Thr	Ile	Ser	Cys	Ser
		145				150						155					160
		Gly	Ser	Ser	Ser	Asn	Ile	Gly	Asn	Asp	Tyr	Val	Ser	Trp	Tyr	Gln	Gln
10					165						170					175	
		Leu	Pro	Gly	Thr	Ala	Pro	Lys	Leu	Leu	Ile	Tyr	Asp	Asn	Asn	Lys	Arg
					180											190	
15		Pro	Ser	Gly	Ile	Pro	Asp	Arg	Phe	Ser	Gly	Ser	Lys	Ser	Gly	Thr	Ser
				195					200						205		
		Ala	Thr	Leu	Gly	Ile	Thr	Gly	Leu	Gln	Thr	Gly	Asp	Glu	Ala	Asn	Tyr
20				210				215					220				
		Tyr	Cys	Ala	Thr	Trp	Asp	Arg	Arg	Pro	Thr	Ala	Tyr	Val	Val	Phe	Gly
		225					230					235					240
		Gly	Gly	Thr	Lys	Leu	Thr	Val	Leu	Gly	Ala	Ala	Ala	Gly	Gln	Pro	Lys
25					245						250					255	
		Ala	Ala	Pro	Ser	Val	Thr	Leu	Phe	Pro	Pro	Ser	Ser	Glu	Glu	Leu	Gln
					260					265					270		
30		Arg	Gly	Ala	Ala	His	His	His	His	His	His	His	His	His	His	His	His
				275											280		
		<210>		81													
		<211>		301													
		<212>		PRT													
35		<213>		Secuencia Artificial													
		<220>															
		<223>		CR6261 scFv (para expresión de mamífero)													
40		<400>		81													
		Met	Asp	Ala	Met	Lys	Arg	Gly	Leu	Cys	Cys	Val	Leu	Leu	Leu	Cys	Gly
		1				5					10					15	
		Ala	Val	Phe	Val	Ser	Pro	Ser	Gln	Ser	Val	Leu	Thr	Gln	Pro	Pro	Ser
45					20					25					30		
		Val	Ser	Ala	Ala	Pro	Gly	Gln	Lys	Val	Thr	Ile	Ser	Cys	Ser	Gly	Ser
				35					40					45			
50		Ser	Ser	Asn	Ile	Gly	Asn	Asp	Tyr	Val	Ser	Trp	Tyr	Gln	Gln	Leu	Pro
			50					55					60				
		Gly	Thr	Ala	Pro	Lys	Leu	Leu	Ile	Tyr	Asp	Asn	Asn	Lys	Arg	Pro	Ser
55						70						75				80	
		Gly	Ile	Pro	Asp	Arg	Phe	Ser	Gly	Ser	Lys	Ser	Gly	Thr	Ser	Ala	Thr
					85						90					95	
60		Leu	Gly	Ile	Thr	Gly	Leu	Gln	Thr	Gly	Asp	Glu	Ala	Asn	Tyr	Tyr	Cys
					100					105						110	

ES 2 733 279 T3

Ala Thr Trp Asp Arg Arg Pro Thr Ala Tyr Val Val Phe Gly Gly Gly  
 115 120 125

5 Thr Lys Leu Thr Val Leu Gly Ala Ala Ala Gly Gln Pro Lys Ala Ala  
 130 135 140

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

10 Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Glu Val Gln Leu Val Glu  
 165 170 175

