(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 614 266**

(51) Int. CI.:

C07K 14/35 (2006.01)
A61K 39/04 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **27.01.2011 PCT/EP2011/051158**

(87) Fecha y número de publicación internacional: **04.08.2011 WO11092253**

(96) Fecha de presentación y número de la solicitud europea: **27.01.2011 E 11700859 (9)**

(97) Fecha y número de publicación de la concesión europea: **21.09.2016 EP 2528621**

(54) Título: **Antígenos de tuberculosis modificados**

(30) Prioridad:

27.01.2010 US 298710 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
30.05.2017

(73) Titular/es:

GLAXOSMITHKLINE BIOLOGICALS S.A. (50.0%)
Rue de l'Institut, 89
1330 Rixensart, BE y
GLAXO GROUP LIMITED (50.0%)

(72) Inventor/es:

BLAIS, NORMAND;
BROWN, JAMES;
GELINAS, ANNE-MARIE;
METTENS, PASCAL y
MURPHY, DENNIS

(74) Agente/Representante:

CARPINTERO LÓPEZ, Mario

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

DESCRIPCIÓN

Antígenos de tuberculosis modificados

Campo de la invención

La presente invención se refiere a proteínas Rv3616c de *Mycobacterium tuberculosis* modificadas, a los polinucleótidos asociados y al uso de tales proteínas y polinucleótidos en el tratamiento o en la prevención de la tuberculosis, en particular al uso en el tratamiento o en la prevención de tuberculosis latente y en la prevención o el retraso de la reactivación de la tuberculosis.

Antecedentes de la invención

La tuberculosis (TB) es una enfermedad infecciosa crónica causada por una infección con *Mycobacterium tuberculosis* y con otras especies de *Mycobacterium*. Es una enfermedad importante en los países en desarrollo, así como un problema creciente en las áreas desarrolladas del mundo. Se cree que más de 2 mil millones de personas están infectadas con bacilos de tuberculosis (TB), con aproximadamente 9,2 millones de nuevos casos de tuberculosis (TB) y 1,7 millones de muertes cada año. El 10 % de aquéllos infectados con bacilos de tuberculosis (TB) desarrollarán la tuberculosis (TB) activa, en donde cada persona con la tuberculosis (TB) activa infectará a un promedio de otras 10 a 15 personas al año. Aunque los índices de incidencia anuales han llegado a su pico globalmente, el número de muertes y casos todavía se está elevando debido al crecimiento de la población (World Health Organisation [Organización Mundial de la Salud] Tuberculosis Facts, 2008).

Mycobacterium tuberculosis infecta a los individuos a través de las vías respiratorias. Los macrófagos alveolares rodean completamente a la bacteria, pero ésta puede sobrevivir y proliferar mediante la inhibición de la fusión del fagosoma con los lisosomas ácidos. Sigue una respuesta inmunitaria compleja que involucra a las células T CD4+ y CD8+, dando como resultado por último la formación de un granuloma. Central para el éxito de la *Mycobacterium tuberculosis* como un patógeno, es el hecho de que la bacteria aislada, pero no erradicada, puede persistir durante largos períodos, dejando a un individuo vulnerable al desarrollo posterior de la tuberculosis (TB) activa.

Menos del 5 % de los individuos infectados desarrollan la tuberculosis (TB) activa en los primeros años después de la infección. El granuloma puede persistir durante décadas, y se cree que contiene la *Mycobacterium tuberculosis* viva en un estado de latencia, privada de oxígeno y nutrientes. Sin embargo, recientemente se ha sugerido que la mayoría de las bacterias en el estado latente se localizan en los tipos de células que no son macrófagos, extendidas a través de todo el cuerpo (Locht y col., Expert Opin. Biol. Ther. 2007 7(11) : 1665-1677). El desarrollo de la tuberculosis (TB) activa se presenta cuando cambia el equilibrio entre la inmunidad natural del huésped y el patógeno, por ejemplo, como un resultado de un evento inmunosupresor (Anderson P, Trends in Microbiology, 2007 15(1) : 7-13; Ehlers S Infection 2009 37(2) : 87-95).

También se ha propuesto una hipótesis dinámica que describe el equilibrio entre la tuberculosis (TB) latente y la tuberculosis (TB) activa (Cardana P-J, Inflammation & Allergy – Drug Targets, 2006, 6: 27-39; Cardana P-J, Infection, 2009 37(2) : 80-86).

Aunque una infección puede ser asintomática durante un período de tiempo considerable, la enfermedad activa se manifiesta más comúnmente como una inflamación aguda de los pulmones, que da como resultado cansancio, pérdida de peso, fiebre, y una tos persistente. Si no se trata, resultan complicaciones serias no tratadas, y típicamente la muerte.

En términos generales, la tuberculosis se puede controlar utilizando una terapia prolongada con antibióticos, aunque este tratamiento no es suficiente para prevenir la extensión de la enfermedad. Los individuos activamente infectados pueden ser en gran parte asintomáticos, pero contagiosos, durante algún tiempo. En adición, aunque es crítico el cumplimiento con el régimen de tratamiento, es difícil monitorear el comportamiento del paciente. Algunos pacientes no completan el curso del tratamiento, lo cual puede conducir a un tratamiento inefectivo y al desarrollo de resistencia al fármaco.

La tuberculosis (TB) resistente a múltiples fármacos (MDR-TB) es una forma que fracasa para responder a los medicamentos de primera línea. El 5 % de todos los casos de tuberculosis (TB) son de tuberculosis (TB) resistente a múltiples fármacos (MDR-TB), presentándose una estimación de 490,000 nuevos casos de tuberculosis (TB) resistente a múltiples fármacos (MDR-TB) cada año. La tuberculosis (TB) extensamente resistente a fármacos (XDR-TB) se presenta cuando se desarrolla resistencia a los medicamentos de segunda línea, en la parte superior de la tuberculosis (TB) resistente a múltiples fármacos (MDR-TB). Se estima que se presentan 40,000 nuevos casos de la tuberculosis (TB) extensamente resistente a fármacos (XDR-TB) virtualmente no tratable anualmente (World Health Organisation [Organización Mundial de la Salud], Tuberculosis Facts, 2008).

Incluso cuando se complete un curso completo de tratamiento con antibióticos, puede no erradicarse la infección con *M. tuberculosis* del individuo infectado, y puede permanecer como una infección latente que se puede reactivar.

Con el objeto de controlar la extensión de la tuberculosis, es de suma importancia un programa de vacunación

efectivo y un diagnóstico oportuno preciso de la enfermedad.

Actualmente, la vacunación con las bacterias vivas es el método más ampliamente empleado para inducir la inmunidad protectora. La *Mycobacterium* más habitual empleada para este propósito es el bacilo de Calmette-Guerin (BCG), una cepa avirulenta de M. bovis que fue primeramente desarrollada hace más de 60 años. Sin embargo, la seguridad y eficacia de el BCG es una fuente de controversia - aunque protege contra la manifestación grave de la enfermedad en los niños, el BCG no previene el establecimiento de la tuberculosis (TB) latente o la reactivación de la enfermedad pulmonar en la vida adulta. Adicionalmente, algunos países, tales como los Estados Unidos, no vacunan al público en general con este agente.

Casi todas las vacunas para tuberculosis (TB) de nueva generación que están actualmente en desarrollo clínico, se han diseñado como vacunas previas a la exposición. Éstas incluyen vacunas subunitarias, las cuales han sido particularmente efectivas para reforzar la inmunidad inducida por la vacunación previa con BCG, y las vacunas micobacterianas vivas avanzadas que tienen como objetivo reemplazar a el BCG con cepas más eficientes y/o más seguras. Aunque estas vacunas tienen como objetivo mejorar la resistencia a la infección, tienen probabilidades de ser menos efectivas como vacunas posteriores a la exposición o como vacunas terapéuticas en los casos de tuberculosis (TB) latente (Lin MY y col., Endocrine, Metabolic & Immune Disorders – Drug Targets, 2008, 8: 15-29).

El Ejemplo 2 del documento US20080269151 da a conocer la clonación, construcción y expresión de ciertas proteínas Rv3616c modificadas, incluyendo: ΔTM-1, un polipéptido de Rv3616c, en donde se han suprimido los residuos 150 a 160 (SEQ ID No: 22 de la Patente de los Estados Unidos de Norteamérica Número US20080269151) ; ΔTM-2, un polipéptido de Rv3616c, en donde se han suprimido los residuos 101 a 203 (SEQ ID No: 24 del documento US20080269151) ; y una secuencia en donde los residuos 150 a 160 de Rv3616c han sido reemplazados por el antígeno TbH9 (SEQ ID No: 60 del documento US20080269151).

Breve sumario de la invención

La presente invención se refiere, en términos generales, al uso de polipéptidos de Rv3616c modificados, o de los polinucleótidos que los codifican, en el campo de las infecciones micobacterianas latentes. Adicionalmente, la presente invención se refiere a proteínas Rv3616c modificadas particulares. Los inventores han descubierto, de una manera sorprendente, que la alteración de la hidrofobicidad de una región particular de una secuencia de Rv3616c puede conducir a una mejor expresión sin un impacto perjudicial a las propiedades inmunogénicas. Las proteínas Rv3616c modificadas son útiles como antígenos de tuberculosis (TB), en particular como antígenos de tuberculosis (TB) latente.

En su aspecto más amplio, la presente invención proporciona una proteína Rv3616c modificada, en donde se ha alterado la hidrofobicidad de los residuos de aminoácidos correspondientes a los residuos 134-183 de la secuencia de H37Rv, de una manera adecuada una proteína Rv3616c modificada, en donde se ha alterado la hidrofobicidad de los residuos de aminoácidos correspondientes a los residuos 135-154 de la secuencia de H37Rv.

En un aspecto de la invención, se proporciona una proteína Rv3616c modificada, comprendiendo dicha proteína Rv3616c modificada, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo C de la proteína Rv3616c modificada en relación con el segundo polipéptido, y en la que:

(i) el primer polipéptido es una secuencia que tiene al menos un 90 % de identidad con los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos un 90 % de identidad con los residuos 184-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están unidos directa o indirectamente.

En un segundo aspecto de la invención, se proporciona una proteína Rv3616c modificada, comprendiendo esta proteína Rv3616c modificada, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo C de la proteína Rv3616c modificada en relación con el segundo polipéptido, y en la que:

(i) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia contigua de al menos 155 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están unidos directa o indirectamente.

En un tercer aspecto de la invención, se proporciona una proteína Rv3616c modificada, comprendiendo esta proteína o, de una manera alternativa, consistiendo esencialmente o consistiendo en, una secuencia de Rv3616c, en donde al menos un aminoácido (por ejemplo, al menos 2) ha sido suprimido de la región correspondiente a los residuos 134-183 de la SEQ ID NO: 1.

Un cuarto aspecto de la invención proporciona una proteína Rv3616c modificada, comprendiendo esta proteína un

primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en la que

(i) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia contigua de al menos 155 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están unidos directa o indirectamente por medio de un tercer polipéptido, correspondiendo dicho tercer polipéptido a los residuos 134-183 de la SEQ ID NO: 1, en la que se ha suprimido al menos 1 aminoácido (por ejemplo, al menos 2).

Un quinto aspecto de la invención proporciona proteínas Rv3616c modificadas, las cuales comprenden un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en en la que:

(i) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-134 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia contigua de al menos 175 aminoácidos dentro de los residuos 155-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están ya sea directamente unidos o bien indirectamente unidos por medio de un tercer polipéptido, en donde este tercer polipéptido corresponde a los residuos 135-154 de la SEQ ID NO: 1, en en la que se ha suprimido al menos 1 aminoácido (por ejemplo, al menos 2).

Un sexto aspecto de la invención proporciona una proteína Rv3616c modificada, comprendiendo dicha proteína un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en la que:

(i) el primer polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 184-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están directamente unidos o indirectamente unidos por medio de un tercer polipéptido, teniendo este tercer polipéptido al menos el 90 % de identidad con una secuencia correspondiente a los residuos 134-183 de la SEQ ID NO: 1, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4).

Un séptimo aspecto de la invención proporciona proteínas Rv3616c modificadas, que comprenden un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en la que:

(i) el primer polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 1-134 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 155-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están ya sea directamente unidos o bien indirectamente unidos por medio de un tercer polipéptido, teniendo este tercer polipéptido al menos el 80 % de identidad con una secuencia correspondiente a los residuos 135-154 de la SEQ ID NO: 1, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4).

En un octavo aspecto de la invención, se proporciona una proteína Rv3616c modificada, comprendiendo dicha proteína una secuencia de Rv3616c, en donde una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4) a partir de la región correspondiente a los residuos 134-183 de la SEQ ID NO: 1 ha sido sustituida con residuos hidrofílicos.

Las proteínas Rv3616c modificadas se pueden basar en una secuencia de proteína Rv3616c de tipo silvestre a partir de cualquier cepa de *M. tuberculosis*. Por ejemplo, una cualquiera de las SEQ ID NO: 3-7, en particular una cualquiera de las SEQ ID NO: 3-6, puede sustituir a la SEQ ID NO: 1 en las realizaciones anteriores.

Las proteínas Rv3616c modificadas de ejemplo de acuerdo con la presente invención, son aquéllas que comprenden las secuencias de aminoácidos proporcionadas en las SEQ ID NOs: 161-169, 179 y 180 (tales como aquéllas que consisten en las secuencias de aminoácidos proporcionadas en las SEQ ID NOs: 161-169, 179 y 180). Son de un interés particular aquéllas que comprenden las secuencias de aminoácidos proporcionadas en las SEQ ID NOs: 161, 163-169, 179 y 180 (tales como aquéllas que consisten en las secuencias de aminoácidos proporcionadas en las SEQ ID NOs: 161, 163-169, 179 y 180).

También se proporcionan estas proteínas Rv3616c modificadas para su uso como medicamentos.

Un aspecto adicional de la invención se refiere a un método para inducir una respuesta inmunitaria en un sujeto, que comprende la administración de una proteína Rv3616c modificada.

Un aspecto adicional de la invención se refiere a un método para el tratamiento, la mejora, o la prevención de tuberculosis (TB), que comprende la administración de una cantidad efectiva de una proteína Rv3616c modificada a un sujeto que lo necesite, en donde este polipéptido induce una respuesta inmunitaria. En un aspecto adicional, el método comprende además inducir una respuesta inmunitaria contra *Mycobacterium tuberculosis*.

El uso de una proteína Rv3616c modificada en la elaboración de un medicamento para el tratamiento, la mejora, o la prevención de tuberculosis (TB), representa otro aspecto de la invención.

La presente invención proporciona un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada. Los polinucleótidos de ejemplo que comprenden una secuencia de ácido nucleico que codifica proteínas Rv3616c modificadas, son aquéllos que comprenden las secuencias de nucleótidos proporcionadas en las SEQ ID NOs: 170-178, tales como aquéllos que consisten en las secuencias de nucleótidos proporcionadas en las SEQ ID NOs: 170-178. Otros polinucleótidos de ejemplo que comprenden una secuencia de ácido nucleico que codifica proteínas Rv3616c modificadas, son aquéllos que comprenden (por ejemplo, que consisten en) una secuencia de nucleótidos que codifica una secuencia de aminoácidos proporcionada en las SEQ ID NOs: 161-169, 179 o 180, tal como en las SEQ ID NOs: 161, 163-169, 179 o 180.

También se proporciona un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada para utilizarse como un medicamento.

Un aspecto adicional de la invención se refiere a un método para inducir una respuesta inmunitaria en un sujeto, que comprende la administración de un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada.

Un aspecto adicional de la invención se refiere a un método para el tratamiento, la mejora, el retardo, o la prevención de la reactivación de tuberculosis, que comprende la administración de una cantidad efectiva de un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada a un sujeto que lo necesite, en donde este polipéptido induce una respuesta inmunitaria. En un aspecto adicional, el método comprende además inducir una respuesta inmunitaria contra *Mycobacterium tuberculosis*.

El uso de un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una proteína Rv3616c modificada en la elaboración de un medicamento para el tratamiento, la mejora, o la prevención de tuberculosis (TB), representa otro aspecto de la invención.

Adicionalmente, se proporciona una composición farmacéutica, que comprende:

(a) una proteína Rv3616c modificada; o
(b) un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada;
y
(c) un vehículo o excipiente farmacéuticamente aceptable.

Además, se proporciona una composición inmunogénica, que comprende:

(a) una proteína Rv3616c modificada; o
(b) un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada;
y
(c) un potenciador no específico de la respuesta inmunitaria.

También se proporciona un vector de expresión, que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada.

Las células huésped, transformadas con este vector de expresión, forman un aspecto adicional de la invención. Adicionalmente, se proporciona una célula huésped que expresa de una manera recombinante una proteína Rv3616c modificada.

Además, se proporciona un método para la producción de una proteína Rv3616c modificada; comprendiendo este método la etapa de expresar de una manera recombinante el polipéptido dentro de una célula huésped.

También se proporcionan kits de diagnóstico, que comprenden:

(a) una proteína Rv3616c modificada;
(b) un aparato suficiente para poner en contacto la proteína Rv3616c modificada con una muestra (por ejemplo,

sangre entera, o de una manera más adecuada, células mononucleares de sangre periférica (PBMC)) a partir de un individuo; y
(c) medios para cuantificar la respuesta de células T de la muestra.

Otro aspecto de la invención se refiere a un kit de diagnóstico, que comprende:

(a) una proteína Rv3616c modificada; y
(b) un aparato suficiente para poner en contacto la proteína Rv3616c modificada con las células dérmicas de un paciente.

Un aspecto adicional de la invención se refiere a un método para detectar la infección por *Mycobacterium tuberculosis* en un sujeto, que comprende:

(a) poner en contacto una muestra a partir del sujeto mencionado, con una proteína Rv3616c modificada; y
(b) detectar, en la muestra biológica, la presencia de anticuerpos que se enlacen a la proteína Rv3616c modificada.

La invención también proporciona un kit de diagnóstico, que comprende:

(a) una proteína Rv3616c modificada, cuya proteína se inmoviliza opcionalmente sobre un soporte sólido; y
(b) un reactivo de detección.

En una realización, el sujeto que recibe una proteína Rv3616c modificada, un polinucleótido o una composición de acuerdo con la invención, puede tener tuberculosis activa (por ejemplo, la infección activa por *M. tuberculosis*). En una segunda realización, el sujeto puede tener tuberculosis latente (por ejemplo, la infección latente por *M. tuberculosis*). En una tercera realización, el sujeto puede estar libre de tuberculosis (por ejemplo, libre de la infección por *M. tuberculosis*).

Un sujeto que recibe una proteína Rv3616c modificada, un polinucleótido o una composición de acuerdo con la invención, puede haber sido vacunado previamente para tuberculosis (por ejemplo, puede haber sido vacunado contra la infección por *M. tuberculosis*), tal como puede haber sido vacunado con un Bacillus Calmette-Guerin (BCG). De una manera alternativa, un sujeto que recibe un polipéptido, polinucleótido o una composición de la invención, puede no haber sido vacunado previamente para tuberculosis (por ejemplo, puede no haber sido vacunado contra la infección por *M. tuberculosis*), tal como puede no haber sido vacunado con un Bacillus Calmette-Guerin (BCG).

Una proteína Rv3616c modificada, un polinucleótido o una composición de acuerdo con la invención, se puede proporcionar para el propósito de:

- tratar la tuberculosis activa;
- prevenir la tuberculosis activa (tal como mediante su administración a un sujeto que no esté infectado, o de una manera alternativa, a un sujeto que tenga una infección latente) ;
- tratar la tuberculosis latente;
- prevenir la tuberculosis latente; o
- prevenir o retardar la reactivación de tuberculosis (en especial la demora de la reactivación de tuberculosis (TB), por ejemplo, por un período de meses, años, o incluso indefinidamente).

También se proporciona un método para el tratamiento de tuberculosis (TB) latente, que comprende las etapas de:

(i) identificar en un sujeto que tiene una infección por tuberculosis (TB) latente (por ejemplo, mediante ensayos basados en PPD o en células T) ; y
(ii) administrar a este sujeto, una cantidad segura y efectiva de una proteína Rv3616c modificada, o de un polinucleótido que codifique una proteína Rv3616c modificada (tal como en la forma de una composición farmacéutica o de una composición inmunogénica).

También se proporciona el uso de un polipéptido de la presente invención, en la elaboración de un kit de diagnóstico para la identificación de tuberculosis (TB) (por ejemplo, tuberculosis (TB) latente) en un sujeto de prueba.

Description de las figuras

Figura 1: Alineación del péptido de Rv3616c con la secuencia de longitud completa.
Figura 2: Respuestas de las células mononucleares de sangre periférica (PBMC) a los péptidos de Rv3616c.
Figura 3: Porcentaje de células CD4 y CD8 a partir de ratones CB6F1 inmunizados que expresan citocinas IFN-gamma y/o IL-2 y/o TNF-alfa en el día 21 (es decir, 7 días después de la segunda inmunización).
Figura 4: Perfil de citocinas en el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD4 específica del antígeno en los ratones CB6F1 inmunizados.
Figura 5: Perfil de citocinas en el día 21 (es decir, 7 días después de la segunda inmunización) de la

respuesta de CD8 específica del antígeno en los ratones CB6F1 inmunizados.

Figura 6: Porcentaje de células CD4 y CD8 a partir de ratones CB6F1 inmunizados que expresan citocinas IFN-gamma y/o IL-2 y/o TNF-alfa en el día 35 (es decir, 7 días después de la tercera inmunización).

Figura 7: Perfil de citocinas en el día 35 (es decir, 7 días después de la tercera inmunización) de la respuesta de CD4 específica del antígeno en los ratones CB6F1 inmunizados.

Figura 8: Perfil de citocinas en el día 35 (es decir, 7 días después de la tercera inmunización) de la respuesta de CD8 específica del antígeno en los ratones CB6F1 inmunizados.

Figura 9: Porcentaje de células CD4 y CD8 a partir de los ratones C57BL/6 inmunizados que expresan citocinas IFN-gamma y/o IL-2 y/o TNF-alfa en el día 21 (es decir, 7 días después de la segunda inmunización).

Figura 10: Perfil de citocinas en el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD4 específica del antígeno en los ratones C57BL/6 inmunizados.

Figura 11: Porcentaje de células CD4 y CD8 a partir de los ratones C57BL/6 inmunizados que expresan citocinas IFN-gamma y/o IL-2 y/o TNF-alfa en el día 35 (es decir, 7 días después de la tercera inmunización).

Figura 12: Perfil de citocinas en el día 35 (es decir, 7 días después de la tercera inmunización) de la respuesta de CD4 específica del antígeno en los ratones C57BL/6 inmunizados.

Figura 13: Perfil de citocinas en el día 35 (es decir, 7 días después de la tercera inmunización) de la respuesta de CD8 específica del antígeno en los ratones C57BL/6 inmunizados.

Figura 14: Respuestas de células T CD4 específicas del antígeno en seres humanos limpios y latentemente infectados.

Figura 15: Alineación de secuencias de proteína Rv3616c de tipo silvestre.

Figuras 16A y 16B: Alineación de secuencias de proteína Rv3616c modificada de ejemplo.

Figura 17: Resultados de SDS-PAGE de los experimentos de expresión inicial del antígeno.

Figura 18: Resultados de SDS-PAGE de los experimentos de otra expresión del antígeno.

Figura 19: Resultados de SDS-PAGE de los experimentos de expresión adicional del antígeno.

Figura 20: Porcentaje de células CD4 a partir de los ratones inmunizados que expresan citocinas IFN-gamma y/o IL-2 y/o TNF-alfa a los 7 días después de la segunda, y a los 7 días después de la tercera inmunización con Rv3616Δ138-145.