15 Ser Gly Ala Glu Val Lys Lys Pro Gly Ser Ser Val Lys Val Ser Cys  
 180 185 190

Lys Ala Ser Gly Gly Pro Phe Arg Ser Tyr Ala Ile Ser Trp Val Arg  
 195 200 205

20 Gln Ala Pro Gly Gln Gly Pro Glu Trp Met Gly Gly Ile Ile Pro Ile  
 210 215 220

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

25 Thr Ala Asp Asp Phe Ala Gly Thr Val Tyr Met Glu Leu Ser Ser Leu  
 245 250 255

Arg Ser Glu Asp Thr Ala Met Tyr Tyr Cys Ala Lys His Met Gly Tyr  
 260 265 270

30 Gln Val Arg Glu Thr Met Asp Val Trp Gly Lys Gly Thr Thr Val Thr  
 275 280 285

Val Ser Ser Arg Gly Ala Ala His His His His His His  
 290 295 300

35 <210> 82  
 <211> 244  
 <212> PRT  
 <213> Secuencia Artificial

40 <220>  
 <223> cadena ligera CR6261 (Fab)

<400> 82

45 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly  
 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Ser Val Leu Thr Gln Pro Pro Ser  
 20 25 30

50 Val Ser Ala Ala Pro Gly Gln Lys Val Thr Ile Ser Cys Ser Gly Ser  
 35 40 45

Ser Ser Asn Ile Gly Asn Asp Tyr Val Ser Trp Tyr Gln Gln Leu Pro  
 50 55 60

Gly Thr Ala Pro Lys Leu Leu Ile Tyr Asp Asn Asn Lys Arg Pro Ser  
 65 70 75 80

60 Gly Ile Pro Asp Arg Phe Ser Gly Ser Lys Ser Gly Thr Ser Ala Thr

ES 2 733 279 T3

				85					90					95			
	Leu	Gly	Ile	Thr	Gly	Leu	Gln	Thr	Gly	Asp	Glu	Ala	Asn	Tyr	Tyr	Cys	
				100					105					110			
5	Ala	Thr	Trp	Asp	Arg	Arg	Pro	Thr	Ala	Tyr	Val	Val	Phe	Gly	Gly	Gly	
			115					120					125				
10	Thr	Lys	Leu	Thr	Val	Leu	Gly	Ala	Ala	Ala	Gly	Gln	Pro	Lys	Ala	Ala	
		130					135					140					
15	Pro	Ser	Val	Thr	Leu	Phe	Pro	Pro	Ser	Ser	Glu	Glu	Leu	Gln	Ala	Asn	
	145					150					155					160	
20	Lys	Ala	Thr	Leu	Val	Cys	Leu	Ile	Ser	Asp	Phe	Tyr	Pro	Gly	Ala	Val	
				165						170					175		
25	Thr	Val	Ala	Trp	Lys	Ala	Asp	Ser	Ser	Pro	Val	Lys	Ala	Gly	Val	Glu	
				180					185					190			
30	Thr	Thr	Thr	Pro	Ser	Lys	Gln	Ser	Asn	Asn	Lys	Tyr	Ala	Ala	Ser	Ser	
			195					200					205				
35	Tyr	Leu	Ser	Leu	Thr	Pro	Glu	Gln	Trp	Lys	Ser	His	Arg	Ser	Tyr	Ser	
		210					215					220					
40	Cys	Gln	Val	Thr	His	Glu	Gly	Ser	Thr	Val	Glu	Lys	Thr	Val	Ala	Pro	
	225					230					235					240	
45	Thr	Glu	Cys	Ser													
50	<210>	83															
	<211>	249															
	<212>	PRT															
55	<213>	Secuencia Artificial															
	<220>																
	<223>	cadena pesada CR6261 (Fab)															
60	<400>	83															
	Met	Asp	Ala	Met	Lys	Arg	Gly	Leu	Cys	Cys	Val	Leu	Leu	Leu	Cys	Gly	
	1				5					10					15		
65	Ala	Val	Phe	Val	Ser	Pro	Ser	Glu	Val	Gln	Leu	Val	Glu	Ser	Gly	Ala	
			20					25					30				
70	Glu	Val	Lys	Lys	Pro	Gly	Ser	Ser	Val	Lys	Val	Ser	Cys	Lys	Ala	Ser	
			35				40						45				
75	Gly	Gly	Pro	Phe	Arg	Ser	Tyr	Ala	Ile	Ser	Trp	Val	Arg	Gln	Ala	Pro	
		50					55					60					
80	Gly	Gln	Gly	Pro	Glu	Trp	Met	Gly	Gly	Ile	Ile	Pro	Ile	Phe	Gly	Thr	
	65					70					75					80	
85	Thr	Lys	Tyr	Ala	Pro	Lys	Phe	Gln	Gly	Arg	Val	Thr	Ile	Thr	Ala	Asp	
				85						90					95		
90	Asp	Phe	Ala	Gly	Thr	Val	Tyr	Met	Glu	Leu	Ser	Ser	Leu	Arg	Ser	Glu	
			100						105						110		

ES 2 733 279 T3

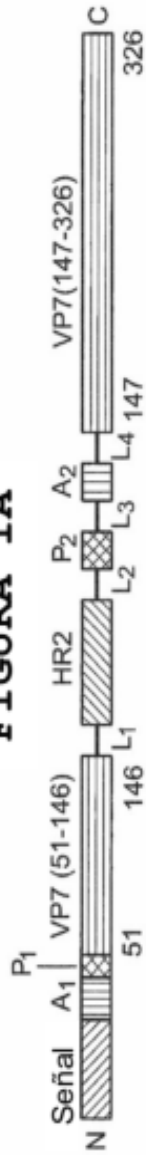
	Asp	Thr	Ala	Met	Tyr	Tyr	Cys	Ala	Lys	His	Met	Gly	Tyr	Gln	Val	Arg
			115					120					125			
5	Glu	Thr	Met	Asp	Val	Trp	Gly	Lys	Gly	Thr	Thr	Val	Thr	Val	Ser	Ser
			130				135					140				
	Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys
10						150					155					160
	Ser	Thr	Ser	Gly	Gly	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr
					165					170					175	
	Phe	Pro	Glu	Pro	Val	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser
15				180					185					190		
	Gly	Val	His	Thr	Phe	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser
20			195					200					205			
	Leu	Ser	Ser	Val	Val	Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr
			210				215					220				
	Tyr	Ile	Cys	Asn	Val	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys
25	225					230					235					240
	Arg	Val	Glu	Pro	Lys	Ser	Cys	Asp	Lys							
					245											
30																