Figura 21: Perfil de citocinas de la respuesta de células T CD4 específica de Rv3616 a los 7 días después de la segunda inmunización con Rv3616Δ138-145.

Figura 22: Perfil de citocinas de la respuesta de células T CD4 específica de Rv3616 a los 7 días después de la tercera inmunización con Rv3616Δ138-145.

Figura 23: Porcentaje de células CD8 a partir de los ratones inmunizados que expresan citocinas IFN-gamma y/o IL-2 y/o TNF-alfa a los 7 días después de la segunda, y a los 7 días después de la tercera inmunización con Rv3616Δ138-145.

Figura 24: Perfil de citocinas de la respuesta de células T CD8 específica de Rv3616 a los 7 días después de la segunda inmunización con Rv3616Δ138-145.

Figura 25: Perfil de citocinas de la respuesta de células T CD8 específica de Rv3616 a los 7 días después de la tercera inmunización con Rv3616Δ138-145.

Descripción de las secuencias del listado

SEQ ID No: 1: secuencia de polipéptido de Rv3616c a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 2: secuencia de polinucleótido de Rv3616c a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 3: secuencia de polipéptido de Rv3616c a partir de *M. tuberculosis* cepa CDC1551.
SEQ ID No: 4: secuencia de polipéptido de Rv3616c a partir de *M. tuberculosis* cepa F11.
SEQ ID No: 5: secuencia de polipéptido de Rv3616c a partir de *M. tuberculosis* cepa Haarlem A.
SEQ ID No: 6: secuencia de polipéptido de Rv3616c a partir de *M. tuberculosis* cepa C.
SEQ ID No: 7: secuencia de polipéptido de Rv3616c a partir de BCG.
SEQ ID No: 8: secuencia de polipéptido de Mtb8.4.
SEQ ID No: 9: secuencia de polipéptido de Mtb9.8.
SEQ ID No: 10: secuencia de polipéptido de Mtb9.9.
SEQ ID No: 11: secuencia de polipéptido de Ra12.
SEQ ID No: 12: secuencia de polipéptido de Ra35.
SEQ ID No: 13: secuencia de polipéptido de TbH9.
SEQ ID No: 14: secuencia de polipéptido de Mtb41.
SEQ ID No: 15: secuencia de polipéptido de ESAT-6.
SEQ ID No: 16: secuencia de polipéptido de Ag85A.
SEQ ID No: 17: secuencia de polipéptido de Ag85B.
SEQ ID No: 18: secuencia de polipéptido de alfa-cristalina.
SEQ ID No: 19: secuencia de polipéptido de MPT64.
SEQ ID No: 20: secuencia de polipéptido de Mtb32A.
SEQ ID No: 21: secuencia de polipéptido de Mtb32A madura mutada en Ser/Ala.
SEQ ID No: 22: secuencia de polipéptido de TB10.4.

SEQ ID No: 23: secuencia de polipéptido de Mtb72f.
SEQ ID No: 24: secuencia de polipéptido de M72.
SEQ ID No: 25: secuencia de polipéptido de Mtb71f.
SEQ ID No: 26: secuencia de polipéptido de fusión M92.
SEQ ID No: 27: secuencia de polipéptido de fusión M103.
SEQ ID No: 28: secuencia de polipéptido de fusión M114.
SEQ ID No: 29: epítopo de células CD4 humano supuesto 1.
SEQ ID No: 30: epítopo de células CD4 humano supuesto 2.
SEQ ID No: 31: epítopo de células CD4 humano supuesto 3.
SEQ ID No: 32: epítopo de células CD4 humano supuesto 4.
SEQ ID No: 33: epítopo de células CD4 humano supuesto 5.
SEQ ID No: 34: epítopo de células CD4 humano supuesto 6.
SEQ ID No: 35: epítopo de células CD4 humano supuesto 7.
SEQ ID No: 36: epítopo de células CD4 humano supuesto 8.
SEQ ID No: 37: epítopo de células CD4 humano supuesto 9.
SEQ ID No: 38: epítopo de células CD4 humano supuesto 10.
SEQ ID No: 39: epítopo de células CD4 humano supuesto 11.
SEQ ID No: 40: epítopo de células CD4 humano supuesto 12.
SEQ ID No: 41: epítopo de células CD4 humano supuesto 13.
SEQ ID No: 42: epítopo de células CD4 humano supuesto 14.
SEQ ID No: 43: epítopo de células CD4 humano supuesto 15.
SEQ ID No: 44: epítopo de células CD4 humano supuesto 16.
SEQ ID No: 45: epítopo de células CD4 humano supuesto 17.
SEQ ID No: 46: epítopo de células CD4 humano supuesto 18.
SEQ ID No: 47: epítopo de células CD4 humano supuesto 19.
SEQ ID No: 48: epítopo de células CD8 humano supuesto 1.
SEQ ID No: 49: epítopo de células CD8 humano supuesto 2.
SEQ ID No: 50: epítopo de células CD8 humano supuesto 3.
SEQ ID No: 51: epítopo de células CD8 humano supuesto 4.
SEQ ID No: 52: epítopo de células CD8 humano supuesto 5.
SEQ ID No: 53: epítopo de células CD8 humano supuesto 6.
SEQ ID No: 54: epítopo de células CD8 humano supuesto 7.
SEQ ID No: 55: epítopo de células CD8 humano supuesto 8.
SEQ ID No: 56: epítopo de células CD8 humano supuesto 9.
SEQ ID No: 57: epítopo de células CD8 humano supuesto 10.
SEQ ID No: 58: epítopo de células CD8 humano supuesto 11.
SEQ ID No: 59: epítopo de células CD8 humano supuesto 12.
SEQ ID No: 60: epítopo de células CD8 humano supuesto 13.
SEQ ID No: 61: epítopo de células CD8 humano supuesto 14.
SEQ ID No: 62: epítopo de células CD8 humano supuesto 15.
SEQ ID No: 63: epítopo de células CD8 humano supuesto 16.
SEQ ID No: 64: epítopo de células CD8 humano supuesto 17.
SEQ ID No: 65: epítopo de células CD8 humano supuesto 18.
SEQ ID No: 66: epítopo de células CD8 humano supuesto 19.
SEQ ID No: 67: epítopo de células CD8 humano supuesto 20.
SEQ ID No: 68: epítopo de células CD8 humano supuesto 21.
SEQ ID No: 69: epítopo de células CD8 humano supuesto 22.
SEQ ID No: 70: epítopo de células CD8 humano supuesto 23.
SEQ ID No: 71: epítopo de células CD8 humano supuesto 24.
SEQ ID No: 72: epítopo de células CD8 humano supuesto 25.
SEQ ID No: 73: epítopo de células CD8 humano supuesto 26.
SEQ ID No: 74: epítopo de células CD8 humano supuesto 27.
SEQ ID No: 75: epítopo de células CD8 humano supuesto 28.
SEQ ID No: 76: epítopo de células CD8 humano supuesto 29.
SEQ ID No: 77: epítopo de células CD8 humano supuesto 30.
SEQ ID No: 78: epítopo de células CD8 humano supuesto 31.
SEQ ID No: 79: epítopo de células CD8 humano supuesto 32.
SEQ ID No: 80: epítopo de células CD8 humano supuesto 33.
SEQ ID No: 81: epítopo de células CD8 humano supuesto 34.
SEQ ID No: 82: epítopo de células CD8 humano supuesto 35.
SEQ ID No: 83: epítopo de células CD8 humano supuesto 36.
SEQ ID No: 84: epítopo de células CD8 humano supuesto 37.
SEQ ID No: 85: epítopo de células CD8 humano supuesto 38.
SEQ ID No: 86: epítopo de células CD8 humano supuesto 39.
SEQ ID No: 87: epítopo de células CD8 humano supuesto 40.
SEQ ID No: 88: epítopo de células CD8 humano supuesto 41.

SEQ ID No: 89: epítopo de células CD8 humano supuesto 42.
SEQ ID No: 90: epítopo de células CD8 humano supuesto 43.
SEQ ID No: 91: epítopo de células CD8 humano supuesto 44.
SEQ ID No: 92: epítopo de células CD8 humano supuesto 45.
SEQ ID No: 93: epítopo de células CD8 humano supuesto 46.
SEQ ID No: 94: epítopo de células CD8 humano supuesto 47.
SEQ ID No: 95: epítopo de células CD8 humano supuesto 48.
SEQ ID No: 96: epítopo de células CD8 humano supuesto 49.
SEQ ID No: 97: epítopo de células CD8 humano supuesto 50.
SEQ ID No: 98: epítopo de células CD8 humano supuesto 51.
SEQ ID No: 99: epítopo de células CD8 humano supuesto 52.
SEQ ID No: 100: epítopo de células CD8 humano supuesto 53.
SEQ ID No: 101: epítopo de células CD8 humano supuesto 54.
SEQ ID No: 102: epítopo de células CD8 humano supuesto 55.
SEQ ID No: 103: epítopo de células CD8 humano supuesto 56.
SEQ ID No: 104: epítopo de células CD8 humano supuesto 57.
SEQ ID No: 105: epítopo de células CD8 humano supuesto 58.
SEQ ID No: 106: epítopo de células CD8 humano supuesto 59.
SEQ ID No: 107: epítopo de células CD8 humano supuesto 60.
SEQ ID No: 108: epítopo de células CD8 humano supuesto 61.
SEQ ID No: 109: epítopo de células CD8 humano supuesto 62.
SEQ ID No: 110: epítopo de células CD8 humano supuesto 63.
SEQ ID No: 111: epítopo de células CD8 humano supuesto 64.
SEQ ID No: 112: epítopo de células CD8 humano supuesto 65.
SEQ ID No: 113: epítopo de células CD8 humano supuesto 66.
SEQ ID No: 114: epítopo de células CD8 humano supuesto 67.
SEQ ID No: 115: epítopo de células CD8 humano supuesto 68.
SEQ ID No: 116: epítopo de células CD8 humano supuesto 69.
SEQ ID No: 117: epítopo de células CD8 humano supuesto 70.
SEQ ID No: 118: epítopo de células CD8 humano supuesto 71.
SEQ ID No: 119: epítopo de células CD8 humano supuesto 72.
SEQ ID No: 120: epítopo de células CD8 humano supuesto 73.
SEQ ID No: 121: epítopo de células CD8 humano supuesto 74.
SEQ ID No: 122: epítopo de células CD8 humano supuesto 75.
SEQ ID No: 123: epítopo de células CD8 humano supuesto 76.
SEQ ID No: 124: epítopo de células CD8 humano supuesto 77.
SEQ ID No: 125: epítopo de células CD8 humano supuesto 78.
SEQ ID No: 126: epítopo de células CD8 humano supuesto 79.
SEQ ID No: 127: péptido 1.
SEQ ID No: 128: péptido 2.
SEQ ID No: 129: péptido 3.
SEQ ID No: 130: péptido 4.
SEQ ID No: 131: péptido 5.
SEQ ID No: 132: péptido 6.
SEQ ID No: 133: péptido 7.
SEQ ID No: 134: péptido 8.
SEQ ID No: 135: péptido 9.
SEQ ID No: 136: péptido 10.
SEQ ID No: 137: péptido 11.
SEQ ID No: 138: péptido 12.
SEQ ID No: 139: péptido 13.
SEQ ID No: 140: péptido 14.
SEQ ID No: 141: péptido 15.
SEQ ID No: 142: péptido 16.
SEQ ID No: 143: péptido 17.
SEQ ID No: 144: péptido 18.
SEQ ID No: 145: péptido 19.
SEQ ID No: 146: péptido 20.
SEQ ID No: 147: péptido 21.
SEQ ID No: 148: péptido 22.
SEQ ID No: 149: péptido 23.
SEQ ID No: 150: péptido 24.
SEQ ID No: 151: péptido 25.
SEQ ID No: 152: péptido 26.
SEQ ID No: 153: péptido 27.
SEQ ID No: 154: péptido 28.

SEQ ID No: 155: péptido 29.
SEQ ID No: 156: péptido 30.
SEQ ID No: 157: secuencia de polipéptido de Rv1753c a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 158: secuencia de polipéptido de Rv2386c a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 159: secuencia de polipéptido de Rv2707c a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 160: secuencia de polinucleótido de codones optimizados de E. coli para Rv3616c a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 161: secuencia de polipéptido de Rv3616cΔ136-183 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 162: secuencia de polipéptido de Rv3616cΔ150-160 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 163: secuencia de polipéptido de Rv3616cΔ136-154 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 164: secuencia de polipéptido de Rv3616cΔ166-182 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 165: secuencia de polipéptido de Rv3616cΔ135-139 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 166: secuencia de polipéptido de Rv3616cΔ142-145 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 167: secuencia de polipéptido de Rv3616cΔ138-145 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 168: secuencia de polipéptido de Rv3616cΔ145-152 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 169: secuencia de polipéptido de Rv3616cΔ149-154 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 170: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ136-183 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 171: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ150-160 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 172: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ136-154 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 173: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ166-182 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 174: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ135-139 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 175: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ142-145 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 176: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ138-145 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 177: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ145-152 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 178: secuencia de polinucleótido de codones optimizados de E. coli que codifica Rv3616cΔ149-154 derivada a partir de *M. tuberculosis* cepa H37Rv.
SEQ ID No: 179: secuencia de polipéptido de proteína Rv3616c modificada basada en la separación y reconfiguración alrededor de los residuos 137-139 a partir de *M. tuberculosis* cepa H37Rv, incluyendo la supresión de Cys138.
SEQ ID No: 180: secuencia de polipéptido de proteína Rv3616c modificada basada en la separación y reconfiguración alrededor de los residuos 152-153 a partir de *M. tuberculosis* cepa H37Rv.

Descripción detallada

La presente invención se refiere en términos generales al uso de polipéptidos de Rv3616c modificados, o de los polinucleótidos que los codifican, en el campo de las infecciones micobacterianas latentes. Adicionalmente, la presente invención se refiere a proteínas Rv3616c modificadas particulares. Los inventores han descubierto, de una manera sorprendente, que la alteración de la hidrofobicidad de una región particular de una secuencia de proteína Rv3616c puede conducir a una mejor expresión sin un impacto perjudicial sustancial a las propiedades inmunogénicas. Las proteínas Rv3616c modificadas son útiles como antígenos de tuberculosis (TB), en particular como antígenos de tuberculosis (TB) latente.

Se ha demostrado que algunas de las proteínas que son fuertemente expresadas durante las primeras etapas de la infección por *Mycobacterium* proporcionan una fuerte eficacia protectora en los modelos animales de vacunación. Sin embargo, la vacunación con antígenos que son altamente expresados durante las primeras etapas de la infección puede no proporcionar una respuesta inmunitaria óptima para tratar con las etapas posteriores de la infección. El control adecuado durante la infección latente puede requerir de células T que sean específicas para los antígenos particulares que se expresen en ese momento.

Las vacunas posteriores a la exposición que se dirigen directamente hacia las bacterias latentes persistentes

pueden ayudar a proteger contra la reactivación de la tuberculosis (TB), mejorando de esta manera el control de la tuberculosis (TB), o haciendo posible incluso la eliminación de la infección. Por consiguiente, una vacuna que se dirija hacia la tuberculosis (TB) latente podría reducir de una manera significativa y económica los índices globales de infección por tuberculosis (TB).

También se podrían utilizar vacunas subunitarias basadas en los antígenos de la etapa tardía en combinación con los antígenos de la etapa temprana para proporcionar una vacuna multifacética. De una manera alternativa, los antígenos de la etapa tardía se podrían utilizar para complementar y mejorar la vacunación con BCG (ya sea mediante el refuerzo de la respuesta a BCG o bien a través del desarrollo de cepas de BCG recombinantes avanzadas).

Aunque se ha demostrado que los macrófagos actúan como los efectores principales de la inmunidad a *Mycobacterium*, las células T son los inductores predominantes de esa inmunidad. La función esencial de las células T en la protección contra la tuberculosis es ilustrada por el aumento en los índices de reactivación de la tuberculosis (TB) en los individuos infectados por el virus de inmunodeficiencia humana, debido a la consunción asociada de las células T CD4+. Adicionalmente, se ha demostrado que la transferencia adoptiva de las células T CD4+ tomada a la altura de la respuesta inmunitaria primaria a *M. tuberculosis*, confiere protección contra *M. tuberculosis* en los ratones deficientes en células T (Orme y col., J. Exp. Med., 1983, 158: 74-83).

Se ha demostrado que las células T CD4+ que reaccionan con *Mycobacterium* son potentes productoras de γ-interferón (IFN-γ), el cual, a su vez, se ha demostrado que desencadena los efectos anti-micobacterianos de los macrófagos en los ratones (Flynn y col., J. Exp. Med., 1993, 178: 2249-2254). Aunque está menos clara la función del IFN-γ en los seres humanos, los estudios han demostrado que la 1,25-dihidroxi-vitamina D3, ya sea sola o bien en combinación con IFN-γ o con el factor de necrosis tumoral-alfa, activa los macrófagos humanos para inhibir la infección por *M. tuberculosis*. Adicionalmente, se sabe que el IFN-γ estimula los macrófagos humanos para elaborar la 1,25-dihidroxi-vitamina D3. De una manera similar, se ha demostrado que la interleucina-12 (IL-12) tiene una función en el estímulo de la resistencia a la infección por *M. tuberculosis*. Para una revisión de la inmunología de la infección por *M. tuberculosis*, véase Chan y Kaufmann, Tuberculosis: Pathogenesis, Protection and Control (Bloom, Editor, 1994), Tuberculosis (2a Edición, Rom y Garay, Editores, 2003), y Harrison's Principles of Internal Medicine, Capítulo 150, páginas 953-966 (16a Edición, Braunwald y col., Editores, 2005).

El diagnóstico de la infección por tuberculosis (TB) latente se lleva a cabo comúnmente utilizando la prueba de tuberculina de piel, la cual involucra la exposición intradérmica al derivado purificado de proteína tuberculina (PPD). Las respuestas de células T específicas del antígeno dan como resultado una induración mensurable en el sitio de la inyección por 48 a 72 horas después de la inyección, lo cual indica la exposición a los antígenos micobacterianos. Sin embargo, con esta prueba, la sensibilidad y la especificidad han sido un problema, y los individuos vacunados con BCG no siempre se pueden distinguir fácilmente de los individuos infectados (esto está en particular importante a la luz del hecho de que el BCG no protege contra la infección latente). En general, los individuos que han recibido BCG pero que no están infectados por *M. tuberculosis*, muestran una reacción a PPD por debajo de 10 milímetros de diámetro, mientras que se considera que las personas que tienen una reacción a PPD por encima de 10 milímetros de diámetro, han sido infectadas por *M. tuberculosis*. Sin embargo, esta regla no es aplicable a los individuos con inmunosupresión debida a la infección por VIH (la cual puede dar como resultado una reacción a PPD por debajo de 10 milímetros de diámetro) ; o en los países endémicos, en donde las personas infectadas por micobacterias que no son de tuberculosis, pueden mostrar una reacción a PPD por encima de 10 milímetros de diámetro.

El progreso durante los años recientes ha visto el desarrollo de ensayos basados en células T *in vitro*, basados en la liberación del interferón-gamma, y utilizan antígenos que son más específicos para *M. tuberculosis* que la PPD, es decir, ESAT-6 y CFP-10. Estas pruebas de alta especificidad parecen ser al menos tan sensibles como la prueba de tuberculina de piel, y también demuestran menos reactividad cruzada debida a la vacunación con BCG. Véase Pai M y col., Expert Rev. Mol. Diagn., 2006, 6(3) : 413-422, para ver una reseña reciente del diagnóstico de tuberculosis (TB) latente. Sin embargo, debido a que los ESAT-6/CFP-10 son los antígenos de la etapa temprana, los ensayos basados en ESAT-6/CFP-10 solamente pueden funcionar de una manera óptima en las personas recientemente infectadas. En consecuencia, la identificación de antígenos específicamente asociados con la tuberculosis latente puede ayudar al desarrollo de ensayos más sensibles que puedan asegurar la detección de infecciones latentes a más largo plazo.

Sigue existiendo una necesidad de estrategias efectivas para el tratamiento y la prevención de tuberculosis, en particular para el tratamiento y la prevención de tuberculosis (TB) latente, y para la prevención de la reactivación de tuberculosis (TB).

Recientemente, se han propuesto una gama de vacunas de *M. tuberculosis* candidatas, basándose en un análisis bioinformático del genoma entero de *M. tuberculosis* (Zvi y col., BMC Medical Genetics, 2008, 1: 18), y en la prueba de proteínas diferencialmente expresadas en los individuos activamente y latentemente infectados (Schuck SD y col., PLoS ONE, 2009, 4(5) : e5590).

La Rv3616c, también conocida como Mtb40, HTCC1 y EspA, está involucrada en el sistema de secreción ESX-1 de

Mycobacterium tuberculosis (Woodsworth y col., Infection and Immunity, 2008, 76(9) : 4199-4205). La Rv3616c ha estado anteriormente implicada en las respuestas inmunitarias asociadas con tuberculosis (véase, por ejemplo, la Publicación Internacional Número WO98/53075). Al-Attiyah y col., Clin. Exp. Immunol., 2004, 138: 139-144 han demostrado que la Rv3616c es bien reconocida (a través de la proliferación de células mononucleares de sangre periférica (PBMC) y de la producción de IFN-gamma) por los pacientes con tuberculosis pulmonar. Mustafa y col., Infect. Immun. 2006, 74(8) : 4566-4572 han investigado el reconocimiento de la Rv3616c por ganado M. bovis infectado vacunado con BCG.

La solicitud internacional de patente número PCT/EP2009/ 059580, publicada como WO2010/010177, describe la identificación de la Rv3616c como un antígeno asociado con la etapa latente de la infección por tuberculosis (TB).

La solicitud internacional de patente número WO2010/121618 propone el uso de proteínas constitutivamente expresadas y de los genes que las codifican, para composiciones inmunológicas, tales como vacunas, incluyendo EspA (es decir, Rv3616c).

Deseablemente se producen antígenos de vacuna que tengan su secuencia de tipo silvestre, asegurando, por consiguiente, que las respuestas inmunológicas solicitadas por la vacuna, correspondan estrechamente con aquéllas requeridas para contrarrestar la infección por un patógeno. No obstante, la producción eficiente de antígenos es un factor importante para reducir los costos asociados con la elaboración de la vacuna. En consecuencia, los antígenos modificados que se expresen de una manera conveniente en altos niveles, pero que eliminen el impacto perjudicial sobre la inmunogenicidad, podrían proporcionar un beneficio sustancial. La presente invención busca proporcionar antígenos de Rv3616c modificados que resuelvan éste y otros problemas.

Sin limitarse por la teoría, se piensa que los residuos de aminoácidos 134-183 de *Mycobacterium tuberculosis* cepa H37Rv, Rv3616c, corresponden a una región transmembrana potencial, una región de baja complejidad, y una bobina enrollada. La alteración de uno, dos, o los tres de estos elementos estructurales hace posible que la secuencia de la proteína Rv3616c modificada resultante se exprese en mayores niveles.

En consecuencia, en su aspecto más amplio, la presente invención proporciona una proteína Rv3616c modificada, en donde se ha alterado la hidrofobicidad de los residuos de aminoácidos correspondientes a los residuos 134-183 de la secuencia de H37Rv, de una manera adecuada una proteína Rv3616c modificada, en donde se ha alterado la hidrofobicidad de los residuos de aminoácidos correspondientes a los residuos 135-154 de la secuencia de H37Rv.

El término ‘alteración de la hidrofobicidad’ significa una modificación de la secuencia que da como resultado una hidrofobicidad suficientemente reducida, de tal manera que se puede expresar más eficientemente la secuencia de la proteína Rv3616c modificada.