## REIVINDICACIONES

- 5 1. Un complejo proteico quimérico que comprende una proteína de superficie VP7 de rotavirus formadora de trímeros unida a una proteína heteróloga, en donde la proteína de superficie VP7 de rotavirus está unida a la proteína heteróloga de manera no covalente mediante un sistema adaptador de dos partes, donde una parte del sistema adaptador es formado por un primer polipéptido adaptador que se fusiona con la proteína de superficie VP7 del rotavirus, opcionalmente a través de una secuencia enlazadora, y la otra parte del adaptador está formado por un segundo polipéptido adaptador que se fusiona con la proteína heteróloga, opcionalmente a través de una secuencia enlazadora, por lo que ambas partes del sistema adaptador forman un complejo estable entre sí y el complejo proteico quimérico puede convertirse en parte de la capa externa de una partícula de rotavirus mediante la recubierta in vitro de partículas de rotavirus de doble capa.
- 10 2. El complejo proteico quimérico de la reivindicación 1, en el que el primer polipéptido adaptador y el segundo polipéptido adaptador comprenden una secuencia de repetición de heptada.
- 15 3. Una partícula de rotavirus que comprende el complejo proteico quimérico de la reivindicación 1 o la reivindicación 2.
- 20 4. La partícula de rotavirus de la reivindicación 3 para uso como medicamento.
5. Una composición de ácido nucleico que comprende:
- 25 (a) un marco de lectura abierto que codifica una proteína de superficie de rotavirus modificada que comprende una proteína de superficie de rotavirus VP7, un primer polipéptido adaptador y, opcionalmente, una secuencia enlazadora; opcionalmente, en el que el marco de lectura abierto está vinculado operativamente a una secuencia promotora; y
- 30 (b) un marco de lectura abierto que codifica una proteína de fusión que comprende una proteína heteróloga, un segundo polipéptido adaptador y, opcionalmente, una secuencia enlazadora; opcionalmente, en el que el marco de lectura abierto está vinculado operativamente a una secuencia promotora; en donde los polipéptidos adaptadores primero y segundo son capaces de formar un complejo estable entre sí.
- 35 6. La composición de ácido nucleico de la reivindicación 5, en la que el polipéptido adaptador comprende una secuencia de repetición de heptada
7. Un kit que comprende:
- 40 (a) (i) un primer ácido nucleico que codifica una proteína de superficie de rotavirus modificada que comprende una proteína de superficie VP7 de rotavirus y un primer polipéptido adaptador, y (ii) un segundo ácido nucleico que comprende una secuencia de nucleótidos que codifica un segundo polipéptido adaptador y un sitio de clonación múltiple, y en la que la inserción de una región codificante para una proteína heteróloga en el sitio de clonación múltiple produce un marco de lectura abierto que codifica una proteína de fusión que comprende la proteína heteróloga y el segundo polipéptido adaptador; o
- 45 (b) (i) un primer ácido nucleico que codifica una proteína de superficie de rotavirus modificada que comprende una proteína de superficie VP7 de rotavirus y un primer polipéptido adaptador y (ii) un segundo ácido nucleico que codifica una proteína de fusión que comprende una proteína heteróloga y un segundo polipéptido adaptador; en el que el primer polipéptido adaptador y el segundo polipéptido adaptador son capaces de formar un complejo estable
- 50 8. Un kit según la reivindicación 7, en el que el kit comprende además una partícula de rotavirus, en la que la partícula se deriva de la misma especie de rotavirus que la proteína de superficie del rotavirus o de una especie diferente de rotavirus
- 55 9. Un método para preparar la partícula de rotavirus de la reivindicación 3, en el que el método comprende propagar una partícula de rotavirus que comprende una capa externa en una célula que crece en un medio de cultivo, purificando la partícula de rotavirus del medio de cultivo, eliminando la capa externa del rotavirus partícula para obtener una partícula de doble capa de rotavirus (DLP), y recubrir la DLP del rotavirus con el complejo de proteínas quiméricas de la reivindicación 1 o la reivindicación 2 para producir la partícula de rotavirus de la reivindicación 3.
- 60 10. Un método para preparar la partícula de rotavirus de la reivindicación 3, en el que el método comprende propagar una partícula de rotavirus que comprende una capa externa en una célula que crece en un medio de cultivo, purificando la partícula de rotavirus del medio de cultivo, eliminando la capa externa del rotavirus partícula para obtener un DLP de rotavirus, y volver a recubrir el DLP de rotavirus con una primera proteína de fusión que
- 65