Deseablemente, el grado de modificaciones en relación con la secuencia de tipo silvestre se debe mantener hasta un mínimo, con el fin de reducir la probabilidad de cualquier impacto perjudicial sobre la inmunogenicidad.

Como se utiliza en la presente, un ‘enlace peptídico directo’ es un enlace peptídico en donde dos péptidos se enlazan por medio de enlaces de péptidos directamente uno con el otro, y sin una secuencia de aminoácidos que intervenga. Un ‘enlace peptídico indirecto’ es un enlace peptídico en donde dos péptidos se enlazan por medio de enlaces de péptidos con un tercer péptido que intervenga.

En el contexto de la presente invención, existen cuatro planteamientos principales para la alteración de la hidrofobicidad – es decir, la separación de los residuos hidrofóbicos, la supresión de los residuos hidrofóbicos, la sustitución de los residuos hidrofóbicos con residuos hidrofílicos, y la adición de residuos hidrofílicos. La persona experta reconocerá que también se puede utilizar una combinación de estos planteamientos. Sin embargo, como se menciona anteriormente, el grado de las modificaciones de la secuencia se debe minimizar idealmente para evitar un impacto perjudicial innecesario sobre la inmunogenicidad.

La separación de los residuos hidrofóbicos se puede lograr dividiendo una secuencia de proteína Rv3616c en una localización entre los aminoácidos correspondientes a los residuos 133 a 184 de la SEQ ID NO: 1 en un fragmento N-terminal y un fragmento C-terminal, seguida por la reconfiguración de esas porciones, de tal manera que el fragmento N-terminal se localiza en la región C-terminal de la proteína Rv3616c modificada, y el fragmento C-terminal se localiza en la región N-terminal de la proteína Rv3616c modificada.

En un aspecto de la invención, se proporciona una proteína Rv3616c modificada, comprendiendo esta proteína Rv3616c modificada, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo C de la proteína Rv3616c modificada en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 184-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están unidos directa o indirectamente.

En algunas realizaciones, la proteína Rv3616c modificada consiste esencialmente en, o de una manera alternativa consiste en, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo C de la proteína Rv3616c modificada en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 184-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están unidos directa o indirectamente.

El primer polipéptido puede ser una secuencia que tenga al menos el 95 % de identidad con los residuos 1-133 de la SEQ ID NO: 1, tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

El segundo polipéptido puede ser una secuencia que tenga al menos el 95 % de identidad con los residuos 184-392 de la SEQ ID NO: 1, tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

De una manera adecuada, el primer polipéptido puede ser una secuencia que tenga al menos el 90 % de identidad con los residuos 1-134 de la SEQ ID NO: 1, en particular al menos el 95 % de identidad, tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

De una manera adecuada, el segundo polipéptido puede ser una secuencia que tenga al menos el 90 % de identidad con los residuos 155-392 de la SEQ ID NO: 1, en particular al menos el 95 % de identidad, tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

De una manera adecuada, la proteína Rv3616c modificada del primer aspecto no comprende una secuencia que tenga al menos el 90 % de identidad con la SEQ ID NO: 1 de longitud completa. De una manera adecuada, la proteína Rv3616c modificada del primer aspecto es de menos de 500 aminoácidos de largo, tal como de menos de 450 aminoácidos de largo, en particular de menos de 400 aminoácidos de largo.

El enlace peptídico puede ser directo. De una manera alternativa, el enlace peptídico puede ser indirecto.

En un segundo aspecto de la invención, se proporciona una proteína Rv3616c modificada, comprendiendo esta proteína Rv3616c modificada, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo C de la proteína Rv3616c modificada en relación con el segundo polipéptido, y en donde:

(iii) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1; y
(iv) el segundo polipéptido es una secuencia contigua de al menos 155 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están unidos directa o indirectamente.

En algunas realizaciones, la proteína Rv3616c modificada consiste esencialmente en, o de una manera alternativa consiste en, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo C de la proteína Rv3616c modificada en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia contigua de al menos 155 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están unidos directa o indirectamente.

El primer polipéptido puede ser una secuencia contigua de al menos 110 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1, tal como de al menos 120 aminoácidos o de al menos 130 aminoácidos, por ejemplo, los residuos 1-133.

El segundo polipéptido puede ser una secuencia contigua de al menos 180 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1, tal como de al menos 190 aminoácidos o de al menos 200 aminoácidos, por ejemplo, los residuos 184-392.

De una manera adecuada, el primer polipéptido puede ser una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-134 de la SEQ ID NO: 1, en particular de al menos 110 aminoácidos, tal como de al menos 120 aminoácidos o de al menos 130 aminoácidos, por ejemplo, los residuos 1-134.

De una manera adecuada, el segundo polipéptido puede ser una secuencia contigua de al menos 175 aminoácidos dentro de los residuos 155-392 de la SEQ ID NO: 1, en particular de al menos 200 aminoácidos, tal como de al menos 210 aminoácidos o de al menos 220 aminoácidos, por ejemplo, los residuos 155-392. Las realizaciones en donde el segundo polipéptido es una secuencia contigua de al menos 235 aminoácidos dentro de los residuos 155-392 de la SEQ ID NO: 1 son también de interés.

De una manera adecuada, la proteína Rv3616c modificada del segundo aspecto no comprende una secuencia contigua de más de 259 aminoácidos a partir de la SEQ ID NO: 1. De una manera alternativa, la proteína Rv3616c modificada del segundo aspecto no comprende una secuencia contigua de más de 257 aminoácidos, una secuencia contigua de más de 255 aminoácidos, o una secuencia contigua de más de 253 aminoácidos. De una manera adecuada, la proteína Rv3616c modificada del segundo aspecto es de menos de 500 aminoácidos de largo, tal como de menos de 450 aminoácidos de largo, en particular de menos de 400 aminoácidos de largo.

El enlace peptídico puede ser un enlace directo o indirecto.

Los ejemplos de los primero y segundo aspectos incluyen las proteínas Rv3616c modificadas, en donde los primero y segundo polipéptidos corresponden a los fragmentos N-terminal y C-terminal resultantes de la división de una secuencia de Rv3616c en una localización entre los aminoácidos correspondientes a los residuos 135-154 de la SEQ ID NO: 1, por ejemplo, los residuos 138-139 o 152-153, por ejemplo, los residuos 138-139 o 152-153 en donde el enlace peptídico es directo. De una manera adecuada, cuando los primero y segundo polipéptidos se reconfiguran, la metionina inicial se deja en el extremo N de la proteína Rv3616c modificada. Véase, por ejemplo, las SEQ ID NOs: 179 y 180, las cuales ilustran este tipo de configuración.

La supresión de los residuos hidrofóbicos se puede lograr a través de la remoción de al menos un aminoácido correspondiente a los residuos 134 a 183 de la SEQ ID NO: 1. Los residuos suprimidos pueden ser no contiguos y/o contiguos.

De una manera adecuada, la supresión de los residuos hidrofóbicos se puede lograr a través de la remoción de al menos dos aminoácidos correspondientes a los residuos 134 a 183 de la SEQ ID NO: 1. La supresión de los residuos hidrofóbicos también se puede lograr a través de la remoción de al menos tres aminoácidos correspondientes a los residuos 134 a 183 de la SEQ ID NO: 1.

Los residuos suprimidos pueden ser no contiguos y/o contiguos.

Se puede observar que las secuencias de Rv3616c de tipo silvestre contienen un residuo Cys en la localización 138. De una manera adecuada, este residuo Cys es suprimido o reemplazado (por ejemplo, C138Q).

En un tercer aspecto de la invención, se proporciona una proteína Rv3616c modificada, comprendiendo esta proteína o, de una manera alternativa, consistiendo esencialmente o consistiendo en, una secuencia de Rv3616c, en donde al menos un aminoácido (por ejemplo, al menos 2) ha sido suprimido de la región correspondiente a los residuos 134-183 de la SEQ ID NO: 1.

La proteína Rv3616c modificada puede comprender o, de una manera alternativa, puede consistir esencialmente o consistir en, una secuencia de Rv3616c, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4) de la región correspondiente a los residuos 134-183 de la SEQ ID NO: 1.

Son de interés las proteínas Rv3616c modificadas particulares que comprenden una secuencia de Rv3616c, en en la que se ha suprimido al menos 1 aminoácido (por ejemplo, al menos 2) de la región correspondiente a los residuos 135-154 de la SEQ ID NO: 1. Otras secuencias de interés son las proteínas Rv3616c modificadas, las cuales comprenden una secuencia de Rv3616c, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4) de la región correspondiente a los residuos 135-154 de la SEQ ID NO: 1.

La porción contigua suprimida puede ser de al menos 5 aminoácidos (por ejemplo, de 5 a 30, tal como de 5 a 20, o de 5 a 15), en especial de al menos 6 aminoácidos (por ejemplo, de 6 a 30, tal como de 6 a 20, o de 6 a 15), en particular de al menos 7 aminoácidos (por ejemplo, de 7 a 30, tal como de 7 a 20, o de 7 a 15), tal como de al menos 8 aminoácidos (por ejemplo, de 8 a 30, tal como de 8 a 20, o de 8 a 15), o de al menos 10 aminoácidos (por ejemplo, de 10 a 30, tal como de 10 a 20, o de 10 a 15).

En ciertas realizaciones, la porción contigua suprimida puede ser de:

- 4 aminoácidos, tales como aquéllos correspondientes a los residuos 142- 145 de la SEQ ID NO: 1;
- 5 aminoácidos, tales como aquéllos correspondientes a los residuos 135- 139 de la SEQ ID NO: 1;
- 6 aminoácidos, tales como aquéllos correspondientes a los residuos 149- 154 de la SEQ ID NO: 1;
- 8 aminoácidos, tales como aquéllos correspondientes a los residuos 138- 145 de la SEQ ID NO: 1, o a los residuos 145- 152 de la SEQ ID NO: 1;
- 11 aminoácidos, tales como aquéllos correspondientes a los residuos 150- 160 de la SEQ ID NO: 1;
- 17 aminoácidos, tales como aquéllos correspondientes a los residuos 166- 182 de la SEQ ID NO: 1;
- 19 aminoácidos, tales como aquéllos correspondientes a los residuos 136- 154 de la SEQ ID NO: 1;

- 31 aminoácidos, tales como aquéllos correspondientes a los residuos 136- 166 de la SEQ ID NO: 1; o
- 48 aminoácidos, tales como aquéllos correspondientes a los residuos 136-183 de la SEQ ID NO: 1.

En otras realizaciones, la porción contigua suprimida puede ser de 3 a 10 residuos de aminoácidos, tales como de 4 a 10, por ejemplo, de 4 a 8. El número particular de aminoácidos suprimidos puede ser de 3, 4, 5, 6, 7, 8, 9 o 10, en especial de 4, 5, 6 u 8.

En otras realizaciones, la porción suprimida puede ser aquélla correspondiente a los residuos 135-138 de la SEQ ID NO: 1, a los residuos 136-138 de la SEQ ID NO: 1, a los residuos 137-138 de la SEQ ID NO: 1, a los residuos 138-140 de la SEQ ID NO: 1, a los residuos 138-141 de la SEQ ID NO: 1, a los residuos 152-154 de la SEQ ID NO: 1, o la supresión de los residuos 149-151 de la SEQ ID NO: 1.

Un cuarto aspecto de la invención proporciona una proteína Rv3616c modificada, comprendiendo esta proteína un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(iii) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1; y
(iv) el segundo polipéptido es una secuencia contigua de al menos 155 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están unidos directa o indirectamente por medio de un tercer polipéptido, correspondiendo este tercer polipéptido a los residuos 134-183 de la SEQ ID NO: 1, en en la que se ha suprimido al menos 1 aminoácido (por ejemplo, al menos 2).

En algunas realizaciones, la proteína Rv3616c modificada consiste esencialmente en, o de una manera alternativa consiste en, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia contigua de al menos 155 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están unidos directa o indirectamente por medio de un tercer polipéptido, correspondiendo este tercer polipéptido a los residuos 134-183 de la SEQ ID NO: 1, en en la que se ha suprimido al menos 1 aminoácido (por ejemplo, al menos 2).

Son de un interés particular las proteínas que comprenden, o de una manera alternativa que consisten esencialmente o que consisten en, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia contigua de al menos 155 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están unidos directa o indirectamente por medio de un tercer polipéptido, correspondiendo este tercer polipéptido a los residuos 134-183 de la SEQ ID NO: 1, en donde al menos se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4).

El primer polipéptido puede ser una secuencia contigua de al menos 110 aminoácidos dentro de los residuos 1-133 de la SEQ ID NO: 1, tal como de al menos 120 aminoácidos o de al menos 130 aminoácidos (por ejemplo, los residuos 1-133).

El segundo polipéptido puede ser una secuencia contigua de al menos 180 aminoácidos dentro de los residuos 184-392 de la SEQ ID NO: 1, tal como de al menos 190 aminoácidos o de al menos 200 aminoácidos (por ejemplo, los residuos 184-392).

La porción contigua suprimida a partir de los residuos correspondientes a los 134-183 de la SEQ ID NO: 1 puede ser de al menos 5 aminoácidos (por ejemplo, de 5 a 30, tal como de 5 a 20, o de 5 a 15), en especial de al menos 6 aminoácidos (por ejemplo, de 6 a 30, tal como de 6 a 20, o de 6 a 15), en particular de al menos 7 aminoácidos (por ejemplo, de 7 a 30, tal como de 7 a 20, o de 7 a 15), tal como de al menos 8 aminoácidos (por ejemplo, de 8 a 30, tal como de 8 a 20, o de 8 a 15), o de al menos 10 aminoácidos (por ejemplo, de 10 a 30, tal como de 10 a 20, o de 10 a 15).

En ciertas realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 134-183 de la SEQ ID NO: 1 puede ser de:

- 4 aminoácidos, tales como aquéllos correspondientes a los residuos 142- 145 de la SEQ ID NO: 1;
- 5 aminoácidos, tales como aquéllos correspondientes a los residuos 135- 139 de la SEQ ID NO: 1;
- 6 aminoácidos, tales como aquéllos correspondientes a los residuos 149- 154 de la SEQ ID NO: 1;
- 8 aminoácidos, tales como aquéllos correspondientes a los residuos 138- 145 de la SEQ ID NO: 1, o a los residuos 145- 152 de la SEQ ID NO: 1;
- 11 aminoácidos, tales como aquéllos correspondientes a los residuos 150- 160 de la SEQ ID NO: 1;
- 17 aminoácidos, tales como aquéllos correspondientes a los residuos 166- 182 de la SEQ ID NO: 1;
- 19 aminoácidos, tales como aquéllos correspondientes a los residuos 136- 154 de la SEQ ID NO: 1;
- 31 aminoácidos, tales como aquéllos correspondientes a los residuos 136- 166 de la SEQ ID NO: 1; o
- 48 aminoácidos, tales como aquéllos correspondientes a los residuos 136-183 de la SEQ ID NO: 1.

En otras realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 134-183 de la SEQ ID NO: 1 puede ser de 3 a 10 residuos de aminoácidos, tal como de 4 a 10, por ejemplo, de 4 a 8. El número particular de aminoácidos suprimidos puede ser de 3, 4, 5, 6, 7, 8, 9 o 10, en especial de 4, 5, 6 u 8.

En otras realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 134-183 de la SEQ ID NO: 1 puede ser aquélla correspondiente a los residuos 135-138 de la SEQ ID NO: 1, a los residuos 136-138 de la SEQ ID NO: 1, a los residuos 137-138 de la SEQ ID NO: 1, a los residuos 138-140 de la SEQ ID NO: 1, a los residuos 138-141 de la SEQ ID NO: 1, a los residuos 152-154 de la SEQ ID NO: 1, o la supresión de los residuos 149-151 de la SEQ ID NO: 1.

El primer polipéptido y el segundo polipéptido, en algunas realizaciones, estarán directamente unidos. En otras realizaciones el primer polipéptido y el segundo polipéptido estarán indirectamente unidos por medio de un tercer polipéptido. El tercer polipéptido puede corresponder a los residuos 134-183 de la SEQ ID NO: 1, en donde la supresión se ha presentado en una sola porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4). Adicionalmente, el tercer polipéptido puede corresponder a los residuos 134-183 de la SEQ ID NO: 1, en donde las supresiones se han presentado en una pluralidad de distintas localizaciones (por ejemplo, de 1 a 10, tal como de 1 a 5, en particular 1 o 2 localizaciones), siendo cada supresión de 1 a 10, tal como de 1 a 5 residuos de aminoácidos.

De una manera adecuada, el tercer polipéptido es de 48 aminoácidos o menos (por ejemplo, de 10 a 48, tal como de 20 a 48, o de 30 a 48 residuos), tal como de 46 aminoácidos o menos (por ejemplo, de 10 a 46, tal como de 20 a 46, o de 30 a 46 residuos), de 44 aminoácidos o menos (por ejemplo, de 10 a 44, tal como de 20 a 44, o de 30 a 44 residuos), o de 42 aminoácidos o menos (por ejemplo, de 10 a 42, tal como de 20 a 42, o de 30 a 42 residuos).

Un quinto aspecto de la invención proporciona proteínas Rv3616c modificadas, las cuales comprenden un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(iii) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-134 de la SEQ ID NO: 1; y
(iv) el segundo polipéptido es una secuencia contigua de al menos 175 aminoácidos dentro de los residuos 155-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están ya sea directamente unidos o bien indirectamente unidos por medio de un tercer polipéptido, en donde este tercer polipéptido corresponde a los residuos 135-154 de la SEQ ID NO: 1, en en la que se ha suprimido al menos 1 aminoácido (por ejemplo, al menos 2).

En algunas realizaciones, la proteína Rv3616c modificada consiste esencialmente en, o de una manera alternativa consiste en, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-134 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia contigua de al menos 175 aminoácidos dentro de los residuos 155-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están ya sea directamente unidos o bien indirectamente unidos por medio de un tercer polipéptido, en donde este tercer polipéptido corresponde a los residuos 135-154 de la SEQ ID NO: 1, en en la que se ha suprimido al menos 1 aminoácido (por ejemplo, al menos 2).

Son de un interés particular las proteínas que comprenden, o de una manera alternativa que consisten esencialmente o que consisten en, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia contigua de al menos 100 aminoácidos dentro de los residuos 1-134 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia contigua de al menos 175 aminoácidos dentro de los residuos 155-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están ya sea directamente unidos o bien indirectamente unidos por medio de un tercer polipéptido, en donde este tercer polipéptido corresponde a los residuos 135-154 de la SEQ ID NO: 1, en donde al menos se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4).

El primer polipéptido también puede ser una secuencia contigua de al menos 110 aminoácidos dentro de los residuos 1-134 de la SEQ ID NO: 1, tal como de al menos 120 aminoácidos o de al menos 130 aminoácidos, por ejemplo, los residuos 1-134.

El segundo polipéptido también puede ser una secuencia contigua de al menos 200 aminoácidos dentro de los residuos 155-392 de la SEQ ID NO: 1, tal como de al menos 210 aminoácidos o de al menos 220 aminoácidos, por ejemplo, los residuos 155-392. Las realizaciones en donde el segundo polipéptido es una secuencia contigua de al menos 235 aminoácidos dentro de los residuos 155-392 de la SEQ ID NO: 1 son también de interés.

La porción contigua suprimida a partir de los residuos correspondientes a los 135-154 de la SEQ ID NO: 1 puede ser de al menos 5 aminoácidos (por ejemplo, de 5 a 20, tal como de 5 a 15 o de 5 a 10), en especial de al menos 6 aminoácidos (por ejemplo, de 6 a 20, tal como de 6 a 15, o de 6 a 10), en particular de al menos 7 aminoácidos (por ejemplo, de 7 a 20, tal como de 7 a 15, o de 7 a 10), tal como de al menos 8 aminoácidos (por ejemplo, de 8 a 20, tal como de 8 a 15, o de 8 a 10), o de al menos 10 aminoácidos (por ejemplo, de 10 a 20, tal como de 10 a 15).

En ciertas realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 135-154 de la SEQ ID NO: 1 puede ser de:

- 4 aminoácidos, tales como aquéllos correspondientes a los residuos 142- 145 de la SEQ ID NO: 1;
- 6 aminoácidos, tales como aquéllos correspondientes a los residuos 149- 154 de la SEQ ID NO: 1;
- 8 aminoácidos, tales como aquéllos correspondientes a los residuos 138- 145 de la SEQ ID NO: 1, o a los residuos 145- 152 de la SEQ ID NO: 1;
- 11 aminoácidos, tales como aquéllos correspondientes a los residuos 150- 160 de la SEQ ID NO: 1; o
- 19 aminoácidos, tales como aquéllos correspondientes a los residuos 136-154 de la SEQ ID NO: 1.

En otras realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 135-154 puede ser de 3 a 10 residuos de aminoácidos, tal como de 4 a 10, por ejemplo, de 4 a 8. El número particular de aminoácidos suprimidos puede ser de 3, 4, 5, 6, 7, 8, 9 o 10, en especial de 4, 5, 6 u 8.

En otras realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 135-154 de la SEQ ID NO: 1 puede ser aquélla correspondiente a los residuos 135-138 de la SEQ ID NO: 1, a los residuos 136-138 de la SEQ ID NO: 1, a los residuos 137-138 de la SEQ ID NO: 1, a los residuos 138-140 de la SEQ ID NO: 1, a los residuos 138-141 de la SEQ ID NO: 1, a los residuos 152-154 de la SEQ ID NO: 1, o la supresión de los residuos 149-151 de la SEQ ID NO: 1.

El primer polipéptido y el segundo polipéptido, en algunas realizaciones, pueden estar directamente unidos. En otras realizaciones, el primer polipéptido y el segundo polipéptido pueden estar indirectamente unidos por medio de un tercer polipéptido. El tercer polipéptido puede corresponder a los residuos 135-154 de la SEQ ID NO: 1, en donde la supresión se ha presentado en una sola porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4). Adicionalmente, el tercer polipéptido puede corresponder a los residuos 135-154 de la SEQ ID NO: 1, en donde las supresiones se han presentado en una pluralidad de distintas localizaciones (por ejemplo, de 1 a 10, tal como de 1 a 5, en particular 1 o 2 localizaciones), siendo cada supresión de 1 a 10, tal como de 1 a 5 residuos de aminoácidos.

De una manera adecuada, el tercer polipéptido es de 20 aminoácidos o menos (por ejemplo, de 5 a 20, tal como de 10 a 20 residuos), tal como de 18 aminoácidos o menos (por ejemplo, de 5 a 18, tal como de 10 a 18 residuos), de 16 aminoácidos o menos (por ejemplo, de 5 a 16, tal como de 10 a 16 residuos), o de 14 aminoácidos o menos (por ejemplo, de 5 a 14, tal como de 10 a 14 residuos).

Un sexto aspecto de la invención proporciona una proteína Rv3616c modificada, comprendiendo esta proteína un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(iii) el primer polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 1-133 de la SEQ ID NO: 1; y
(iv) el segundo polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 184-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están directamente unidos o indirectamente unidos por medio de un tercer polipéptido, teniendo este tercer polipéptido al menos el 90 % de identidad con una secuencia correspondiente a los residuos 134-183 de la SEQ ID NO: 1, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4).

En algunas realizaciones, la proteína Rv3616c modificada consiste esencialmente en, o de una manera alternativa

consiste en, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 1-133 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 184-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están directamente unidos o indirectamente unidos por medio de un tercer polipéptido, teniendo este tercer polipéptido al menos el 90 % de identidad con una secuencia correspondiente a los residuos 134-183 de la SEQ ID NO: 1, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4).