- 5 comprende una proteína de superficie VP7 de rotavirus formadora de trímeros, un primer polipéptido adaptador que comprende una secuencia de repetición de heptada, y opcionalmente una secuencia enlazadora y mezclar el DLP de rotavirus recubierto con una la segunda proteína de fusión que comprende una proteína heteróloga formadora de trímeros, un segundo polipéptido adaptador que comprende una secuencia de repetición de heptada y, opcionalmente, una secuencia enlazadora para producir la partícula de rotavirus de la reivindicación 3.
- 10 11. Un método para preparar la partícula de rotavirus de la reivindicación 3, que comprende mezclar una partícula de rotavirus que comprende una primera proteína de fusión que comprende una proteína de superficie VP7 de rotavirus formadora de trímeros, un primer polipéptido adaptador que comprende una secuencia de repetición de heptada y, opcionalmente, una secuencia enlazadora con una segunda proteína de fusión que comprende una proteína heteróloga formadora de trímeros, un segundo polipéptido adaptador que comprende una secuencia de repetición de heptada y, opcionalmente, una secuencia enlazadora.
- 15 12. Un método para determinar la estructura de una proteína heteróloga, en donde el método comprende los pasos de (i) recubrir un DLP de rotavirus con un complejo proteico quimérico de acuerdo con la reivindicación 1 o la reivindicación 2 para producir una suspensión de partículas de rotavirus que muestran la proteína de superficie quimérica, (ii) congelación de la suspensión, (iii) imágenes del rotavirus  
20 partículas que utilizan crio-EM para obtener una pluralidad de micrografías, y (iv) analizar la pluralidad de micrografías para obtener un modelo tridimensional del complejo proteico quimérico.
- 25 13. Un método para determinar la estructura de una proteína heteróloga en complejo con una molécula, en donde el método comprende las etapas de (i) recubrir una DLP de rotavirus con un complejo de proteína quimérica de acuerdo con la reivindicación 1 o la reivindicación 2 para producir una suspensión de partículas de rotavirus mostrando el complejo proteico quimérico, (ii) agregando a la suspensión una molécula que se une específicamente a la proteína heteróloga, en donde la molécula forma un complejo con el complejo proteico quimérico, (iii) congelando la suspensión, (iv) formando imágenes de las partículas del rotavirus usando crio -EM para obtener una pluralidad de micrografías, y (vi) analizar la pluralidad de micrografías para obtener un modelo tridimensional de la proteína quimérica complejada con la molécula.
- 30 14. El método de la reivindicación 13, en el que la molécula es (a) un anticuerpo o fragmento del mismo, en el que el anticuerpo o fragmento se une específicamente a la proteína heteróloga; o (b) un receptor de superficie celular, en el que la proteína heteróloga es una proteína de entrada celular viral y la molécula proteica está unida por la proteína de entrada celular viral.
- 35

**FIGURA 1A**

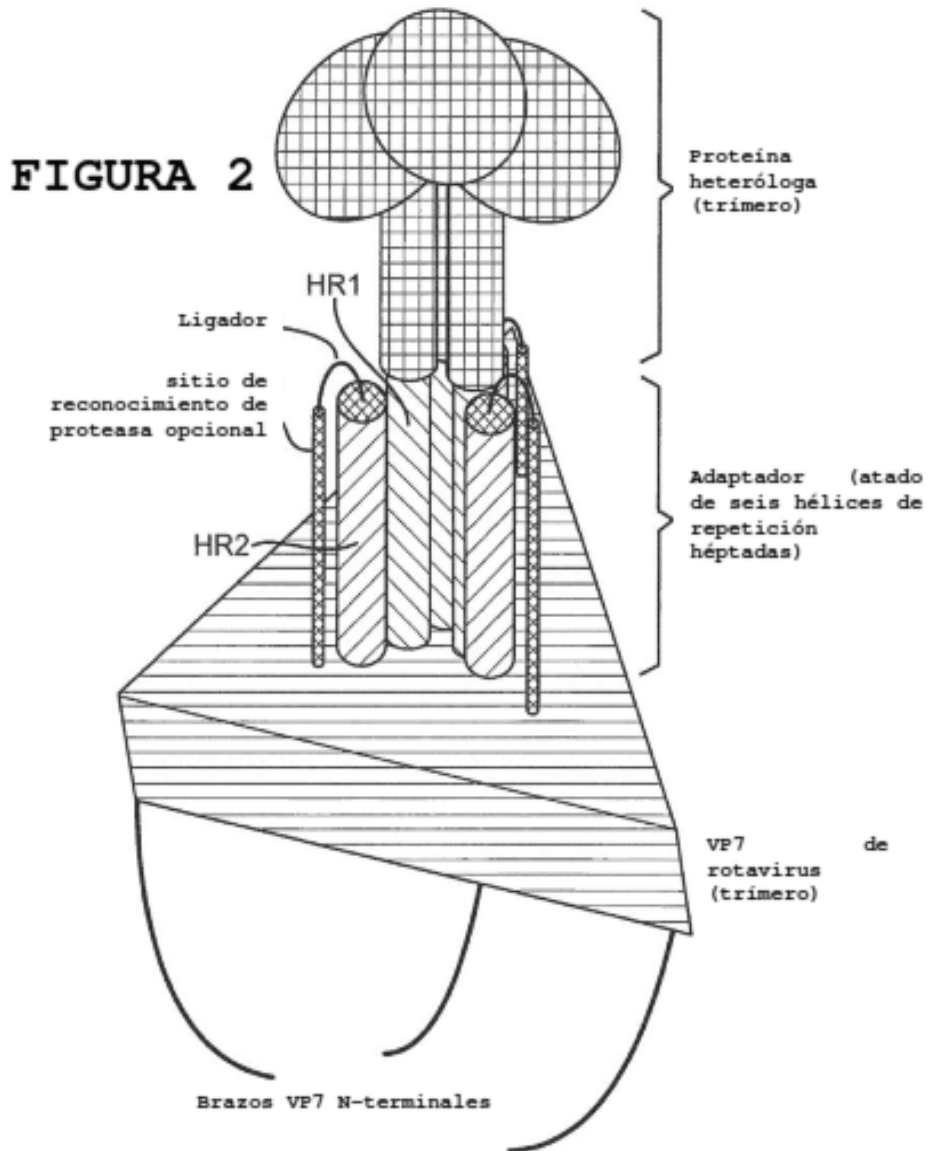


**FIGURA 1B**

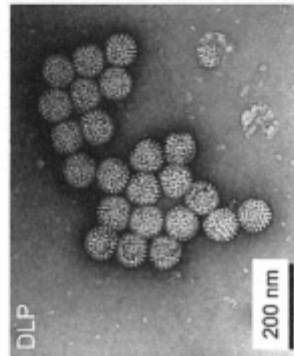


Señal=secuencia d péptidos señal heteróloga  
 A=etiqueta de afinidad (opcional)  
 P=sitio de reconocimiento de proteasa (opcional)  
 L=ligador  
 T=etiqueta de trimerización (opcional)  
 HR1, HR2=secuencias de repetición héptadas