El primer polipéptido puede ser una secuencia que tenga al menos el 95 % de identidad con los residuos 1-133 de la SEQ ID NO: 1, tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

El segundo polipéptido puede ser una secuencia que tenga al menos el 95 % de identidad con los residuos 184-392 de la SEQ ID NO: 1, tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

El primer polipéptido y el segundo polipéptido, en algunas realizaciones, pueden estar directamente unidos. En otras realizaciones, el primer polipéptido y el segundo polipéptido estarán indirectamente unidos por medio de un tercer polipéptido. El tercer polipéptido puede ser una secuencia que tenga al menos el 95 % de identidad con una secuencia correspondiente a los residuos 134-183 de la SEQ ID NO: 1, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4), tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

La porción contigua suprimida a partir de los residuos correspondientes a los 134-183 de la SEQ ID NO: 1 puede ser de al menos 5 aminoácidos (por ejemplo, de 5 a 30, tal como de 5 a 20, o de 5 a 15), en especial de al menos 6 aminoácidos (por ejemplo, de 6 a 30, tal como de 6 a 20, o de 6 a 15), en particular de al menos 7 aminoácidos (por ejemplo, de 7 a 30, tal como de 7 a 20, o de 7 a 15), tal como de al menos 8 aminoácidos (por ejemplo, de 8 a 30, tal como de 8 a 20, o de 8 a 15), o de al menos 10 aminoácidos (por ejemplo, de 10 a 30, tal como de 10 a 20, o de 10 a 15).

En ciertas realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 134-183 de la SEQ ID NO: 1 puede ser de:

- 4 aminoácidos, tales como aquéllos correspondientes a los residuos 142- 145 de la SEQ ID NO: 1;
- 5 aminoácidos, tales como aquéllos correspondientes a los residuos 135- 139 de la SEQ ID NO: 1;
- 6 aminoácidos, tales como aquéllos correspondientes a los residuos 149- 154 de la SEQ ID NO: 1;
- 8 aminoácidos, tales como aquéllos correspondientes a los residuos 138- 145 de la SEQ ID NO: 1, o a los residuos 145- 152 de la SEQ ID NO: 1;
- 11 aminoácidos, tales como aquéllos correspondientes a los residuos 150- 160 de la SEQ ID NO: 1;
- 17 aminoácidos, tales como aquéllos correspondientes a los residuos 166- 182 de la SEQ ID NO: 1;
- 19 aminoácidos, tales como aquéllos correspondientes a los residuos 136- 154 de la SEQ ID NO: 1;
- 31 aminoácidos, tales como aquéllos correspondientes a los residuos 136- 166 de la SEQ ID NO: 1; o
- 48 aminoácidos, tales como aquéllos correspondientes a los residuos 136-183 de la SEQ ID NO: 1.

En otras realizaciones, la porción contigua suprimida puede ser de 3 a 10 residuos de aminoácidos, tal como de 4 a 10, por ejemplo, de 4 a 8. El número particular de aminoácidos suprimidos puede ser de 3, 4, 5, 6, 7, 8, 9 o 10, en especial de 4, 5, 6 u 8.

En otras realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 134-183 de la SEQ ID NO: 1 puede ser aquélla correspondiente a los residuos 135-138 de la SEQ ID NO: 1, a los residuos 136-138 de la SEQ ID NO: 1, a los residuos 137-138 de la SEQ ID NO: 1, a los residuos 138-140 de la SEQ ID NO: 1, a los residuos 138-141 de la SEQ ID NO: 1, a los residuos 152-154 de la SEQ ID NO: 1, o la supresión de los residuos 149-151 de la SEQ ID NO: 1.

De una manera adecuada, el tercer polipéptido es de 48 aminoácidos o menos (por ejemplo, de 10 a 48, tal como de 20 a 48, o de 30 a 48 residuos), tal como de 46 aminoácidos o menos (por ejemplo, de 10 a 46, tal como de 20 a 46, o de 30 a 46 residuos), de 44 aminoácidos o menos (por ejemplo, de 10 a 44, tal como de 20 a 44, o de 30 a 44 residuos), o de 42 aminoácidos o menos (por ejemplo, de 10 a 42, tal como de 20 a 42, o de 30 a 42 residuos).

Un séptimo aspecto de la invención proporciona proteínas Rv3616c modificadas, las cuales comprenden un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 1-134 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 155-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están ya sea directamente unidos o bien indirectamente unidos por medio de un tercer polipéptido, teniendo este tercer polipéptido al menos el 80 % de identidad con una secuencia correspondiente a los residuos 135-154 de la SEQ ID NO: 1, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4).

En algunas realizaciones, la proteína Rv3616c modificada consiste esencialmente en, o de una manera alternativa consiste en, un primer polipéptido y un segundo polipéptido, localizándose el primer polipéptido hacia el extremo N en relación con el segundo polipéptido, y en donde:

(i) el primer polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 1-134 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos el 90 % de identidad con los residuos 155-392 de la SEQ ID NO: 1;

en donde los primero y segundo polipéptidos están ya sea directamente unidos o bien indirectamente unidos por medio de un tercer polipéptido, teniendo este tercer polipéptido al menos el 80 % de identidad con una secuencia correspondiente a los residuos 135-154 de la SEQ ID NO: 1, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4).

El primer polipéptido puede ser una secuencia que tenga al menos el 95 % de identidad con los residuos 1-134 de la SEQ ID NO: 1, tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

El segundo polipéptido puede ser una secuencia que tenga al menos el 95 % de identidad con los residuos 155-392 de la SEQ ID NO: 1, tal como al menos el 97 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

El primer polipéptido y el segundo polipéptido, en algunas realizaciones, estarán directamente unidos. En otras realizaciones, el primer polipéptido y el segundo polipéptido estarán indirectamente unidos por medio de un tercer polipéptido. El tercer polipéptido puede ser una secuencia que tenga al menos el 90 % de identidad con una secuencia correspondiente a los residuos 135-154 de la SEQ ID NO: 1, en donde se ha suprimido una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4), tal como al menos el 95 % de identidad, al menos el 98 % de identidad, al menos el 99 % de identidad, o incluso el 100 % idéntica.

La porción contigua suprimida a partir de los residuos correspondientes a los 135-154 de la SEQ ID NO: 1 puede ser de al menos 5 aminoácidos (por ejemplo, de 5 a 20, tal como de 5 a 15 o de 5 a 10), en especial de al menos 6 aminoácidos (por ejemplo, de 6 a 20, tal como de 6 a 15, o de 6 a 10), en particular de al menos 7 aminoácidos (por ejemplo, de 7 a 20, tal como de 7 a 15, o de 7 a 10), tal como de al menos 8 aminoácidos (por ejemplo, de 8 a 20, tal como de 8 a 15, o de 8 a 10), o de al menos 10 aminoácidos (por ejemplo, de 10 a 20, tal como de 10 a 15).

En ciertas realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 135-154 de la SEQ ID NO: 1 puede ser de:

- 4 aminoácidos, tales como aquéllos correspondientes a los residuos 142- 145 de la SEQ ID NO: 1;
- 6 aminoácidos, tales como aquéllos correspondientes a los residuos 149- 154 de la SEQ ID NO: 1;
- 8 aminoácidos, tales como aquéllos correspondientes a los residuos 138- 145 de la SEQ ID NO: 1, o a los residuos 145- 152 de la SEQ ID NO: 1;
- 11 aminoácidos, tales como aquéllos correspondientes a los residuos 150- 160 de la SEQ ID NO: 1; o
- 19 aminoácidos, tales como aquéllos correspondientes a los residuos 136-154 de la SEQ ID NO: 1.

En otras realizaciones, la porción contigua suprimida puede ser de 3 a 10 residuos de aminoácidos, tal como de 4 a 10, por ejemplo, de 4 a 8. El número particular de aminoácidos suprimidos puede ser de 3, 4, 5, 6, 7, 8, 9 o 10, en especial de 4, 5, 6 u 8.

En otras realizaciones, la porción contigua suprimida a partir de los residuos correspondientes a los 135-154 de la SEQ ID NO: 1 puede ser aquélla correspondiente a los residuos 135-138 de la SEQ ID NO: 1, a los residuos 136-138 de la SEQ ID NO: 1, a los residuos 137-138 de la SEQ ID NO: 1, a los residuos 138-140 de la SEQ ID NO: 1, a los residuos 138-141 de la SEQ ID NO: 1, a los residuos 152-154 de la SEQ ID NO: 1, o la supresión de los residuos 149-151 de la SEQ ID NO: 1.

De una manera adecuada, el tercer polipéptido es de 20 aminoácidos o menos (por ejemplo, de 5 a 20, tal como de 10 a 20 residuos), tal como de 18 aminoácidos o menos (por ejemplo, de 5 a 18, tal como de 10 a 18 residuos), de 16 aminoácidos o menos (por ejemplo, de 5 a 16, tal como de 10 a 16 residuos), o de 14 aminoácidos o menos (por

ejemplo, de 5 a 14, tal como de 10 a 14 residuos).

La sustitución de los residuos hidrofóbicos se puede lograr a través del reemplazo de al menos un (por ejemplo, de al menos 2) aminoácido correspondiente a los residuos 134 a 183 de la SEQ ID NO: 1 con un residuo hidrofílico. En este aspecto, los residuos hidrofílicos adecuados típicamente serán Gln (Q), Asp (D), Glu (E), Asn (N), His (H), Lys (K), Arg (R), Ser (S) o Thr (T).

Es de un interés particular el reemplazo de al menos un (por ejemplo, de al menos 2) aminoácido correspondiente a los residuos 135 a 154 de la SEQ ID NO: 1 con un residuo hidrofílico. En este aspecto, los residuos hidrofílicos adecuados típicamente serán Gln (Q), Asp (D), Glu (E), Asn (N), His (H), Lys (K), Arg (R), Ser (S) o Thr (T).

Los residuos sustituidos pueden ser no contiguos, aunque de una manera adecuada son contiguos.

En un octavo aspecto de la invención, se proporciona una proteína Rv3616c modificada, comprendiendo esta proteína una secuencia de Rv3616c, en donde una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4) a partir de la región correspondiente a los residuos 134-183 de la SEQ ID NO: 1 ha sido sustituida con residuos hidrofílicos.

En algunas realizaciones, la proteína Rv3616c modificada consiste esencialmente en, o de una manera alternativa consiste en una secuencia de Rv3616c, en donde una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4) a partir de la región correspondiente a los residuos 134-183 de la SEQ ID NO: 1 ha sido sustituida con residuos hidrofílicos.

Son de interés las proteínas Rv3616c modificadas particulares que comprenden una secuencia de Rv3616c, en donde una porción contigua de al menos 3 aminoácidos (por ejemplo, de al menos 4) a partir de la región correspondiente a los residuos 135-154 de la SEQ ID NO: 1 ha sido sustituida con residuos hidrofílicos.

La porción contigua sustituida puede ser de al menos 5 aminoácidos (por ejemplo, de 5 a 30, tal como de 5 a 20, o de 5 a 15), en especial de al menos 6 aminoácidos (por ejemplo, de 6 a 30, tal como de 6 a 20, o de 6 a 15), en particular de al menos 7 aminoácidos (por ejemplo, de 7 a 30, tal como de 7 a 20, o de 7 a 15), tal como de al menos 8 aminoácidos (por ejemplo, de 8 a 30, tal como de 8 a 20, o de 8 a 15), o de al menos 10 aminoácidos (por ejemplo, de 10 a 30, tal como de 10 a 20, o de 10 a 15).

En ciertas realizaciones, la porción contigua sustituida puede ser de:

- 4 aminoácidos, tales como aquéllos correspondientes a los residuos 142- 145 de la SEQ ID NO: 1;
- 5 aminoácidos, tales como aquéllos correspondientes a los residuos 135- 139 de la SEQ ID NO: 1;
- 6 aminoácidos, tales como aquéllos correspondientes a los residuos 149- 154 de la SEQ ID NO: 1;
- 8 aminoácidos, tales como aquéllos correspondientes a los residuos 138- 145 de la SEQ ID NO: 1, o a los residuos 145- 152 de la SEQ ID NO: 1;
- 11 aminoácidos, tales como aquéllos correspondientes a los residuos 150- 160 de la SEQ ID NO: 1;
- 17 aminoácidos, tales como aquéllos correspondientes a los residuos 166- 182 de la SEQ ID NO: 1;
- 19 aminoácidos, tales como aquéllos correspondientes a los residuos 136- 154 de la SEQ ID NO: 1;
- 31 aminoácidos, tales como aquéllos correspondientes a los residuos 136- 166 de la SEQ ID NO: 1; o
- 48 aminoácidos, tales como aquéllos correspondientes a los residuos 136-183 de la SEQ ID NO: 1.

En otras realizaciones, la porción contigua sustituida puede ser de 3 a 10 residuos de aminoácidos, tal como de 4 a 10, por ejemplo, de 4 a 8. El número particular de aminoácidos sustituidos puede ser de 3, 4, 5, 6, 7, 8, 9 o 10, en especial de 4, 5, 6 u 8.

En otras realizaciones, la porción sustituida puede ser aquélla correspondiente a los residuos 135-138 de la SEQ ID NO: 1, a los residuos 136-138 de la SEQ ID NO: 1, a los residuos 137-138 de la SEQ ID NO: 1, a los residuos 138-140 de la SEQ ID NO: 1, a los residuos 138-141 de la SEQ ID NO: 1, a los residuos 152-154 de la SEQ ID NO: 1, o la supresión de los residuos 149-151 de la SEQ ID NO: 1.

La alteración de la hidrofobicidad también se puede lograr mediante la adición de residuos hidrofílicos, por ejemplo, la adición de al menos un residuo de aminoácido hidrofílico (por ejemplo, de al menos 2, tal como de 2 a 10) en una localización entre los residuos correspondientes a los residuos 133 a 184 de la SEQ ID NO: 1. De una manera adecuada, se pueden agregar al menos 3 residuos hidrofílicos (por ejemplo, de 3 a 20, tal como de 3 a 15, en especial de 3 a 10), tal como de al menos 4 residuos (por ejemplo, de 4 a 20, tal como de 4 a 15, en especial de 4 a 10), en particular de al menos 5 residuos (por ejemplo, de 5 a 20, tal como de 5 a 15, en especial de 5 a 10), opcionalmente de al menos 6 residuos (por ejemplo, de 6 a 20, tal como de 6 a 15, en especial de 6 a 10). En este aspecto, los residuos hidrofílicos adecuados típicamente serán Gln (Q), Asp (D), Glu (E), Asn (N), His (H), Lys (K), Arg (R), Ser (S) o Thr (T).

Los residuos hidrofílicos adicionales típicamente se localizarán entre los residuos correspondientes a los residuos 133 a 184 de la SEQ ID NO: 1, en especial entre los residuos correspondientes a los residuos 134 a 155 de la SEQ ID NO: 1 (tal como entre los residuos correspondientes a los residuos 135 a 154 de la SEQ ID NO: 1).

Los residuos hidrofílicos adicionales se pueden distribuir en diferentes posiciones entre los residuos correspondientes a los residuos 133 a 184 de la SEQ ID NO: 1 (por ejemplo, de 1 a 10 localizaciones, tal como de 1 a 5, en particular 1 o 2 localizaciones), teniendo cada localización de 1 a 10 residuos hidrofílicos adicionales, tal como de 1 a 5 residuos adicionales. Los residuos hidrofílicos adicionales se localizarán de una manera adecuada en un grupo contiguo.

En las realizaciones particulares de las proteínas Rv3616c modificadas descritas en los diferentes aspectos anteriores, la proteína Rv3616c modificada no es la SEQ ID NO: 162 (Rv3616cΔ150-160). En otras realizaciones, la proteína Rv3616c modificada no comprende la SEQ ID NO: 162 (Rv3616cΔ150-160).

Las proteínas Rv3616c modificadas se pueden basar en una secuencia de proteína Rv3616c de tipo silvestre a partir de cualquier cepa de *M. tuberculosis*. Por ejemplo, cualquiera de las SEQ ID NOs: 3-7, en particular cualquiera de las SEQ ID NOs: 3-6, puede sustituir a la SEQ ID NO: 1 en las realizaciones anteriores.

Las proteínas de los diferentes aspectos discutidos anteriormente son colectivamente referidas en la presente como proteínas Rv3616c modificadas. También se proporcionan estas proteínas Rv3616c modificadas para utilizarse como medicamentos, tales como un medicamento para el tratamiento o la prevención de tuberculosis (TB).

Un aspecto adicional de la invención se refiere a un método para inducir una respuesta inmunitaria en un sujeto, que comprende la administración de una proteína Rv3616c modificada.

Un aspecto adicional de la invención se refiere a un método para el tratamiento, la mejora, o la prevención de tuberculosis (TB), que comprende la administración de una cantidad segura y efectiva de una proteína Rv3616c modificada a un sujeto que lo necesite, en donde este polipéptido induce una respuesta inmunitaria. En un aspecto adicional, el método comprende además inducir una respuesta inmunitaria contra *Mycobacterium tuberculosis*.

Un aspecto adicional de la invención se refiere a un método para el tratamiento, la mejora, el retardo, o la prevención de la reactivación de tuberculosis, que comprende la administración de una cantidad efectiva de una proteína Rv3616c modificada a un sujeto que lo necesite, en donde este polipéptido induce una respuesta inmunitaria. En un aspecto adicional, el método comprende además inducir una respuesta inmunitaria contra *Mycobacterium tuberculosis*.

El uso de una proteína Rv3616c modificada en la elaboración de un medicamento para el tratamiento, la mejora, o la prevención de tuberculosis (TB), representa otro aspecto de la invención.

La presente invención proporciona un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada. También se proporciona un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada para utilizarse como un medicamento, tal como un medicamento para el tratamiento, la mejora, o la prevención de tuberculosis (TB).

Un aspecto adicional de la invención se refiere a un método para inducir una respuesta inmunitaria en un sujeto, que comprende la administración de un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada.

Un aspecto adicional de la invención se refiere a un método para el tratamiento, la mejora, o la prevención de tuberculosis (TB), que comprende la administración de una cantidad segura y efectiva de un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada a un sujeto que lo necesite, en donde este polinucleótido induce una respuesta inmunitaria. En un aspecto adicional, la presente invención proporciona un método para inducir una respuesta inmunitaria contra *Mycobacterium tuberculosis*.

Un aspecto adicional de la invención se refiere a un método para el tratamiento, la mejora, el retardo, o la prevención de la reactivación de tuberculosis, que comprende la administración de una cantidad efectiva de un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada a un sujeto que lo necesite, en donde este polipéptido induce una respuesta inmunitaria. En un aspecto adicional, el método comprende además inducir una respuesta inmunitaria contra *Mycobacterium tuberculosis*.

El uso de un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una proteína Rv3616c modificada en la elaboración de un medicamento para el tratamiento, la mejora, o la prevención de tuberculosis (TB), representa otro aspecto de la invención.

Adicionalmente, se proporciona una composición farmacéutica, que comprende:

(a) una proteína Rv3616c modificada; o
(b) un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada;
y
(c) un vehículo o excipiente farmacéuticamente aceptable.

Además, se proporciona una composición inmunogénica, que comprende:

(a) una proteína Rv3616c modificada; o
(b) un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada;
y
(c) un potenciador no específico de la respuesta inmunitaria.

También se proporciona un vector de expresión, que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada.

Las células huésped, transformadas con este vector de expresión, forman un aspecto adicional de la invención. Adicionalmente, se proporciona una célula huésped que expresa de una manera recombinante una proteína Rv3616c modificada.

Además, se proporciona un método para la producción de una proteína Rv3616c modificada; comprendiendo este método la etapa de expresar de una manera recombinante el polipéptido dentro de una célula huésped.

También se proporcionan kits de diagnóstico, que comprenden:

(a) una proteína Rv3616c modificada;
(b) un aparato suficiente para poner en contacto la proteína Rv3616c modificada con una muestra (por ejemplo, sangre entera, o de una manera más adecuada, células mononucleares de sangre periférica (PBMC)) a partir de un individuo; y
(c) medios para cuantificar la respuesta de células T de la muestra.

Otro aspecto de la invención se refiere a un kit de diagnóstico, que comprende:

(a) una proteína Rv3616c modificada; y
(b) un aparato suficiente para poner en contacto la proteína Rv3616c modificada con las células dérmicas de un paciente.

Un aspecto adicional de la invención se refiere a un método para detectar la infección por *Mycobacterium tuberculosis* en un sujeto, que comprende:

(a) poner en contacto una muestra a partir del sujeto mencionado, con una proteína Rv3616c modificada; y
(b) detectar, en la muestra biológica, la presencia de anticuerpos que se enlacen a la proteína Rv3616c modificada.

La invención también proporciona un kit de diagnóstico, que comprende:

(a) una proteína Rv3616c modificada, cuya proteína se inmoviliza opcionalmente sobre un soporte sólido; y
(b) un reactivo de detección.

En una realización, el sujeto que recibe una proteína Rv3616c modificada, un polinucleótido o una composición de acuerdo con la invención, puede tener tuberculosis activa (por ejemplo, la infección activa por *M. tuberculosis*). En una segunda realización, el sujeto puede tener tuberculosis latente (por ejemplo, la infección latente por *M. tuberculosis*). En una tercera realización, el sujeto puede estar libre de tuberculosis (por ejemplo, libre de la infección por *M. tuberculosis*).

Un sujeto que recibe una proteína Rv3616c modificada, un polinucleótido o una composición de acuerdo con la invención, puede haber sido vacunado previamente para tuberculosis (por ejemplo, puede haber sido vacunado contra la infección por *M. tuberculosis*), tal como puede haber sido vacunado con un Bacillus Calmette-Guerin (BCG). De una manera alternativa, un sujeto que recibe un polipéptido, polinucleótido o una composición de la invención, puede no haber sido anteriormente vacunado para tuberculosis (por ejemplo, no vacunado contra la infección por *M. tuberculosis*), tal como que no ha sido vacunado con un Bacillus Calmette-Guerin (BCG).

Una proteína Rv3616c modificada, un polinucleótido o una composición de acuerdo con la invención, se puede proporcionar para el propósito de:

- tratar la tuberculosis activa;
- prevenir la tuberculosis activa (tal como mediante su administración a un sujeto que no esté infectado, o de una manera alternativa un sujeto que tenga la infección latente) ;
- tratar la tuberculosis latente;
- prevenir la tuberculosis latente; o
- prevenir o retardar la reactivación de tuberculosis (en especial la demora de la reactivación de tuberculosis (TB), por ejemplo, por un período de meses, años, o incluso indefinidamente).

También se proporciona un método para el tratamiento de tuberculosis (TB) latente, que comprende las etapas de:

(i) identificar a un sujeto por tener una infección por tuberculosis (TB) latente (por ejemplo, mediante ensayos

basados en PPD o en células T) ; y
(ii) administrar a este sujeto, una cantidad segura y efectiva de una proteína Rv3616c modificada, o de un polinucleótido que codifique una proteína Rv3616c modificada (tal como en la forma de una composición farmacéutica o de una composición inmunogénica).

También se proporciona el uso de un polipéptido de la presente invención, en la elaboración de un kit de diagnóstico para la identificación de tuberculosis (TB) (por ejemplo, tuberculosis (TB) latente) en un sujeto de prueba.

El término “especie de *Mycobacterium* del complejo de tuberculosis” incluye las especies tradicionalmente consideradas como causantes de la enfermedad de tuberculosis, así como las especies medio-ambientales y oportunistas de *Mycobacterium* que provocan tuberculosis y enfermedad pulmonar en los pacientes inmuno-comprometidos, tales como los pacientes con SIDA, por ejemplo, *M. tuberculosis*, M. bovis, o M. africanum, BCG, M. avium, M. intracellulare, M. celatum, M. genavense, M. haemophilum, M. kansasii, M. simiae, M. vaccae, M. fortuitum, y M. scrofulaceum (véase, por ejemplo, Harrison’s Principles of Internal Medicine, Capítulo 150, páginas 953-966 (16a Edición, Braunwald y col., Editores, 2005). La presente invención se refiere en particular a la infección con *M. tuberculosis*.

El término “infección activa” se refiere a una infección (por ejemplo, infección por *M. tuberculosis*) con los síntomas y/o lesiones evidentes de la enfermedad (de una manera adecuada con los síntomas evidentes de la enfermedad).

Los términos “infección inactiva”, “infección en reposo" o "infección latente” se refieren a una infección (por ejemplo, a una infección por *M. tuberculosis*) sin los síntomas y/o lesiones evidentes de la enfermedad (de una manera adecuada sin los síntomas evidentes de la enfermedad). Un sujeto con la infección latente de una manera adecuada será uno que se pruebe positivo para la infección (por ejemplo, mediante ensayos basados en PPD o en células T), pero el cual no haya demostrado los síntomas y/o lesiones de la enfermedad que estén asociados con una infección activa.

El término “tuberculosis primaria” se refiere a la enfermedad clínica (por ejemplo, la manifestación de los síntomas de la enfermedad) directamente en seguida de la infección (por ejemplo, de la infección por *M. tuberculosis*). Véase, Harrison’s Principles of Internal Medicine, Capítulo 150, páginas 953-966 (16a Edición, Braunwald y col., Editores, 2005).

Los términos “tuberculosis secundaria" o "tuberculosis post-primaria” se refieren a la reactivación de una infección en reposo, inactiva, o latente (por ejemplo, de una infección por *M. tuberculosis*). Véase, Harrison’s Principles of Internal Medicine, Capítulo 150, páginas 953-966 (16a Edición, Braunwald y col., Editores, 2005).

El término “reactivación de tuberculosis” se refiere a la manifestación posterior de los síntomas de la enfermedad en un individuo que pruebe ser positivo para la infección (por ejemplo, en una prueba de tuberculina de piel, de una manera adecuada en un ensayo basado en células T *in vitro*), pero que no tenga los síntomas evidentes de la enfermedad. De una manera adecuada, el individuo no se habrá vuelto a exponer a la infección. La prueba de diagnóstico positiva indica que el individuo está infectado; sin embargo, el individuo puede o no haber manifestado anteriormente los síntomas evidentes de la enfermedad activa, que hubieran sido tratados suficientemente para llevar la tuberculosis hasta un estado inactivo o latente. Se reconocerá que los métodos para la prevención, el retardo, o el tratamiento de la reactivación de tuberculosis, se pueden iniciar en un individuo que manifieste los síntomas de la enfermedad activa.

El término tuberculosis “resistente a fármacos” se refiere a una infección (por ejemplo, a una infección por *M. tuberculosis*), en donde la cepa infecciosa no se mantiene estática o aniquilada (es decir, es resistente a) uno o más de los denominados como agentes quimioterapéuticos "de la línea frontal" efectivos en el tratamiento de tuberculosis (por ejemplo, isoniazida, rifampina, etambutol, estreptomicina y pirazinamida).

El término tuberculosis “resistente a múltiples fármacos” se refiere a una infección (por ejemplo, a una infección por *M. tuberculosis*), en donde la cepa infecciosa es resistente a dos o más de los agentes quimioterapéuticos "de la línea frontal" efectivos en el tratamiento de tuberculosis.

Un “agente quimioterapéutico” se refiere a un agente farmacológico conocido y usado en la técnica, para tratar la tuberculosis (por ejemplo, la infección por *M. tuberculosis*). Los agentes farmacológicos ejemplificados utilizados para tratar la tuberculosis incluyen, pero no se limitan a, amicacina, ácido amino-salicílico, capreomicina, cicloserina, etambutol, etionamida, isoniazida, kanamicina, pirazinamida, rifamicinas (es decir, rifampina, rifapentina y rifabutina), estreptomicina, ofloxacina, ciprofloxacina, claritromicina, azitromicina y fluoro-quinolonas. Los agentes quimioterapéuticos "de primera línea" o "de línea frontal" utilizados para tratar la tuberculosis que no es resistente a fármacos incluyen isoniazida, rifampina, etambutol, estreptomicina y pirazinamida. Los agentes quimioterapéuticos "de segunda línea" utilizados para tratar la tuberculosis que ha demostrado resistencia al fármaco para uno o más fármacos "de primera línea" incluyen ofloxacina, ciprofloxacina, etionamida, ácido amino-salicílico, cicloserina, amicacina, kanamicina y capreomicina. Estos agentes farmacológicos son revisados en el Capítulo 48 de Goodman and Gilman’s The Pharmacological Basis of Therapeutics, Hardman y Limbird, Editores, 2001.

Los términos “polipéptido”, “péptido" y "proteína" se utilizan indistintamente en la presente para referirse a un

polímero de residuos de aminoácidos. Los aminoácidos que se presentan naturalmente son aquéllos codificados por el código genético, así como aquellos aminoácidos que se modifican posteriormente, por ejemplo, hidroxi-prolina, γ-carboxi-glutamato, y O-fosfoserina. De una manera adecuada, un polipéptido de acuerdo con la presente invención consistirá solamente en residuos de aminoácidos que se presentan naturalmente, en especial aquellos aminoácidos codificados por el código genético.

"Ácido nucleico" se refiere a los desoxi-ribonucleótidos o ribonucleótidos y polímeros de los mismos, ya sea en la forma de una sola cadena o bien de doble cadena. El término "ácido nucleico" se utiliza de una manera intercambiable con gen, ADNc, ARNm, oligonucleótido, y polinucleótido.

Los aminoácidos pueden ser referidos en la presente mediante ya sea sus símbolos de tres letras comúnmente conocidos, o bien mediante los símbolos de una letra recomendados por la IUPAC-IUB Biochemical Nomenclature Commission. Los nucleótidos, de la misma manera, pueden ser referidos mediante sus códigos de una sola letra comúnmente aceptados.

El término 'secuencia de proteína Rv3616c', como se utiliza en la presente, significa la secuencia del polipéptido de Rv3616c proporcionada en la SEQ ID NO: 1 o un homólogo de la misma, a partir de una especie de *Mycobacterium* del complejo de tuberculosis, por ejemplo, una especie tal como *M. tuberculosis*, M. bovis, o M. africanum, o una especie de *Mycobacterium* que sea medio-ambiental y oportunista y que provoque infecciones oportunistas, tales como infecciones de pulmón en los huéspedes inmuno-comprometidos (por ejemplo, en los pacientes con SIDA), por ejemplo, BCG, M. avium, M. intracellulare, M. celatum, M. genavense, M. haemophilum, M. kansasii, M. simiae, M. vaccae, M. fortuitum, y M. scrofulaceum (véase, por ejemplo, Harrison's Principles of Internal Medicine, Capítulo 150, páginas 953-966, 16a Edición, Braunwald y col., Editores, 2005).

Con el fin de asegurar un alto índice de eficacia entre los huéspedes vacunados, los componentes de una vacuna deben estar bien conservados entre las cepas de significado clínico. De una manera adecuada, la proteína Rv3616c se deriva a partir de *M. tuberculosis* H37Rv (es decir, la secuencia de polipéptido proporcionada en la SEQ ID NO: 1) o un homólogo de la misma a partir de otra cepa de *M. tuberculosis* (tal como las cepas CDC1551, F11, Haarlem A y C). Las cepas de *M. tuberculosis* que están asociadas con resistencia a fármacos (por ejemplo, MDR o en especial XDR) son una base particularmente valiosa para la secuencia de proteína Rv3616c de tipo silvestre. Las cepas de interés incluyen:

CDC1551 - cepa transmisible y virulenta.
Familia Haarlem (tal como Haarlem A) - Cepas resistentes al fármaco que se encuentran en las poblaciones humanas populosas. Los miembros de la familia Haarlem de cepas de *M. tuberculosis* se han encontrado en muchas partes del mundo. El primer representante de la familia fue descubierto en Haarlem, Holanda.
KZN4207 - Aislado sensible a fármacos a partir de los pacientes de KwaZulu-Natal, Sudáfrica.
KZN1435 - Aislado resistente a múltiples fármacos (MDR) a partir de los pacientes de KwaZulu-Natal, Sudáfrica.
KZN605 - Aislado extensamente resistente a fármacos (XDR) a partir de los pacientes de KwaZulu-Natal, Sudáfrica.
C - Altamente transmitida en la Ciudad de Nueva York. En un estudio, se encontró que esta cepa es más común entre los usuarios de fármacos en inyección, y es resistente a los intermediarios de nitrógeno reactivo (Friedman y col., J. Infect. Dis. 1997, 176(2) : 478-84)
94_M4241A - aislada en San Francisco en 1994 a partir de un paciente nacido en China. Esta cepa fue anteriormente analizada mediante el análisis de supresión genómica (Gagneux y col., PNAS 2006, 103(8) : 2869-2873).
02_1987 - aislada en San Francisco en 2002 a partir de un paciente nacido en Corea del Sur. Esta cepa fue anteriormente analizada mediante el análisis de supresión genómica (Gagneux y col., PNAS 2006, 103(8) : 2869-2873).
T92 - aislada en San Francisco en 1999 a partir de un paciente nacido en Las Filipinas. Esta cepa fue publicada en Hirsh y col., PNAS 2004, 101: 4871–4876).
T85 - aislada en San Francisco en 1998 a partir de un paciente nacido en China. Esta cepa fue publicada en Hirsh y col., PNAS 2004, 101: 4871–4876).
EAS054 - aislada en San Francisco en 1993 a partir de un paciente nacido en India. Esta cepa fue anteriormente analizada mediante el análisis de supresión genómica (Gagneux y col., PNAS 2006, 103(8) : 2869-2873).
Gagneux y col., PNAS 2006, 103(8) : 2869-2873 y Herbert y col., Infect. Immun. 2007, 75(12) : 5798-5805 proporcionan valiosos antecedentes sobre la gama de cepas de *M. tuberculosis* que se sabe que existen.
De una manera más adecuada, la proteína Rv3616c se selecciona a partir de la secuencia de polipéptidos proporcionada en las SEQ ID NOs: 1 y 3-7, en particular en las SEQ ID NOs: 1 y 3-6, tal como en la SEQ ID NO: 1. En la Figura 15 se proporciona una alineación de las SEQ ID NOs: 1 y 3-7.
Las proteínas Rv3616c modificadas de un interés particular son aquéllas que comprenden (por ejemplo, que consisten en) las SEQ ID NOs: 161-169.
Los polinucleótidos de un interés particular son aquéllos derivados a partir de las secuencias de tipo silvestre correspondientes a las cepas de *M. tuberculosis* discutidas anteriormente, tales como aquéllas derivadas a partir de la SEQ ID NO: 2 o su SEQ ID No: 160 de codones optimizados de E. coli relacionada.

COMBINACIONES

Una secuencia que contiene las proteínas Rv3616c modificadas (o los polinucleótidos asociados) de la presente invención puede comprender además otros componentes diseñados para potenciar su inmunogenicidad o para mejorar estos antígenos en otros aspectos. Por ejemplo, se puede facilitar un mejor aislamiento de los antígenos del polipéptido a través de la adición de un estiramiento de residuos de histidina (comúnmente conocido como una marca-His) hacia un extremo del antígeno.

El término “marca-His” se refiere a un cordón de residuos de histidina, típicamente de seis residuos, que se insertan dentro de la secuencia de referencia. Para minimizar la alteración de la actividad asociada con la secuencia de referencia, típicamente se inserta una marca-His en el término-N, usualmente inmediatamente después del residuo de metionina iniciador, o de otra manera, en el término C. Normalmente son heterólogas para la secuencia nativa, pero se incorporan debido a que facilitan el aislamiento al mejorar en enlace de la proteína a las resinas de la cromatografía de afinidad con metal inmovilizado (IMAC). Hablando en términos generales, la presencia o ausencia de una marca-His no es significativa desde el punto de vista de provocar una respuesta inmunitaria deseable contra la proteína de referencia. Sin embargo, con el objeto de evitar el riesgo de una reacción adversa contra la marca-His misma, se considera mejor minimizar la longitud de la marca-His, por ejemplo, hasta cuatro o menos residuos, en particular hasta dos residuos, o excluir el uso de una marca-His enteramente.

Con el fin de mejorar la magnitud y/o amplitud de la respuesta inmunitaria provocada, se pueden preparar composiciones, polipéptidos (y ácidos nucleicos que los codifican) que comprendan múltiples secuencias de Rv3616c modificadas y/o polipéptidos heterólogos adicionales o los polinucleótidos que los codifiquen, a partir de las especies de *Mycobacterium* (en particular *M. tuberculosis*).

Un experto en la técnica reconocerá que, cuando se utilizan un número de componentes en combinación, se puede variar la presentación precisa. Por ejemplo, se podrían presentar un componente de la secuencia de Rv3616c modificada y una copia adicional del antígeno o un componente de antígeno heterólogo adicional:

(1) como dos componentes de polipéptido individuales;
(2) como una proteína de fusión que comprenda ambos componentes del polipéptido;
(3) como un polipéptido y un componente de polinucleótido;
(4) como dos componentes de polinucleótido individuales;
(5) como un solo polinucleótido que codifique dos componentes de polipéptido individuales; o
(6) como un solo polinucleótido que codifique una proteína de fusión que comprenda ambos componentes del polipéptido.

Esta flexibilidad se aplica igualmente a las situaciones en donde se utilizan tres o más componentes en combinación. Sin embargo, para mayor conveniencia, con frecuencia es deseable que cuando estén presentes un número de componentes, estén contenidos dentro de una sola proteína de fusión o dentro de un polinucleótido que codifique una sola proteína de fusión. En una realización de la invención, todos los componentes de antígeno se proporcionan como polipéptidos (por ejemplo, dentro de una sola proteína de fusión). En una realización alternativa de la invención, todos los componentes de antígeno se proporcionan como polinucleótidos (por ejemplo, un solo polinucleótido, tal como uno que codifique una sola proteína de fusión).

El término “heterólogas”, cuando se utiliza con referencia a porciones de un ácido nucleico, indica que el ácido nucleico comprende dos o más subsecuencias que no se encuentran en la misma relación una con la otra en la naturaleza. Por ejemplo, el ácido nucleico típicamente se produce de una manera recombinante, que tiene dos o más secuencias a partir de genes no relacionados configurados para la elaboración de un nuevo ácido nucleico funcional, por ejemplo, un promotor a partir de una fuente, y una región codificante a partir de otra fuente. De una manera similar, una proteína heteróloga indica que la proteína comprende dos o más subsecuencias que no se encuentran en la misma relación una con la otra en la naturaleza (por ejemplo, una proteína de fusión).

“Polipéptido de fusión" o "proteína de fusión" se refiere a una proteína que tiene al menos dos polipéptidos heterólogos (por ejemplo, al menos dos polipéptidos de *Mycobacterium* sp.) covalentemente unidos, ya sea directamente o bien por medio de un enlazador de aminoácidos. Los polipéptidos que forman la proteína de fusión típicamente enlazan el extremo C al término-N, aunque también pueden enlazar el extremo C al término-C, el extremo N al término-N, o el extremo N al término-C. Los polipéptidos de la proteína de fusión pueden estar en cualquier orden. Este término también se refiere a las variantes conservadoramente modificadas, las variantes polimórficas, los alelos, los mutantes, los fragmentos inmunogénicos, y los homólogos inter-especies de los antígenos que conforman la proteína de fusión. Los antígenos de *Mycobacterium tuberculosis* se describen en Cole y col., Nature 393: 537 (1998), el cual da a conocer el genoma entero de *Mycobacterium tuberculosis*. Se pueden identificar antígenos a partir de otras especies de *Mycobacterium* que correspondan a los antígenos de *M. tuberculosis*, por ejemplo, utilizando algoritmos de comparación de secuencias, como se describe en la presente, u otros métodos conocidos por los expertos en este campo, por ejemplo, ensayos de hibridación y ensayos de enlace de anticuerpos.

El término “fusionado” se refiere al enlace covalente entre dos polipéptidos en una proteína de fusión. Los polipéptidos típicamente se unen por medio de un enlace peptídico, ya sea directamente uno al otro, o bien por medio de un enlazador de aminoácidos. Opcionalmente, los péptidos se pueden unir por medio de enlaces

covalentes no peptídicos conocidos por los expertos en este campo.

Los antígenos de *M. tuberculosis* de ejemplo que se pueden combinar con una secuencia de Rv3616c modificada incluyen uno o más (por ejemplo, de 1 a 5, tal como de 1 a 3, en particular 1) de los siguientes:

(i) Mtb8.4 (también conocido como DPV y Rv1174c), la secuencia de polipéptido que se describe en la SEQ ID NO: 102 de la Publicación Internacional Número WO97/09428 (ADNc de la SEQ ID NO: 101), y en Coler y col., Journal of Immunology, 1998, 161: 2356-2364. Es de un interés particular la secuencia de Mtb8.4 madura, la cual está ausente del péptido de señal líder (es decir, de los residuos de aminoácidos 15-96 a partir de la SEQ ID NO: 102 de la Publicación Internacional Número WO97/09428). La secuencia del polipéptido de longitud completa para Mtb8.4 se muestra de la SEQ ID NO: 8;

(ii) Mtb9.8 (también conocido como MSL y Rv0287), la secuencia de polipéptido que se describe en la SEQ ID NO: 109 de la Publicación Internacional Número WO98/53075 (los fragmentos de MSL se dan a conocer en las SEQ ID NOs: 110-124 de la Publicación Internacional Número WO98/53075, siendo de un interés particular las SEQ ID NOs: 119 y 120), y también en Coler y col., Vaccine, 2009, 27: 223-233 (en particular los fragmentos reactivos mostrados en la Figura 2 del mismo). La secuencia del polipéptido de longitud completa para Mtb9.8 se muestra de la SEQ ID NO: 9;

(iii) Mtb9.9 (también conocido como Mtb9.9A, MTI, MTI-A y Rv1793), la secuencia de polipéptido que se describe en la SEQ ID NO: 19 de la Publicación Internacional Número WO98/53075, y en Alderson y col., Journal of Experimental Medicine, 2000, 7: 551-559 (los fragmentos de MTI se dan a conocer en las SEQ ID NOs: 17 y 51-66 de la Publicación Internacional Número WO98/ 53075, siendo de un interés particular las SEQ ID NOs: 17, 51, 52, 53, 56 y 62-65). Un número de las variantes de polipéptidos de MTI se describen en las SEQ ID NOs: 21, 23, 25, 27, 29 y 31 de la Publicación Internacional Número WO98/53075, y en Alderson y col., Journal of Experimental Medicine, 2000, 7: 551-559. La secuencia del polipéptido de longitud completa para Mtb9.9 se muestra de la SEQ ID NO: 10;

(iv) Ra12 (también conocido como el antígeno C-terminal de Mtb32A), la secuencia de polipéptido que se describe en la SEQ ID NO: 10 de la Publicación Internacional Número WO01/98460, y en Skeiky y col., Journal of Immunology, 2004, 172: 7618-7682. La secuencia del polipéptido de longitud completa para Ra12 se muestra de la SEQ ID NO: 11;

(v) Ra35 (también conocido como el antígeno N-terminal de Mtb32A), la secuencia de polipéptido que se describe en la SEQ ID NO: 8 de la Publicación Internacional Número WO01/98460, y en Skeiky y col., Journal of Immunology, 2004, 172: 7618-7682. La secuencia del polipéptido de longitud completa para Ra35 se muestra de la SEQ ID NO: 12;

(vi) TbH9 (también conocido como Mtb39, Mtb39A, TbH9FL y Rv1196), la secuencia de polipéptido que se describe en la SEQ ID NO: 107 de la Publicación Internacional Número WO97/09428, y también en Dillon y col., Infection and Immunity, 1999, 67(6) : 2941-2950, y en Skeiky y col., Journal of Immunology, 2004, 172: 7618-7682. La secuencia del polipéptido de longitud completa para TbH9 se muestra de la SEQ ID NO: 13;

(vii) Mtb41 (también conocido como MTCC2 y Rv0915c), la secuencia de polipéptido que se describe en la SEQ ID NO: 142 de la Publicación Internacional Número WO98/53075 (ADNc de la SEQ ID NO: 140), y en Skeiky y col., Journal of Immunology, 2000, 165: 7140-7149. La secuencia del polipéptido de longitud completa para Mtb41 se muestra de la SEQ ID NO: 14;

(viii) ESAT-6 (también conocido como esxA y Rv3875), la secuencia de polipéptido que se describe en la SEQ ID NO: 103 de la Publicación Internacional Número WO97/09428 (ADNc de la SEQ ID NO: 104), y en Sorensen y col. Infection and Immunity, 1995, 63(5) : 1710-1717. La secuencia del polipéptido de longitud completa para ESAT-6 se muestra de la SEQ ID NO: 15.L. Sorensen, S. Nagai, G. Houen, P. Andersen and A.B. Andersen, Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*, Infect Immun 63 (1995) (5), pp. 1710–1717;

(ix) Antígenos complejos Ag85 (por ejemplo, Ag85A, también conocido como fbpA y Rv3804c; o Ag85B, también conocido como fbpB y Rv1886c), los cuales se discuten, por ejemplo, en Content y col., Infection and Immunity, 1991, 59: 3205-3212, y en Huygen y col., Nature Medicine, 1996, 2(8) : 893-898. La secuencia del polipéptido de longitud completa para Ag85A se muestra de la SEQ ID NO: 16 (siendo de un interés particular la proteína madura de los residuos 43-338, es decir, que carece del péptido de señal). La secuencia del polipéptido de longitud completa para Ag85B se muestra de la SEQ ID NO: 17 (siendo de un interés particular la proteína madura de los residuos 41-325, es decir, que carece del péptido de señal) ;

(x) Alfa-cristalina (también conocida como hspX y Rv2031c), la cual se describe en Verbon y col., Journal of Bacteriology, 1992, 174: 1352-1359, y en Friscia y col., Clinical and Experimental Immunology, 1995, 102: 53-57 (son de un interés particular los fragmentos correspondientes a los residuos 71-91, 21-40, 91-110 y 111-130). La secuencia del polipéptido de longitud completa para la alfa-cristalina se muestra de la SEQ ID NO: 18;

(xi) Mpt64 (también conocido como Rv1980c), el cual se describe en Roche y col., Scandinavian Journal of Immunology, 1996, 43: 662-670. La secuencia del polipéptido de longitud completa para MPT64 se muestra de la SEQ ID NO: 19 (siendo de un interés particular la proteína madura de los residuos 24-228, es decir, que carece del péptido de señal) :

(xii) Mtb32A, la secuencia de polipéptido que se describe en la SEQ ID NO: 2 (longitud completa), y los residuos 8-330 de la SEQ ID NO: 4 (madura) de la Publicación Internacional Número WO01/ 98460, en especial las variantes que tienen al menos uno de los triads catalíticos mutados (por ejemplo, el residuo de serina catalítico, el cual, por ejemplo, se puede mutar hasta alanina). La secuencia del polipéptido de longitud completa para Mtb32A se muestra de la SEQ ID NO: 20. La forma madura de Mtb32A que tiene una mutación de Ser/Ala se muestra de la SEQ ID NO: 21;

(xiii) TB10.4, la secuencia del polipéptido de longitud completa para TB10.4 se muestra de la SEQ ID NO: 22; .L. Sorensen, S. Nagai, G. Houen, P. Andersen and A.B. Andersen, Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*, Infect Immun 63 (1995) (5), pp. 1710–1717;

(xiv) Rv1753c, la secuencia del polipéptido de longitud completa para Rv1753c a partir de *Mycobacterium tuberculosis* H37Rv se muestra de la SEQ ID NO: 157;