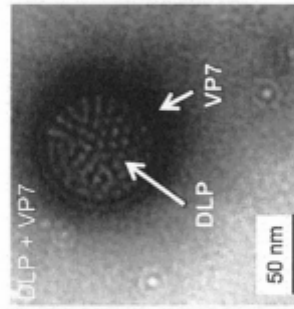




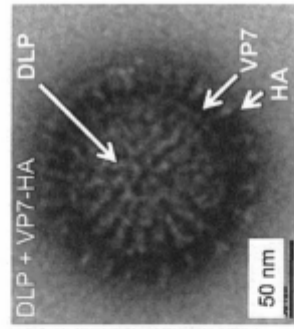
**FIGURA 3A**



**FIGURA 3B**



**FIGURA 3C**



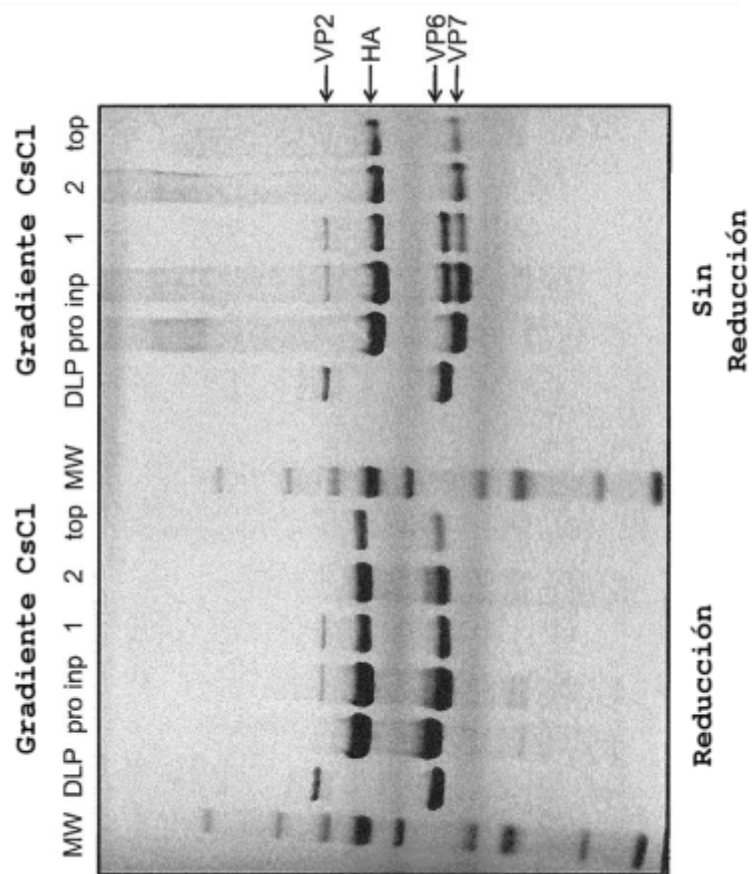
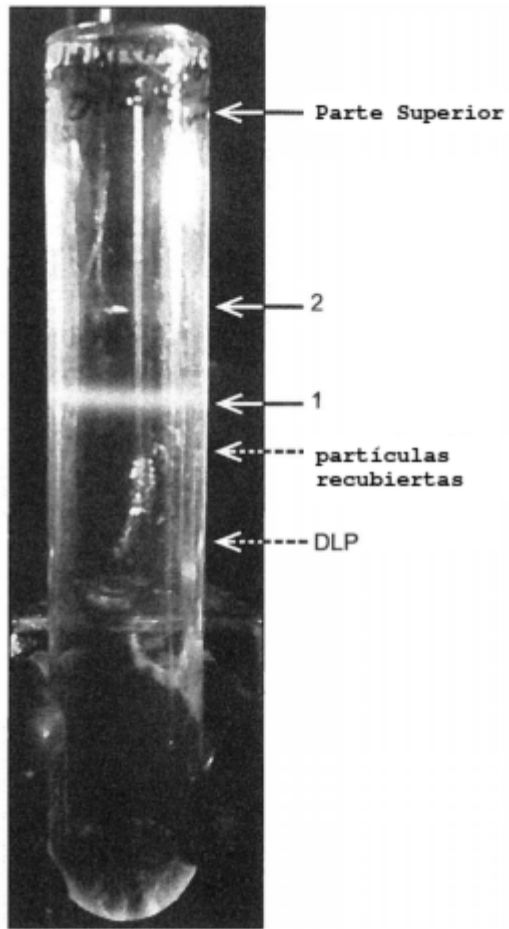
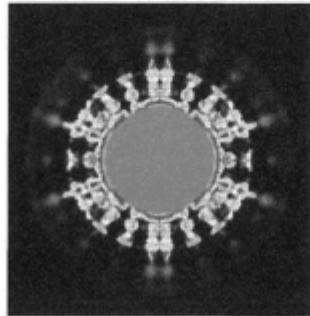