(xv) Rv2386c, la secuencia del polipéptido de longitud completa para Rv2386c a partir de *Mycobacterium tuberculosis* H37Rv se muestra de la SEQ ID NO: 158; y/o

(xvi) Rv2707c, la secuencia del polipéptido de longitud completa para Rv2707c a partir de *Mycobacterium tuberculosis* H37Rv se muestra de la SEQ ID NO: 159;

o las combinaciones de los mismos, tales como:

(a) una combinación de los componentes Ra12, TbH9 y Ra35, por ejemplo, en la forma de una proteína de fusión, tal como Mtb72f. La secuencia de polipéptido de Mtb72f se describe en la SEQ ID NO: 6 de la Publicación Internacional Número WO2006/117240 (ADNc de la SEQ ID NO: 5), y en Skeiky y col., Journal of Immunology, 2004, 172: 7618-7682 (en donde incorpora una marca-His opcional para ayudar a la purificación, cuando se utiliza en la presente invención, de una manera adecuada, Mtb72f está ausente de los residuos opcionales de histidina). La secuencia de polipéptido para Mtb72f se muestra de la SEQ ID NO: 23;

(b) una combinación de los componentes Ra12, TbH9 y Ra35 mutado en Ser/Ala (es decir, en donde el residuo de serina catalítico ha sido reemplazado con alanina), por ejemplo, en la forma de una proteína de fusión, tal como M72. La secuencia de polipéptido de M72 se describe en la SEQ ID NO: 4 de la Publicación Internacional Número WO2006/117240 (ADNc de la SEQ ID NO: 3), en donde incorpora una doble histidina opcional, para ayudar a la elaboración, cuando se utiliza en la presente invención M72 también puede incorporar una doble histidina, aunque, de una manera adecuada, M72 está ausente de la doble histidina opcional (es decir, los residuos 4-725 a partir de la SEQ ID NO: 4 de la Publicación Internacional Número WO2006/117240 son de un interés particular). La secuencia de polipéptido para M72 se muestra de la SEQ ID NO: 24;

(c) una combinación de los componentes Mtb8.4, Mtb9.8, Mtb9.9 y Mtb41, por ejemplo, en la forma de una proteína de fusión, tal como Mtb71f. La secuencia de polipéptido de Mtb71f se describe en la SEQ ID NO: 16 de la Publicación Internacional Número WO99/051748 (ADNc de la SEQ ID NO: 15), en donde incorpora una marca-His opcional para ayudar a la purificación, cuando se utiliza en la presente invención, de una manera adecuada, Mtb71f corresponde a los residuos de aminoácidos 9-710 de la SEQ ID NO: 16 a partir de la Publicación Internacional Número WO99/051748. La secuencia de polipéptido para Mtb71f se muestra de la SEQ ID NO: 25;

(d) una combinación de Mtb72f o M72 (de una manera adecuada sin residuos de histidina opcionales para ayudar a la expresión) con Mtb9.8 y Mtb9.9, por ejemplo, en una proteína de fusión. La secuencia de polipéptido para una fusión de M72-Mtb9.9-Mtb9.8 se muestra de la SEQ ID NO: 26 (fusión de M92), cuando se utilizan en la presente invención, la fusión de M72-Mtb9.9-Mtb9.8 puede incorporar opcionalmente una doble histidina en seguida del residuo de metionina iniciador, para ayudar a la elaboración;

(e) una combinación de Mtb72f o M72 (de una manera adecuada sin residuos de histidina opcionales para ayudar a la expresión) con Ag85B, por ejemplo, en una proteína de fusión, tal como Mtb103f. La secuencia de polipéptido de Mtb103f se describe en la SEQ ID NO: 18 de la Publicación Internacional Número WO03/ 070187 (ADNc de la SEQ ID NO: 10), en donde se incorpora una marca-His opcional para ayudar a la purificación, cuando se utiliza en la presente invención, de una manera adecuada, Mtb103f corresponde a los residuos de aminoácidos 8-1016 de la SEQ ID NO: 18 a partir de la Publicación Internacional Número WO03/070187. También es de un interés particular M103, es decir, Mtb103f que incorpora una mutación de Ser/Ala en el componente Ra35, cuando se utiliza en la presente invención, de una manera adecuada, M103 corresponde a los residuos de aminoácidos 8-1016 de la SEQ ID NO: 18 a partir de la Publicación Internacional Número WO03/070187, en donde el residuo de Ser en la posición 710 ha sido reemplazado con Ala. La secuencia de polipéptido para M103 se muestra de la SEQ ID NO: 27, cuando se utiliza en la presente invención, la fusión de

M72-Mtb9.9-Mtb9.8 puede incorporar opcionalmente una doble histidina en seguida del residuo de metionina iniciador, para ayudar a la elaboración;

(f) una combinación de Mtb72f o M72 (de una manera adecuada sin residuos de histidina opcionales para ayudar a la expresión) con Mtb41, por ejemplo, en una proteína de fusión, tal como Mtb114f. La secuencia de polipéptido de Mtb114f se describe en la SEQ ID NO: 16 de la Publicación Internacional Número WO03/ 070187 (ADNc de la SEQ ID NO: 9), en donde incorpora una marca-His opcional para ayudar a la purificación, cuando se utiliza en la presente invención, de una manera adecuada, Mtb114f corresponde a los residuos de aminoácidos 8-1154 de la SEQ ID NO: 16 a partir de la Publicación Internacional Número WO03/070187. También es de un interés particular M114, es decir, Mtb114f que incorpora una mutación de Ser/Ala en el componente Ra35, cuando se utiliza en la presente invención, de una manera adecuada, M114 corresponde a los residuos de aminoácidos 8-1154 de la SEQ ID NO: 16 a partir de la Publicación Internacional Número WO03/070187, en donde el residuo de Ser en la posición 710 ha sido reemplazado con Ala. La secuencia de polipéptido para M114 se muestra de la SEQ ID NO: 28, cuando se utiliza en la presente invención, la fusión de M72-Mtb9.9-Mtb9.8 puede incorporar opcionalmente una doble histidina en seguida del residuo de metionina iniciador, para ayudar a la elaboración;

(g) una combinación de los componentes Ag85B y ESAT-6, tal como en una fusión descrita en Doherty y col., Journal of Infectious Diseases, 2004, 190: 2146–2153; y/o

(h) una combinación de los componentes Ag85B y TB10.4, tal como en una fusión descrita en Dietrich y col., Journal of Immunology, 2005, 174(10) : 6332-6339, 190: 2146–2153.

Las combinaciones de un componente de una secuencia de Rv3616c modificada y un componente de Rv1753c, son de un interés particular. Obviamente estas combinaciones podrían contener opcionalmente otros componentes de antígeno adicionales (por ejemplo, un componente M72).

Otra combinación de interés comprende un componente de una secuencia de Rv3616c modificada y un componente M72.

Una combinación de interés adicional comprende un componente de una secuencia de Rv3616c modificada y un componente Rv2386c.

Otras combinaciones de interés incluyen aquéllas que comprenden un componente de una secuencia de Rv3616c modificada y un componente Rv2707c.

Una combinación de interés adicional comprende un componente de una secuencia de Rv3616c modificada y un componente de alfa-cristalina.

La persona experta reconocerá que las combinaciones no necesitan apoyarse en las secuencias específicas descritas anteriormente en (i) -(xvi) y (a) -(h), y que se pueden utilizar las variantes conservadoramente modificadas (por ejemplo, que tengan al menos el 70 % de identidad, tal como al menos el 80 % de identidad, en particular al menos el 90 % de identidad, y en especial al menos el 95 % de identidad), o los fragmentos inmunogénicos (por ejemplo, de al menos el 20 % del antígeno de longitud completa, tal como de al menos el 50 % del antígeno, en particular de al menos el 70 por ciento, y en especial de al menos el 80 por ciento) de las secuencias descritas para lograr el mismo efecto práctico.

Cada una de las secuencias de antígeno individuales anteriores también se da a conocer en Cole y col., Nature, 1998, 393: 537-544, y en Camus, Microbiology, 2002, 148: 2967-2973. El genoma de *M. tuberculosis* H37Rv está públicamente disponible, por ejemplo, en el sitio web Welcome Trust Sanger Institute (www.sanger.ac.uk/Projects/M_tuberculosis/), y en cualquier otra parte.

Muchos de los antígenos anteriores también se dan a conocer en las Solicitudes de Patente de los Estados Unidos de Norteamérica Números 08/523,435, 08/523,436, 08/658,800, 08/659,683, 08/818,111, 08/818,112, 08/942,341, 08/942,578, 08/858,998, 08/859,381, 09/056,556, 09/072,596, 09/072,967, 09/073,009, 09/073,010, 09/223,040, 09/287,849 y en las Solicitudes de Patente del TCP Números PCT/US98/10407, PCT/US98/10514, PCT/US99/03265, PCT/US99/03268, PCT/US99/07717, WO97/09428 y WO97/09429, WO98/16645, WO98/16646, cada una de las cuales se incorpora a la presente como referencia.

Las composiciones, polipéptidos, y ácidos nucleicos de la invención también pueden comprender polipéptidos adicionales a partir de otras fuentes. Por ejemplo, las composiciones y las proteínas de fusión de la invención pueden incluir polipéptidos o ácidos nucleicos que codifiquen los polipéptidos, en donde el polipéptido mejora la expresión del antígeno, por ejemplo, NS1, una proteína de virus de influenza (véanse, por ejemplo, las Publicaciones Internacionales Números WO99/40188 y WO93/04175).

Los ácidos nucleicos de la invención se pueden diseñar basándose en la preferencia de codones en una especie de elección, por ejemplo, en seres humanos (en el caso de la expresión in vivo) o en una bacteria particular (en el caso de la producción de polipéptidos). La SEQ ID NO: 160, por ejemplo, proporciona un polinucleótido de codones optimizados para la expresión de Rv3616c a partir de H37Rv en E. coli.

El componente de la secuencia de Rv3616c modificada también se puede administrar con uno o más agentes quimioterapéuticos efectivos contra la tuberculosis (por ejemplo, la infección por *M. tuberculosis*). Los ejemplos de estos agentes quimioterapéuticos incluyen, pero no se limitan a, amicacina, ácido amino-salicílico, capreomicina, cicloserina, etambutol, etionamida, isoniazida, kanamicina, pirazinamida, rifamicinas (es decir, rifampina, rifapentina y rifabutina), estreptomicina, ofloxacina, ciprofloxacina, claritromicina, azitromicina y fluoro-quinolonas. Esta quimioterapia es determinada por el juicio del médico tratante utilizando las combinaciones preferidas de fármacos. Los agentes quimioterapéuticos "de primera línea" utilizados para tratar la tuberculosis (por ejemplo, la infección por *M. tuberculosis*) que no es resistente a fármacos, incluyen isoniazida, rifampina, etambutol, estreptomicina y pirazinamida. Los agentes quimioterapéuticos "de segunda línea" utilizados para tratar la tuberculosis (por ejemplo, la infección por *M. tuberculosis*) que ha demostrado resistencia al fármaco para uno o más fármacos "de primera línea" incluyen ofloxacina, ciprofloxacina, etionamida, ácido amino-salicílico, cicloserina, amicacina, kanamicina y capreomicina.

Los agentes quimioterapéuticos convencionales se administran en términos generales durante un período relativamente largo (de aproximadamente 9 meses). La combinación de los agentes quimioterapéuticos convencionales con la administración de un componente de una secuencia de Rv3616c modificada de acuerdo con la presente invención, puede hacer posible que se reduzca el período de tratamiento quimioterapéutico (por ejemplo, hasta 8 meses, 7 meses, 6 meses, 5 meses, 4 meses, 3 meses, o menos) sin una disminución en la eficacia.

Es de un interés particular el uso de un componente de una secuencia de Rv3616c modificada en conjunto con Bacillus Calmette-Guerin (BCG). Por ejemplo, en la forma de un BCG modificado que exprese de una manera recombinante una proteína Rv3616c modificada. De una manera alternativa, el componente de la secuencia de Rv3616c modificada se puede utilizar para mejorar la respuesta de un sujeto a la vacunación con BCG, ya sea mediante la co-administración o mediante el refuerzo de una vacunación previa con BCG. Cuando se utiliza para potenciar la respuesta de un sujeto a la vacunación con BCG, el componente de la secuencia de Rv3616c modificada se puede proporcionar obviamente en la forma de un polipéptido o de un polinucleótido (opcionalmente en conjunto con componentes antigénicos adicionales, como se describe anteriormente).

La persona experta reconocerá que las combinaciones de los componentes no necesitan administrarse juntas, y se pueden aplicar: por separado o en combinación; al mismo tiempo, en secuencia o dentro de un período corto; a través de las mismas o a través de diferentes vías. No obstante, para mayor conveniencia en general es deseable (en donde los regímenes de administración sean compatibles) administrar una combinación de componentes como una sola composición.

Los polipéptidos, polinucleótidos y composiciones de la presente invención normalmente se administrarán a seres humanos, pero se puede esperar que sean efectivos en otros mamíferos, incluyendo mamíferos domésticos (por ejemplo, perros, gatos, conejos, ratas, ratones, cobayos, hámsters, chinchillas), y mamíferos agrícolas (por ejemplo, reses, cerdos, ovejas, cabras, caballos).

VARIANTES

Los epítopos de células T son estiramientos de aminoácidos contiguos cortos que son reconocidos por las células T (por ejemplo, las células T CD4+ o CD8+). La identificación de los epítopos de células T se puede lograr a través de los experimentos de mapeo de epítopos, los cuales son bien conocidos por la persona experta en la materia (véase, por ejemplo, Paul, Fundamental Immunology, 3a Edición, 243-247 (1993) ; Beißbarth y col., Bioinformatics, 2005, 21 (Suplemento 1) : i29-i37). De una manera alternativa, los epítopos se pueden predecir o mapear utilizando los planteamientos discutidos en los Ejemplos.

En una población exógama diversa, tal como en los seres humanos, los diferentes tipos de HLA significan que algunos epítopos particulares pueden no ser reconocidos por todos los miembros de la población. Como un resultado del involucramiento crucial de la respuesta de células T en la tuberculosis, con el fin de maximizar el nivel de reconocimiento y la escala de respuesta inmunitaria, una proteína Rv3616c modificada óptima es una que contiene la mayoría de (o de una manera adecuada todos) los epítopos de células T intactos.

“Variantes" o "variantes conservadoramente modificadas” se aplica a las secuencias tanto de aminoácidos como de ácidos nucleicos. Con respecto a las secuencias de ácidos nucleicos particulares, las variantes conservadoramente modificadas se refieren a los ácidos nucleicos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o en donde el ácido nucleico no codifica una secuencia de aminoácidos, para secuencias esencialmente idénticas.

Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier proteína dada. Por ejemplo, los codones GCA, GCC, GCG y GCU, codifican todos el aminoácido alanina. Por consiguiente, en cada posición en donde una alanina es especificada por un codón, el codón se puede alterar hasta cualquiera de los codones correspondientes descritos sin alterar el polipéptido codificado. Las variaciones de ácidos nucleicos conducen a las variantes “silenciosas" o "degeneradas”, las cuales son una especie de variaciones conservadoramente modificadas. Cada secuencia de ácido nucleico en la presente, que codifique un polipéptido, también describe cada posible variación silenciosa del ácido nucleico. Un experto en la materia

reconocerá que cada codón en un ácido nucleico (excepto AUG, el cual es ordinariamente el único codón para metionina, y TGG, el cual es ordinariamente el único codón para triptófano) puede ser modificado para proporcionar una molécula funcionalmente idéntica. De conformidad con lo anterior, cada variación silenciosa de un ácido nucleico que codifique un polipéptido es implícita en cada secuencia descrita.

Las variaciones no silenciosas son aquéllas que dan como resultado un cambio en la secuencia de aminoácidos codificada (ya sea a través de la sustitución, supresión o adición de los residuos de aminoácidos). Los expertos en este campo reconocerán que una secuencia de polinucleótido particular puede contener variaciones conservadoras silenciosas y no silenciosas.

Con respecto a las variantes de una secuencia de proteína, la persona experta reconocerá que las sustituciones, supresiones o adiciones individuales al polipéptido, que alteren, agreguen, o supriman un solo aminoácido o un pequeño porcentaje de aminoácidos, es una "variante conservadoramente modificada" en donde las alteraciones dan como resultado la sustitución de un aminoácido con un aminoácido funcionalmente similar o la sustitución/supresión/adición de los residuos que no tengan un impacto sustancial sobre la función biológica de la variante.

Las tablas de sustituciones conservadoras que proporcionan aminoácidos funcionalmente similares son bien conocidas en la materia. Estas variantes conservadoramente modificadas son en adición a, y no excluyen, las variantes polimórficas, los homólogos inter-especies, y los alelos de la invención.

En general, estas sustituciones conservadoras caerán dentro de una de las agrupaciones de aminoácidos especificadas más adelante, aunque, en algunas circunstancias son posibles otras sustituciones sin afectar sustancialmente a las propiedades inmunogénicas del antígeno. Los siguientes ocho grupos contienen cada uno, aminoácidos que son típicamente sustituciones conservadoras unos por otros:

1) Alanina (A), Glicina (G) ;
2) Ácido aspártico (D), ácido glutámico (E) ;
3) Asparagina (N), Glutamina (Q) ;
4) Arginina (R), Lisina (K) ;
5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V) ;
6) Fenilalanina (F), Tirosina (Y), Triptófano (W) ;
7) Serina (S), Treonina (T) ; y
8) Cisteína (C), Metionina (M) ;

(véase, por ejemplo, Creighton, Proteins, 1984).

De una manera adecuada, estas sustituciones no se presentan en la región de un epítopo, y, por consiguiente, no tienen un impacto significativo sobre las propiedades inmunogénicas del antígeno.

Las variantes de proteínas también pueden incluir aquéllas en donde se insertan aminoácidos adicionales, comparándose con la secuencia de referencia. De una manera adecuada, estas inserciones no se presentan en la región de un epítopo, y, por consiguiente, no tienen un impacto significativo sobre las propiedades inmunogénicas del antígeno. Un ejemplo de las inserciones incluye un estiramiento corto de los residuos de histidina (por ejemplo, de 2 a 6 residuos) para ayudar a la expresión y/o purificación del antígeno en cuestión.

Las variantes de proteínas incluyen aquéllas en donde se han suprimido aminoácidos, comparándose con la secuencia de referencia. De una manera adecuada, estas supresiones no se presentan en la región de un epítopo, y, por consiguiente, no tienen un impacto significativo sobre las propiedades inmunogénicas del antígeno.

La persona experta reconocerá que una variante de proteína particular puede comprender sustituciones, supresiones y adiciones (o cualquier combinación de las mismas).

Los términos "idéntica" o porcentaje de "identidad", en el contexto de dos o más secuencias de ácidos nucleicos o de polipéptidos, se refieren a dos o más secuencias o sub-secuencias que son iguales o que tienen un porcentaje especificado de residuos de aminoácidos o de nucleótidos que son iguales (es decir, el 70 % de identidad, opcionalmente el 75 por ciento, el 80 por ciento, el 85 por ciento, el 90 por ciento, el 95 por ciento, el 98 % o el 99 % de identidad sobre una región especificada), cuando se comparan y se alinean para una máxima correspondencia sobre una ventana de comparación, o una región designada, como se mide utilizando uno de los siguientes algoritmos de comparación de secuencias, o mediante alineación manual e inspección visual. Entonces se dice que estas secuencias son "sustancialmente idénticas". Esta definición también se refiere al cumplimiento de una secuencia de prueba. Opcionalmente, la identidad existe sobre una región que es de al menos aproximadamente 25 a aproximadamente 50 aminoácidos o nucleótidos de longitud, u opcionalmente sobre una región que es de 75 a 100 aminoácidos o nucleótidos de longitud. De una manera adecuada, la comparación se lleva a cabo sobre una ventana correspondiente a la longitud entera de la secuencia de referencia (opuestamente a la secuencia variante).

Para la comparación de secuencias, típicamente una secuencia actúa como una secuencia de referencia, con la cual se comparan las secuencias de prueba. Cuando se utiliza un algoritmo de comparación de secuencias, se introducen en una computadora las secuencias de prueba y de referencia, se designan las coordenadas de

subsecuencias, si es necesario, y se designan los parámetros del programa de algoritmos de secuencias. Se pueden utilizar los parámetros del programa por omisión, o se pueden designar parámetros alternativos. Entonces el algoritmo de comparación de secuencias calcula el porcentaje de identidad de secuencia para las secuencias de prueba en relación con la secuencia de referencia, basándose en los parámetros del programa.

Una "ventana de comparación", como se utiliza en la presente, hace referencia a un segmento en donde una secuencia se puede comparar con una secuencia de referencia del mismo número de posiciones contiguas después de que se alinean óptimamente las dos secuencias. Los métodos de alineación de secuencias para comparación son bien conocidos en este campo. La alineación óptima de secuencias para comparación se puede conducir, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman, Adv. Appl. Math., 2: 482 (1981), mediante el algoritmo de alineación de homología de Needleman y Wunsch, J. Mol. Biol., 48: 443 (1970), mediante el método de búsqueda de similitudes de Pearson y Lipman, Proc. Nat'l. Acad. Sci., EUA, 85: 2444 (1988), mediante las implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante alineación manual e inspección visual (véase, por ejemplo, Current Protocols in Molecular Biology (Ausubel y col., Editores, 1995, Suplemento)).

Un ejemplo de un algoritmo útil es PILEUP. PILEUP crea una alineación de múltiples secuencias a partir de un grupo de secuencias relacionadas utilizando alineaciones progresivas por pares, para mostrar la relación y el porcentaje de identidad de secuencias. También grafica un árbol o dendograma que muestra las relaciones agrupadas utilizadas para crear la alineación. PILEUP utiliza una simplificación del método de alineación progresiva de Feng y Doolittle, J. Mol. Evol., 35: 351-360 (1987). El método empleado es similar al método descrito por Higgins y Sharp, CABIOS 5: 151-153 (1989). El programa puede alinear hasta 300 secuencias, cada una de una longitud máxima de 5,000 nucleótidos o aminoácidos. El procedimiento de múltiples alineaciones empieza con la alineación por pares de las dos secuencias más similares, produciendo un racimo de dos secuencias alineadas. Este racimo se alinea entonces con la siguiente secuencia o racimo de secuencias alineadas más relacionadas. Dos racimos de secuencias se alinean mediante una simple extensión de la alineación por pares de dos secuencias individuales. La alineación final se logra mediante una serie de alineaciones progresivas por pares. El programa se ejecuta designando las secuencias específicas y sus coordenadas de aminoácidos o de nucleótidos para las regiones de comparación de secuencias, y designando los parámetros del programa. Utilizando PILEUP, una secuencia de referencia se compara con otras secuencias de prueba para determinar el porcentaje de relación de identidad de secuencias utilizando los siguientes parámetros: peso de hueco por omisión (3,00), peso de longitud de hueco por omisión (0,10), y huecos de extremo ponderados. PILEUP se puede obtener a partir del paquete de software de análisis de secuencias GCG, por ejemplo, versión 7.0 (Devereaux y col., Nuc. Acids Res., 12: 387-395 (1984).

Otro ejemplo de algoritmo que es adecuado para determinar el porcentaje de identidad de secuencias y la similitud de secuencias, son los algoritmos BLAST y BLAST 2.0, los cuales se describen en Altschul y col., Nuc. Acids Res., 25: 3389-3402 (1977), y Altschul y col., J. Mol. Biol., 215: 403-410 (1990), respectivamente. El software para llevar a cabo los análisis BLAST está públicamente disponible a través del National Center for Biotechnology Information (sitio web en www.ncbi.nlm.nih.gov/). Este algoritmo involucra primero identificar los pares de secuencias de alto puntaje (HSPs) mediante la identificación de palabras cortas de una longitud W en la secuencia pedida, las cuales concuerdan o satisfacen algún puntaje umbral de valor positivo T cuando se alinean con una palabra de la misma longitud en una secuencia de la base de datos. T es referido como el umbral de puntaje de palabra vecina (Altschul y col., supra). Estos impactos de palabra vecina iniciales actúan como semillas para iniciar las búsquedas con el fin de encontrar HPSs más largos que las contengan. Los impactos de palabras se extienden en ambas direcciones a lo largo de cada secuencia por tanto como se pueda incrementar el puntaje de alineación acumulativo. Los puntajes acumulativos se calculan utilizando, para las secuencias de nucleótido, los parámetros M (puntaje de recompensa para un par de residuos emparejados; siempre > 0), y N (puntaje de penalidad para los residuos mal emparejados; siempre < 0). Para las secuencias de aminoácidos, se utiliza una matriz de puntuación para calcular el puntaje acumulativo. La extensión de los impactos de palabras en cada dirección se detiene cuando: el puntaje de alineación acumulativo cae fuera por la cantidad X a partir de su máximo valor alcanzado; el puntaje acumulativo llega a cero o más abajo, debido a la acumulación de una o más alineaciones de residuos de puntuación negativa; o se llega al final de cualquier secuencia. Los parámetros W, T, y X del algoritmo BLAST determinan la sensibilidad y la velocidad de la alineación. El programa BLASTN (para las secuencias de nucleótidos) utiliza por omisión una longitud de palabra (W) de 11, una expectativa (E) de 10, M = 5, N = -4, y una comparación de ambas cadenas. Para las secuencias de aminoácidos, el programa BLASTP utiliza por omisión una longitud de palabra de 3, y una expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, Proc. Natl. Acad. Sci., EUA, 89: 10915 (1989)) alineaciones (B) de 50, expectativa (E) de 10, M = 5, N = -4, y una comparación de ambas cadenas.

El algoritmo BLAST también lleva a cabo un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, Proc. Nat'l. Acad. Sci., EUA, 90: 5873-5787 (1993)). Una medida de similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la que se presentaría un emparejamiento entre dos secuencias de nucleótidos o de aminoácidos por azar. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de prueba con el ácido nucleico de referencia es menor de aproximadamente 0,2, más preferiblemente menor de aproximadamente 0,01, y de una manera muy preferible menor de aproximadamente 0,001.

En cualquier caso, las variantes de un secuencia de polipéptido tendrán esencialmente la misma actividad que la secuencia de referencia (en el caso de los polinucleótidos, las secuencias de polinucleótidos variantes codificarán un polipéptido que tendrá esencialmente la misma actividad que la secuencia de referencia). Esencialmente la misma actividad significa al menos el 50 por ciento, de una manera adecuada, al menos el 75 por ciento, y en especial al menos el 90 % de actividad de la secuencia de referencia en un ensayo de reestimulación *in vitro* de células mononucleares de sangre periférica (PBMC) o de sangre entera con antígenos específicos (por ejemplo, la reestimulación durante un período de entre varias horas y hasta dos semanas, tal como hasta un día, de 1 día a 1 semana, o de 1 a 2 semanas), el cual mide la activación de las células por medio de linfoproliferación, producción de citocinas en el sobrenadante de cultivo (medida mediante ELISA, CBA, etc.), o caracterización de respuestas de células T y B mediante teñido intra- y extra-celular (por ejemplo, utilizando anticuerpos específicos para inmuno-marcadores, tales como CD3, CD4, CD8, IL2, TNFa, IFNg, CD40L, CD69, etc.), seguida por análisis con a citómetro de flujo. De una manera adecuada, por esencialmente la misma actividad significa al menos el 50 por ciento, de una manera adecuada, al menos el 75 por ciento, y en especial al menos el 90 % de actividad de la secuencia de referencia en un ensayo de proliferación de células T y/o de producción de IFN-gamma.

Como será entendido por los expertos en la materia, los polinucleótidos útiles en la presente invención pueden incluir secuencias genómicas, secuencias extra-genómicas y codificadas por plásmidos, y segmentos genéticos diseñados más pequeños que expresen, o que se puedan adaptar para expresar, proteínas, polipéptidos, péptidos y similares. Estos segmentos se pueden aislar naturalmente, o se pueden modificar sintéticamente mediante la mano del hombre.

Como será reconocido por el experto, los polinucleótidos pueden ser de una sola cadena (codificantes o antisentido) o de doble cadena, y pueden ser moléculas de ADN (genómico, ADNc, o sintético) o de ARN. Las moléculas de ARN incluyen moléculas de HnARN, que contienen intrones y que corresponden a una molécula de ADN de una manera de una a una, y moléculas de ARNm, que no contienen intrones. Puede haber, pero no necesita haber, secuencias codificantes o no codificantes adicionales presentes dentro de un polinucleótido de la presente invención, y un polinucleótido se puede enlazar, pero no necesita enlazarse, a otras moléculas y/o materiales de soporte.

Los polinucleótidos pueden comprender una secuencia nativa (es decir, una secuencia endógena que codifique un antígeno de *Mycobacterium* o una porción del mismo), o pueden comprender una variante, o un equivalente biológico o funcional de esta secuencia. Las variantes de polinucleótidos pueden contener una o más sustituciones, adiciones, supresiones y/o inserciones, de tal manera que no disminuye la inmunogenicidad del polipéptido codificado en relación con la proteína de referencia.

IDENTIFICACIÓN Y CARACTERIZACIÓN DE POLINUCLEÓTIDOS

Los polinucleótidos se pueden identificar, preparar y/o manipular utilizando cualquiera de una variedad de técnicas bien establecidas. Por ejemplo, un polinucleótido se puede identificar, como se describe con mayor detalle más adelante, mediante el rastreo de un microarreglo de ADNcs. Estos rastreos se pueden llevar a cabo, por ejemplo, utilizando un microarreglo Synteni (Palo Alto, CA) de acuerdo con las instrucciones del fabricante (y esencialmente como es descrito por Schena y col., Proc. Natl. Acad. Sci., EUA, 93: 10614-10619 (1996), y por Heller y col., Proc. Natl. Acad. Sci., EUA, 94: 2150-2155 (1997)). De una manera alternativa, los polinucleótidos se pueden amplificar a partir del ADNc preparado a partir de las células que expresen las proteínas descritas en la presente, tales como las células de *M. tuberculosis*. Estos polinucleótidos se pueden amplificar por medio de reacción en cadena de la polimerasa (PCR). Para este planteamiento, se pueden diseñar cebadores específicos de la secuencia basándose en las secuencias proporcionadas en la presente, y se pueden comprar o se pueden sintetizar.

Se puede utilizar una porción amplificada de un polinucleótido para aislar un gen de longitud completa a partir de una biblioteca adecuada (por ejemplo, una biblioteca de ADNc de *M. tuberculosis*) utilizando técnicas bien conocidas. Dentro de estas técnicas, una biblioteca (ADNc o genómico) se rastrea utilizando una o más sondas o cebadores de polinucleótido adecuados para la amplificación. De preferencia, una biblioteca se selecciona por tamaños para incluir las moléculas más grandes. También se pueden preferir las bibliotecas con cebadores aleatorios para identificar las regiones 5' y corriente arriba de los genes. Se prefieren las bibliotecas genómicas para obtener intrones y para extender las secuencias 5'.

Para las técnicas de hibridación, una secuencia parcial se puede marcar (por ejemplo, mediante traducción de muesca o marcado de extremo con 32P) utilizando técnicas bien conocidas. Entonces se rastrea en términos generales una biblioteca bacteriana o de bacteriófagos mediante filtros de hibridación que contengan colonias bacterianas desnaturalizadas (o prados que contengan placas de fagos) con la sonda marcada (véase Sambrook y col., Molecular Cloning: A Laboratory Manual (2000)). Las colonias o placas de hibridación se seleccionan y se expanden, y el ADN se aísla para su análisis adicional. Los clones de ADNc se pueden analizar para determinar la cantidad de secuencia adicional, por ejemplo, mediante reacción en cadena de la polimerasa (PCR), utilizando un cebador a partir de la secuencia parcial, y un cebador a partir del vector. Se pueden generar mapas de restricción y secuencias parciales para identificar uno o más clones traslapados. Entonces se puede determinar la secuencia completa utilizando las técnicas convencionales, las cuales pueden involucrar la generación de una serie de clones de supresión. Entonces se pueden ensamblar las secuencias traslapadas resultantes en una sola secuencia contigua. Se puede generar un ADNc de longitud completa mediante el ligamiento de los fragmentos adecuados,

utilizando técnicas bien conocidas.

De una manera alternativa, existen numerosas técnicas de amplificación para obtener una secuencia de codificación de longitud completa a partir de una secuencia de ADNc parcial. Dentro de estas técnicas, la amplificación se lleva a cabo en general por medio de reacción en cadena de la polimerasa (PCR). Se puede utilizar cualquiera de una variedad de kits comercialmente disponibles para llevar a cabo la etapa de amplificación. Los cebadores se pueden diseñar utilizando, por ejemplo, softwares bien conocidos en la materia. Los cebadores son de preferencia de 22 a 30 nucleótidos de longitud, tienen un contenido de GC de al menos el 50 por ciento, y se templan a la secuencia objetivo a temperaturas de aproximadamente 68°C a 72°C. La región amplificada se puede secuenciar como se describe anteriormente, y las secuencias traslapadas se ensamblan en una secuencia contigua.

Una técnica de amplificación es la reacción en cadena de la polimerasa (PCR) inversa (véase Triglia y col., Nucl. Acids Res., 16: 8186 (1988)), la cual utiliza enzimas de restricción para generar un fragmento en la región conocida del gen. El fragmento entonces se circulariza mediante ligamiento intramolecular, y se utiliza como una plantilla para la reacción en cadena de la polimerasa (PCR) con cebadores divergentes derivados a partir de la región conocida. Dentro de un planteamiento alternativo, se pueden recuperar las secuencias adyacentes a una secuencia parcial mediante amplificación con un cebador para una secuencia enlazadora y un cebador específico para una región conocida. Las secuencias amplificadas típicamente se someten a una segunda ronda de amplificación con el mismo cebador enlazador y un segundo cebador específico para la región conocida. Una variación sobre este procedimiento, que emplea dos cebadores que inician la extensión en direcciones opuestas a partir de la secuencia conocida, se describe en la Publicación Internacional Número WO 96/38591. Otra técnica se conoce como "amplificación rápida de los extremos del ADNc" o RACE. Esta técnica involucra el uso de un cebador interno y un cebador externo, que se hibrida a una región o secuencia de vector poliA, para identificar las secuencias que son 5' y 3' de una secuencia conocida. Las técnicas adicionales incluyen reacción en cadena de la polimerasa (PCR) de captura (Lagerstrom y col., PCR Methods Applic., 1: 111-19 (1991)), y reacción en cadena de la polimerasa (PCR) caminante (Parker y col., Nucl. Acids. Res., 19: 3055-60 (1991)). También se pueden usar otros métodos que emplean amplificación para obtener una secuencia de ADNc de longitud completa.

En ciertas instancias, es posible obtener una secuencia de ADNc de longitud completa mediante el análisis de secuencias proporcionado en una base de datos de marca de secuencia expresada (EST), tal como aquélla disponible en GenBank. Las búsquedas de las ESTs traslapadas se pueden llevar a cabo en términos generales utilizando programas bien conocidos (por ejemplo, búsquedas en NCBI BLAST), y estas ESTs se pueden utilizar para generar una secuencia contigua de longitud completa. También se pueden obtener secuencias de ADN de longitud completa mediante el análisis de los fragmentos genómicos.

EXPRESIÓN DE POLINUCLEÓTIDOS EN LAS CÉLULAS HUÉSPED

Las secuencias de polinucleótidos o los fragmentos de las mismas que codifican los polipéptidos, o las proteínas de fusión o los equivalentes funcionales de las mismas, se pueden utilizar en las moléculas de ADN recombinante para dirigir la expresión de un polipéptido en las células huésped apropiadas. Debido a la degeneración inherente del código genético, se pueden producir otras secuencias de ADN que codifiquen sustancialmente la misma o una secuencia de aminoácidos funcionalmente equivalente, y estas secuencias se pueden utilizar para clonar y expresar un polipéptido dado.

Como será entendido por los expertos en este campo, puede ser conveniente en algunas instancias producir secuencias de nucleótidos que codifiquen el polipéptido que posean codones que no se presenten naturalmente. Por ejemplo, se pueden seleccionar los codones preferidos por un huésped procariótico o eucariótico particular para aumentar el índice de expresión de proteína, o para producir una transcripción de ARN recombinante que tenga propiedades deseables, tales como una vida media que sea más larga que aquélla de una transcripción generada a partir de la secuencia que se presente naturalmente.

Más aún, las secuencias de polinucleótidos se pueden diseñar utilizando métodos generalmente conocidos en la materia, con el objeto de alterar las secuencias que codifiquen el polipéptido por una variedad de razones, incluyendo, pero no limitándose a, alteraciones que modifiquen la clonación, el procesamiento, y/o la expresión del producto genético. Por ejemplo, se puede utilizar la mezcla del ADN mediante fragmentación aleatoria, y el reensamble mediante reacción en cadena de la polimerasa (PCR) de los fragmentos genéticos y de los oligonucleótidos sintéticos para diseñar las secuencias de nucleótidos. En adición, se puede utilizar mutagénesis dirigida al sitio para insertar nuevos sitios de restricción, para alterar los patrones de glicosilación, para cambiar la preferencia de codones, para producir variantes de empalme, o para introducir mutaciones, y así sucesivamente.

Las secuencias de ácidos nucleicos naturales, modificadas, o recombinantes se pueden ligar con una secuencia heteróloga para codificar una proteína de fusión. Por ejemplo, para rastrear las bibliotecas de péptidos con el fin de buscar inhibidores de la actividad del polipéptido, puede ser útil codificar una proteína quimérica que pueda ser reconocida por un anticuerpo comercialmente disponible. También se pueden diseñar una proteína de fusión para contener un sitio de disociación localizado entre la secuencia de codificación del polipéptido y la secuencia de la proteína heteróloga, de tal manera que el polipéptido se pueda disociar y purificar para deshacerse de la fracción heteróloga.

Las secuencias que codifiquen un polipéptido deseado se pueden sintetizar, del todo o en parte, empleando métodos químicos bien conocidos en la materia (véase Caruthers, M. H. y col., Nucl. Acids Res. Symp. Ser., páginas 215-223 (1980), Horn y col., Nucl. Acids Res. Symp. Ser., páginas 225-232 (1980)). De una manera alternativa, la proteína misma se puede producir utilizando métodos químicos para sintetizar la secuencia de aminoácidos de un polipéptido, o una porción de la misma. Por ejemplo, la síntesis del péptido se puede llevar a cabo utilizando diferentes técnicas en fase sólida (Roberge y col., Science, 269: 202-204 (1995)), y se puede lograr la síntesis automatizada, por ejemplo, utilizando el Sintetizador de Péptidos ABI 431A (Perkin Elmer, Palo Alto, CA).

Un péptido recién sintetizado se puede purificar sustancialmente mediante cromatografía de líquidos de alto rendimiento de preparación (por ejemplo, Creighton, Proteins, Structures and Molecular Principles (1983)) o mediante otras técnicas comparables disponibles en este campo. La composición de los péptidos sintéticos se puede confirmar mediante análisis o secuenciación de aminoácidos (por ejemplo, el procedimiento de degradación de Edman). Adicionalmente, la secuencia de aminoácidos de un polipéptido, o cualquier parte de la misma, se puede alterar durante la síntesis directa y/o se puede combinar utilizando métodos químicos con secuencias a partir de otras proteínas, o cualquier parte de las mismas, para producir un polipéptido variante.

Con el objeto de expresar un polipéptido deseado, las secuencias de nucleótidos que codifican el polipéptido, o los equivalentes funcionales, se pueden insertar en un vector de expresión apropiado, es decir, un vector que contenga los elementos necesarios para la transcripción y traducción de la secuencia de codificación insertada. Se pueden emplear los métodos que son bien conocidos por los expertos en este campo para construir vectores de expresión que contengan secuencias que codifiquen un polipéptido de interés y los elementos de control de transcripción y de traducción apropiados. Estos métodos incluyen las técnicas de ADN recombinante *in vitro*, las técnicas sintéticas, y la recombinación genética in vivo. Estas técnicas se describen en Sambrook y col., Molecular Cloning, A Laboratory Manual (2000), y en Ausubel y col., Current Protocols in Molecular Biology (actualizado anualmente).

Se pueden utilizar una variedad de sistemas de vector de expresión/huésped para contener y expresar las secuencias de polinucleótidos. Éstos incluyen, pero no se limitan a, microorganismos, tales como bacterias transformadas con vectores de expresión de ADN recombinante de bacteriófagos, plásmidos, o cósmidos; levadura transformada con vectores de expresión de levadura; sistemas celulares de insectos infectados con vectores de expresión de virus (por ejemplo, baculovirus) ; sistemas celulares de plantas transformados con vectores de expresión de virus (por ejemplo, virus de mosaico de coliflor, CaMV; virus de mosaico de tabaco, TMV), o con vectores de expresión bacteriana (por ejemplo, plásmidos Ti o pBR322) ; o sistemas celulares de animales.

Los “elementos de control" o las "secuencias reguladoras” presentes en un vector de expresión son las regiones no traducidas del vector -- potenciadores, promotores, regiones no traducidas 5’ y 3’-- que interactúan con las proteínas celulares del huésped para llevar a cabo la transcripción y la traducción. Estos elementos pueden variar en su fuerza y especificidad. Dependiendo del sistema del vector y del huésped utilizado, se puede utilizar cualquier número de elementos de transcripción y traducción adecuados, incluyendo promotores constitutivos e inducibles. Por ejemplo, cuando se clona en sistemas bacterianos, se pueden utilizar los promotores inducibles, tales como el promotor híbrido lacZ del fagémido PBLUESCRIPT (Stratagene, la Jolla, Calif.) o el plásmido PSPORT1 (Gibco BRL, Gaithersburg, MD), y similares. En los sistemas celulares de mamífero, se prefieren en general los promotores a partir de genes de mamífero o a partir de virus de mamífero. Si es necesario generar una línea celular que contenga múltiples copias de la secuencia que codifique un polipéptido, convenientemente se pueden utilizar los vectores basados en SV40 o EBV con un marcador seleccionable apropiado.

En los sistemas bacterianos, se pueden seleccionar un número de vectores de expresión, dependiendo del uso pretendido para el polipéptido expresado. Por ejemplo, cuando se necesitan grandes cantidades, por ejemplo, para la inducción de anticuerpos, se pueden utilizar vectores que dirijan una expresión de alto nivel de las proteínas de fusión que se purifiquen fácilmente. Estos vectores incluyen, pero no se limitan a, los vectores multifuncionales de clonación y expresión de E. coli, tales como BLUESCRIPT (Stratagene), en donde la secuencia que codifica el polipéptido de interés se puede ligar en el vector dentro del marco con las secuencias para la Met amino-terminal y los siguientes 7 residuos de ß-galactosidasa, de tal manera que se produzca una proteína híbrida; vectores pIN (Van Heeke y Schuster, J. Biol. Chem., 264: 5503-5509 (1989)) ; y similares. También se pueden utilizar los vectores pGEX (Promega, Madison, Wis.) para expresar polipéptidos extraños como las proteínas de fusión con S-transferasa de glutationa (GST). En general, estas proteínas de fusión son solubles y se pueden purificar fácilmente para deshacerse de las células lisadas, mediante adsorción en perlas de glutationa-agarosa, seguida por elución en la presencia de glutationa libre. Las proteínas hechas en estos sistemas se pueden diseñar para incluir heparina, trombina, o sitios de disociación de proteasa del factor XA, de tal manera que se pueda liberar el polipéptido de interés clonado a partir de la fracción GST como sea deseado.

En la levadura, Saccharomyces cerevisiae, se pueden utilizar un número de vectores que contengan promotores constitutivos o inducibles, tales como factor alfa, oxidasa de alcohol, y PGH. Otros vectores que contienen promotores constitutivos o inducibles incluyen GAP, PGK, GAL y ADH. Para las reseñas, véase Ausubel y col. (supra), y Grant y col., Methods Enzymol., 153: 516-544 (1987), y Romas y col., Yeast, 8 423-88 (1992).

En los casos en donde se utilizan vectores de expresión de plantas, la expresión de las secuencias que codifiquen los polipéptidos se puede impulsar mediante cualquiera de un número de promotores. Por ejemplo, se pueden

utilizar los promotores virales, tales como los promotores 35S y 19S de CaMV solos o en combinación con la secuencia líder omega a partir de TMV (Takamatsu, EMBO J. 6: 307-311 (1987)). De una manera alternativa, se pueden utilizar los promotores de plantas, tales como la subunidad pequeña de RUBISCO o los promotores de choque por calor (Coruzzi y col., EMBO J., 3: 1671-1680 (1984) ; Broglie y col., Science, 224: 838-843 (1984) ; y Winter y col., Results Probl. Cell Differ., 17: 85-105 (1991)). Estas construcciones se pueden introducir en las células de plantas mediante transformación directa del ADN o mediante transfección mediada por patógenos. Estas técnicas se describen en un número de reseñas generalmente disponibles (véase, por ejemplo, Hobbs en McGraw Hill Yearbook of Science and Technology, páginas 191-196 (1992)).

También se puede utilizar un sistema de insecto para expresar un polipéptido de interés. Por ejemplo, en este sistema, se utiliza el virus de polihedrosis nuclear de Autographa californica (AcNPV) como un vector para expresar genes extraños en células de Spodoptera frugiperda o en larvas de Trichoplusia. Las secuencias que codifican el polipéptido se pueden clonar en una región no esencial del virus, tal como el gen de polihedrina, y se colocan bajo el control del promotor de polihedrina. La inserción con éxito de la secuencia que codifique el polipéptido hará que se inactive el gen de polihedrina y produzca una proteína de recubrimiento que carezca del virus recombinante. Entonces se pueden utilizar los virus recombinantes para infectar, por ejemplo, las células de S. frugiperda o las larvas de Trichoplusia, en donde se puede expresar el polipéptido de interés (Engelhard y col., Proc. Natl. Acad. Sci., EUA, 91: 3224-3227 (1994)).

En las células huésped de mamífero, en general hay un número de sistemas de expresión basados en virus disponibles. Por ejemplo, en los casos en donde se utiliza un adenovirus como un vector de expresión, las secuencias que codifiquen un polipéptido de interés se pueden ligar en un complejo de transcripción/traducción de adenovirus que consiste en el promotor tardío y la secuencia líder tripartita. Se puede utilizar la inserción en una región E1 o E3 no esencial del genoma viral, para obtener un virus viable, el cual es capaz de expresar el polipéptido en las células huésped infectadas (Logan y Shenk, Proc. Natl. Acad. Sci., EUA, 81: 3655-3659 (1984)). En adición, se pueden utilizar potenciadores de transcripción, tales como el potenciador del virus de sarcoma de Rous (RSV), para aumentar la expresión en las células huésped de mamífero. Los métodos y protocolos para trabajar con vectores de adenovirus se revisan en Wold, Adenovirus Methods and Protocols, 1998. Se pueden encontrar referencias adicionales con respecto al uso de los vectores de adenovirus en Adenovirus: A Medical Dictionary, Bibliography, y Annotated Research Guide to Internet References, 2004.