FIGURA 4A



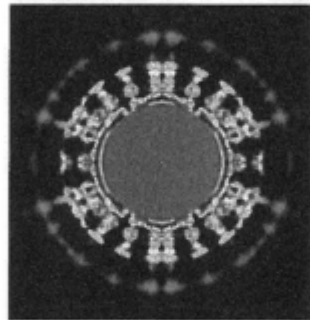
**FIGURA 4B**

**FIGURA 5A**



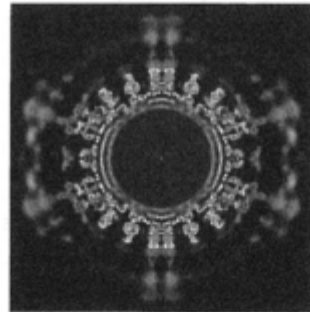
DLP + VP7-HA

**FIGURA 5B**

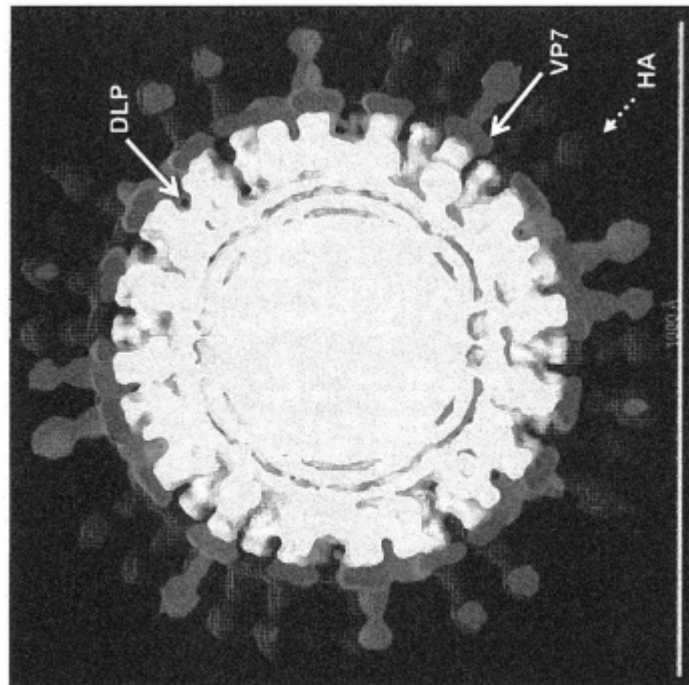


DLP + VP7-HA-  
CR6261ScFv

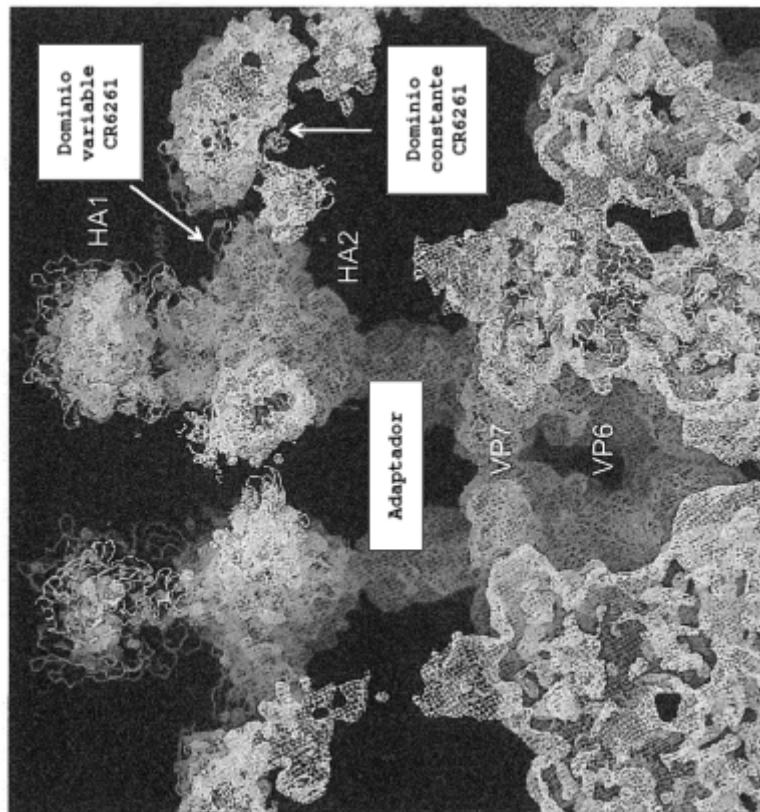
**FIGURA 5C**



DLP + VP7-HA-  
CR6261Fab



**FIGURA 6**



**FIGURA 7**