También se pueden utilizar señales de inicio específicas para lograr una traducción más eficiente de las secuencias que codifiquen un polipéptido de interés. Estas señales incluyen el codón de inicio de ATG y las secuencias adyacentes. En los casos en donde se inserten las secuencias que codifiquen el polipéptido, su codón de inicio, y las secuencias corriente arriba en el vector de expresión apropiado, pueden no ser necesarias las señales de control de transcripción o de traducción adicionales. Sin embargo, en los casos en donde solamente se inserte la secuencia de codificación, o una porción de la misma, se deben proporcionar señales de control de traducción exógenas, incluyendo el codón de inicio de ATG. Adicionalmente, el codón de inicio debe estar en el marco de lectura correcto, con el fin de asegurar la traducción del inserto entero. Los elementos de traducción exógenos y los codones de inicio pueden ser de distintos orígenes, tanto naturales como sintéticos. La eficiencia de la expresión se puede mejorar mediante la inclusión de potenciadores que sean apropiados para el sistema celular particular que se utilice, tal como aquéllos descritos en la literatura (Scharf. y col., Results Probl. Cell Differ., 20: 125-162 (1994)).

En adición, se puede seleccionar una cepa de la célula huésped por su capacidad para modular la expresión de las secuencias insertadas o para procesar la proteína expresada en la forma deseada. Estas modificaciones del polipéptido incluyen, pero no se limitan a, acetilación, carboxilación, glicosilación, fosforilación, lipidación, y acilación. También se puede utilizar el procesamiento posterior a la traducción, el cual disocia una forma “prepro” de la proteína para facilitar la inserción, el pliegue y/o la función correcta. Se pueden seleccionar diferentes células huésped, tales como CHO, HeLa, MDCK, HEK293, y WI38, las cuales tengan la maquinaria celular específica y los mecanismos característicos para estas actividades posteriores a la traducción, con el fin de asegurar la modificación y el procesamiento correctos de la proteína extraña.

Para la producción de alto rendimiento a largo plazo de las proteínas recombinantes, en general se prefiere la expresión estable. Por ejemplo, las líneas celulares que expresen establemente un polinucleótido de interés, se pueden transformar utilizando vectores de expresión, los cuales pueden contener orígenes de réplica virales y/o elementos de expresión endógenos, y un gen marcador seleccionable sobre el mismo vector o sobre un vector separado. En seguida de la introducción del vector, se puede permitir que las células crezcan durante 1 a 2 días en un medio enriquecido antes de cambiarse al medio selectivo. El propósito del marcador seleccionable es conferir resistencia a la selección, y su presencia permite el crecimiento y la recuperación de las células que expresen con éxito las secuencias introducidas. Los clones resistentes de las células establemente transformadas pueden proliferar utilizando técnicas de cultivo de tejido apropiadas para el tipo de células.

Se puede utilizar cualquier número de sistemas de selección para recuperar las líneas celulares transformadas. Éstos incluyen, pero no se limitan a, los genes de la cinasa de timidina del virus de herpes símplex (Wigler y col., Cell, 11: 223-32 (1977)), y de la fosforribosil-transferasa de adenina (Lowy y col., Cell, 22: 817-23 (1990)), los cuales se pueden emplear en las células tk.sup.- o aprt.sup.-, respectivamente. También, se puede utilizar resistencia a antimetabolitos, antibióticos o herbicidas, como la base para la selección; por ejemplo, dhfr, que confiere resistencia

al metotrexato (Wigler y col., Proc. Natl. Acad. Sci., EUA, 77: 3567-70 (1980)) ; npt, que confiere resistencia a los aminoglicósidos, a la neomicina, y al G-418 (Colbere-Garapin y col., J. Mol. Biol., 150: 1-14 (1981)) ; y als o pat, que confieren resistencia al clorosulfurón y a la acetil-transferasa de fosfinotricina, respectivamente (Murry, supra). Se han descrito genes seleccionables adicionales, por ejemplo, trpB, que permite que las células utilicen indol en lugar de triptófano, o hisD, que permite que las células utilicen histinol en lugar de histidina (Hartman y Mulligan, Proc. Natl. Acad. Sci., EUA, 85: 8047-51 (1988)). Recientemente, el uso de marcadores visibles ha ganado popularidad con los marcadores tales como antocianinas, ß-glucuronidasa y sus sustrato GUS, y luciferasa y su sustrato luciferina, que se utiliza ampliamente no solo para identificar los transformantes, sino también para cuantificar la cantidad de expresión de proteína transitoria o estable que se pueda atribuir a un sistema de vector específico (Rhodes y col., Methods Mol. Biol., 55: 121-131 (1995)).

Aunque la presencia/ausencia de la expresión del gen marcador sugiere que también está presente el gen de interés, puede no necesitarse confirmar su presencia y expresión. Por ejemplo, si la secuencia que codifica un polipéptido se inserta dentro de una secuencia del gen marcador, se pueden identificar las células recombinantes que contengan las secuencias, por la ausencia de la función del gen marcador. De una manera alternativa, un gen marcador se puede colocar en fila con una secuencia que codifique el polipéptido bajo el control de un solo promotor. La expresión del gen marcador en respuesta a la inducción o selección, usualmente indica la expresión del gen en fila también.

De una manera alternativa, las células huésped que contienen y expresan una secuencia de polinucleótido deseada, se pueden identificar mediante una variedad de procedimientos conocidos por los expertos en este campo. Estos procedimientos incluyen, pero no se limitan a, hibridaciones de ADN-ADN o ADN-ARN, y técnicas de bioensayo de proteínas o de inmunoensayos, las cuales incluyen las tecnologías basadas en membrana, en solución, o en chip para la detección y/o cuantificación del ácido nucleico o de la proteína.

En la materia se conocen una variedad de protocolos para detectar y medir la expresión de los productos codificados por el polinucleótido, utilizando cualquiera de los anticuerpos policlonales o monoclonales específicos para el producto. Los ejemplos incluyen el análisis de inmunoabsorción ligado a enzimas (ELISA), radioinmunoensayo (RIA), y selección de células activada por fluorescencia (FACS). Para algunas aplicaciones, se puede preferir un inmunoensayo basado en anticuerpos monoclonales, de dos sitios, que utilice anticuerpos monoclonales que reaccionen con dos epítopos que no interfieran sobre un polipéptido dado, pero también se puede emplear un ensayo de enlace competitivo. Estos y otros ensayos se describen, entre otros lugares, en Hampton y col., Serological Methods, a Laboratory Manual (1990), y Maddox y col., J. Exp. Med., 158: 1211-1216 (1983).

Los expertos en la materia conocen una amplia variedad de marcas y técnicas de conjugación, y se puede utilizar en diferentes ensayos de ácidos nucleicos y de aminoácidos. Los medios para producir sondas marcadas de hibridación o de reacción en cadena de la polimerasa (PCR), para detectar las secuencias relacionadas con los polinucleótidos, incluyen oligomarcado, traducción de muesca, marcado de extremos, o amplificación mediante reacción en cadena de la polimerasa (PCR) utilizando un nucleótido marcado. De una manera alternativa, las secuencias, o cualesquiera porciones de las mismas, se pueden clonar en un vector para la producción de una sonda de ARNm. Estos vectores son conocidos en la materia, están comercialmente disponibles, y se pueden utilizar para sintetizar sondas de ARN *in vitro* mediante la adición de una polimerasa de ARN apropiada, tal como T7, T3, o SP6 y los nucleótidos marcados. Estos procedimientos se pueden conducir utilizando una variedad de kits comercialmente disponibles. Las moléculas reporteras o marcas adecuadas que se pueden utilizar incluyen radionúclidos, enzimas, agentes fluorescentes, quimiluminiscentes, o cromogénicos, así como sustratos, co-factores, inhibidores, partículas magnéticas, y similares.

Las células huésped transformadas con una secuencia de polinucleótido de interés se pueden cultivar bajo condiciones adecuadas para la expresión y recuperación de la proteína a partir del cultivo celular. La proteína producida mediante una célula recombinante se puede secretar o puede ser contenida intracelularmente, dependiendo de la secuencia y/o del vector utilizado. Como será entendido por los expertos en este campo, los vectores de expresión que contienen polinucleótidos se pueden diseñar para contener secuencias de señales que dirijan la secreción del polipéptido codificado a través de una membrana celular procariótica o eucariótica. Se pueden utilizar otras construcciones recombinantes para unir las secuencias que codifiquen un polipéptido de interés para la secuencia de nucleótidos que codifique un dominio de polipéptido que facilite la purificación de las proteínas solubles. Estos dominios que facilitan la purificación incluyen, pero no se limitan a, péptidos quelantes de metales, tales como módulos de histidina-triptófano, que permiten la purificación sobre metales inmovilizados, dominios de proteína A, que permiten la purificación sobre inmunoglobulina inmovilizada, y el dominio utilizado en el sistema de purificación de extensión/afinidad FLAGS (Immunex Corp., Seattle, Wash.). Se puede utilizar la inclusión de las secuencias enlazadoras disociables, tales como aquéllas específicas para el Factor XA o la enterocinasa (Invitrogen. San Diego, Calif.) entre el dominio de purificación y el polipéptido codificado para facilitar la purificación. Uno de estos vectores de expresión proporciona la expresión de una proteína de fusión que contiene un polipéptido de interés y un ácido nucleico que codifica 6 residuos de histidina precediendo a un sitio de disociación de tiorredoxina o de enterocinasa. Los residuos de histidina facilitan la purificación sobre IMIAC (cromatografía de afinidad de ion de metal inmovilizado), como se describe en Porath y col., Prot. Exp. Purif., 3: 263-281 (1992), mientras que el sitio de disociación de enterocinasa proporciona un medio para la purificación del polipéptido deseado a partir de la proteína de fusión. Se proporciona una discusión de los vectores que contienen las proteínas de fusión en Kroll y col., DNA

Cell Biol., 12: 441-453 (1993)).

TÉCNICAS DE SUMINISTRO DE POLINUCLEÓTIDOS IN VIVO

En las realizaciones adicionales, las construcciones genéticas que comprenden uno o más de los polinucleótidos de la invención se introducen en las células in vivo. Esto se puede lograr utilizando cualquiera de una variedad de planteamientos bien conocidos, algunos de los cuales se ilustran a continuación para el propósito de ilustración.

1. ADENOVIRUS

Uno de los métodos preferidos para el suministro in vivo de una o más secuencias de ácidos nucleicos involucra el uso de un vector de expresión de adenovirus. "Vector de expresión de adenovirus" pretende incluir las construcciones que contienen secuencias de adenovirus suficientes para: (a) soportar el empaque de la construcción, y (b) expresar un polinucleótido que se haya clonado en la misma en una orientación en sentido o antisentido. Desde luego, en el contexto de una construcción antisentido, la expresión no requiere que se sintetice el producto genético.

El vector de expresión comprende una forma genéticamente diseñada de un adenovirus. El conocimiento de la organización genética del adenovirus, un virus de ADN de doble cadena, lineal, de 36 kb, permite la sustitución de grandes pedazos de ADN adenoviral con secuencias extrañas de hasta 7 kb (Grunhaus y Horwitz, 1992). En contraste con el retrovirus, la infección adenoviral de las células huésped no da como resultado la integración cromosómica, debido a que el ADN adenoviral puede replicarse de una manera episomal sin una genotoxicidad potencial. También, los adenovirus son estructuralmente estables, y no se ha detectado ninguna reconfiguración del genoma después de una extensa amplificación. Los adenovirus pueden infectar virtualmente todas las células epiteliales independientemente de la etapa de su ciclo celular. Hasta ahora, la infección adenoviral parece estar vinculada solamente a la enfermedad leve, tal como la enfermedad respiratoria aguda en los seres humanos.

El adenovirus está en particular adecuado para utilizarse como un vector de transferencia genética, debido a su genoma de tamaño mediano, a su facilidad de manipulación, a su alta titulación, a su amplia gama de células objetivo, y a su alta infectividad. Ambos extremos del genoma viral contienen repeticiones invertidas (ITRs) de 100 a 200 pares de bases, que son los elementos cis necesarios para la réplica y el empaque del ADN viral. Las regiones temprana (E) y tardía (L) del genoma contienen diferentes unidades de transcripción que se dividen entre el establecimiento de la réplica del ADN viral. La región E1 (E1A y E1B) codifica las proteínas responsables de la regulación de la transcripción del genoma viral y unos cuantos genes celulares. La expresión de la región E2 (E2A y E2B) da como resultado la síntesis de las proteínas para la réplica del ADN viral. Estas proteínas están involucradas en la réplica del ADN, en la expresión genética tardía, y en el cierre de la célula huésped (Renan, 1990). Los productos de los genes tardíos, incluyendo la mayoría de las proteínas de cápsida viral, se expresan solamente después de un procesamiento significativo de una sola transcripción primaria expedida por el promotor tardío mayor (MLP). El MLP, (localizado en 16,8 m.u.) es particularmente eficiente durante la fase tardía de la infección, y todos los ARNms expedidos a partir de este promotor poseen una secuencia líder tripartita-5' (TPL), lo cual los convierte en los ARNms preferidos para la traducción.

En un sistema actual, se genera el adenovirus recombinante a partir de la recombinación homóloga entre el vector de mezcla y el vector de provirus. Debido a la posible recombinación entre dos vectores provirales, se puede generar el adenovirus de tipo silvestre a partir de este proceso. Por consiguiente, es crítico aislar un solo clon del virus a partir de una placa individual, y examinar su estructura genómica.

La generación y propagación de los vectores de adenovirus actuales, que son deficientes en réplica, dependen de una línea celular auxiliar única, designada como 293, la cual se transformó a partir de células de riñón embrionario humano, mediante fragmentos de ADN Ad5, y expresan constitutivamente las proteínas E1 (Graham y col., 1977). Debido a que se puede prescindir de la región E3 del genoma del adenovirus (Jones y Shenk, 1978), los vectores de adenovirus actuales, con la ayuda de las células 293, llevan el ADN extraño a cualquiera de las regiones E1, D3, o ambas (Graham y Prevec, 1991). En la naturaleza, el adenovirus puede empacar aproximadamente el 105 % del genoma de tipo silvestre (Ghosh-Choudhury y col., 1987), proporcionando capacidad para aproximadamente 2 kB extra de ADN. Combinada con los aproximadamente 5,5 kB de ADN que pueden ser reemplazados en las regiones E1 y E3, la máxima capacidad del vector de adenovirus actual está por debajo de 7,5 kB, o aproximadamente el 15 % de la longitud total del vector. Más del 80 % del genoma viral del adenovirus permanece en la estructura base del vector, y es la fuente de la citotoxicidad surgida del vector. También, la deficiencia en réplica del virus con la E1 suprimida es incompleta. Por ejemplo, se ha observado una fuga de expresión genética viral con los vectores actualmente disponibles en altas multiplicidades de infección (MOI) (Mulligan, 1993).

Las líneas celulares auxiliares se pueden derivar a partir de células humanas, tales como células de riñón embrionario humano, células de músculo, células hematopoyéticas, u otras células mesenquimatosas o epiteliales embrionarias humanas. De una manera alternativa, Las células auxiliares se pueden derivar a partir de las células de otras especies de mamíferos que sean permisivas para el adenovirus humano. Estas células incluyen, por ejemplo, células Vero células u otras células mesenquimatosas o epiteliales embrionarias de mono. Como se menciona anteriormente, la línea celular auxiliar actualmente preferida es la 293.

Racher y col. (1995) han dado a conocer métodos mejorados para cultivar las células 293, y para propagar el adenovirus. En un formato, se cultivan aglomerados de células naturales mediante la inoculación de las células individuales en matraces centrífugos siliconizados de 1 litro (Techne, Cambridge, Reino Unido) que contienen de 100 a 200 mililitros del medio. En seguida de la agitación a 40 revoluciones por minuto, se estima la viabilidad celular con azul de tripano. En otro formato, se emplean microvehículos Fibra-Cel (Bibby Sterlin, Stone, Reino Unido) (5 gramos/litro) como sigue. Un inóculo celular, resuspendido en 5 mililitros del medio, se agrega al vehículo (50 mililitros) en un matraz Erlenmeyer de 250 mililitros, y se deja estacionario, con agitación ocasional, durante 1 a 4 horas. Entonces se reemplaza el medio con 50 mililitros de medio fresco, y se inicia la agitación. Para la producción del virus, las células se dejan crecer hasta una confluencia de aproximadamente el 80 por ciento; después de cuyo tiempo, se reemplaza el medio (hasta el 25 % del volumen final), y se agrega el adenovirus a una multiplicidad de infección (MOI) de 0,05. Los cultivos se dejan estacionarios durante la noche, en seguida de lo cual, se aumenta el volumen hasta el 100 por ciento, y se comienza la agitación durante otras 72 horas.

Para algo que no sea el requerimiento de que el vector de adenovirus sea defectuoso en réplica, o al menos condicionalmente defectuoso, no se cree que la naturaleza del vector de adenovirus sea crucial para la práctica con éxito de la invención. El adenovirus puede ser de cualquiera de los 42 diferentes serotipos conocidos o subgrupos A-F. El Adenovirus tipo 5 del subgrupo C es el material de partida preferido con el fin de obtener un vector de adenovirus defectuoso en réplica condicional, para utilizarse en la presente invención, debido a que el Adenovirus tipo 5 es un adenovirus humano acerca del cual se conoce una gran cantidad de información bioquímica y genética, e históricamente se ha utilizado para la mayoría de las construcciones que emplean adenovirus como un vector.

Como se menciona anteriormente, el vector típico de acuerdo con la presente invención es defectuoso en réplica, y no tendrá la región de adenovirus E1. Por consiguiente, será más conveniente introducir el polinucleótido que codifique el gen de interés en la posición a partir de la cual se hayan removido las secuencias de codificación de E1. Sin embargo, la posición de inserción de la construcción dentro de las secuencias de adenovirus no es crítica para la invención. El polinucleótido que codifique el gen de interés también se puede insertar en lugar de la región E3 suprimida en los vectores de reemplazo de E3, como es descrito por Karlsson y col. (1986) o en la región E4 en donde una línea celular auxiliar o un virus auxiliar complemente el defecto de E4.

El adenovirus es fácil de cultivar y manipular, y exhibe una amplia gama de huéspedes *in vitro* e in vivo. Este grupo de virus se puede obtener en altas titulaciones, por ejemplo, de 109 a 1011 unidades formadoras de placas por mililitro, y son altamente inefectivos. El ciclo de vida del adenovirus no requiere de su integración en el genoma de la célula huésped. Los genes extraños suministrados por los vectores de adenovirus son episomales y, por consiguiente, tienen una baja genotoxicidad para las células huésped. No se han reportado efectos secundarios en los estudios de vacunación con adenovirus de tipo silvestre (Couch y col., 1963; Top y col., 1971), demostrando su seguridad y su potencial terapéutico como vectores de transferencia genética in vivo.

Los vectores de adenovirus se han utilizado en la expresión genética eucariótica (Levrero y col., 1991; Gomez-Foix y col., 1992), y en el desarrollo de vacunas (Grunhaus y Horwitz, 1992; Graham y Prevec, 1992). Recientemente, los estudios con animales sugirieron que se podría utilizar el adenovirus recombinante para la terapia genética (Stratford-Perricaudet y Perricaudet, 1991; Stratford-Perricaudet y col., 1990; Rich y col., 1993). Los estudios de la administración del adenovirus recombinante a diferentes tejidos incluyen instilación en la tráquea (Rosenfeld y col., 1991; Rosenfeld y col., 1992), inyección muscular (Ragot y col., 1993), inyecciones intravenosas periféricas (Herz y Gerard, 1993), e inoculación estereotáctica en el cerebro (Le Gal La Salle y col., 1993).

Los vectores de adenovirus se pueden originar a partir del adenovirus humano. De una manera alternativa, se pueden originar a partir de adenovirus de otras especies, por ejemplo, de chimpancé, el cual puede tener la ventaja de que los vectores virales no son neutralizados por los anticuerpos contra el adenovirus humano que circula en muchos sujetos humanos (véase, por ejemplo: Tatsis N y col., Gene, Therapy, 2006, 13: 421-429).

El Adenovirus tipo 35, el cual es relativamente poco común y, por consiguiente, existen bajos niveles de inmunidad previamente existente al vector mismo, se ha utilizado como un sistema de suministro en ciertas vacunas de tuberculosis que se están desarrollando (véase, por ejemplo, Radosevic y col., Infection and Immunity, 2007, 75(8) : 4105-4115). El Adenovirus tipo 35 también puede ser de un valor particular en la presente invención como un vector de suministro.

2. RETROVIRUS

Los retrovirus son un grupo de virus de ARN de una sola cadena caracterizados por una capacidad para convertir su ARN en ADN de cadena doble en el ADN de células infectadas mediante un proceso de transcripción inversa (Coffin, 1990). El ADN resultante se integra establemente dentro de los cromosomas celulares como un provirus y dirige la síntesis de proteínas virales. La integración da como resultado la retención de las secuencias del gen viral en la célula receptora y sus descendientes. El genoma retroviral contiene tres genes, gag, pol, y env que codifican para proteínas cápsidas, enzima polimerasa, y componentes de envoltura, respectivamente. Una secuencia encontrada corriente arriba del gen gag contiene una señal para empacar el genoma dentro de viriones. Dos secuencias de repetición terminal larga (LTR) están presentes en los extremos 5' y 3' del genoma viral. Estas contienen secuencias de promotor fuerte y mejoradoras y también son requeridas para la integración en el genoma

de la célula huésped (Coffin, 1990).

Con el fin de construir un vector retroviral, se inserta un ácido nucleico que codifica una o más secuencias de oligonucleótidos o polinucleótidos de interés dentro del genoma viral en el lugar de ciertas secuencias virales para producir un virus que carece de replicación. Con el fin de producir viriones, se construye una línea celular de empaque que contiene los genes gag, pol y env pero sin la LTR y los componentes de empaque (Mann y col., 1983). Cuando un plásmido recombinante que contiene un ADNc, junto con la LTR retroviral y secuencias de empaque se introduce dentro de esta línea celular (mediante precipitación de fosfato de calcio por ejemplo), la secuencia de empaque permite que la transcripción de ARN del plásmido recombinante sea empacada dentro de partículas virales, que se secretan entonces en el medio de cultivo (Nicolas y Rubenstein, 1988; Temin, 1986; Mann y col., 1983). El medio que contiene el retrovirus recombinante se recolecta, opcionalmente se concentra, y se usa para transferencia de genes. Los vectores retrovirales pueden infectar una amplia variedad de tipos de células. Sin embargo, la integración y la expresión estable requieren la división de células huéspedes (Paskind y col., 1975).

Un enfoque novedoso diseñado para permitir un direccionamiento específico de vectores de retrovirus recientemente se desarrolló basado en la modificación química de un retrovirus por la adición química de residuos de lactosa a la envoltura viral. Esta modificación podría permitir la infección específica de hepatocitos vía los receptores de sialoglicoproteína.

Se diseñó un enfoque diferente para dirigirse a los retrovirus recombinante en el cual se usaron anticuerpos biotinilados contra una proteína de envoltura retroviral y contra un receptor celular específico. Los anticuerpos se acoplaron vía los componentes de biotina usando estreptavidina (Roux y col., 1989). Usando anticuerpos contra los antígenos clase I y clase II del complejo mayor de histocompatibilidad, ellos demostraron la infección de una variedad de células humanas que llevan esos antígenos de superficie con un virus ecotrópico *in vitro* (Roux y col., 1989).

3. VIRUS ADENO-ASOCIADOS

AAV (Ridgeway, 1988; Hermonat y Muzycska, 1984) es un parovirus, descubierto como una contaminación de cepas adenovirales. Es un virus ubicuo (los anticuerpos están presentes en el 85 % de la población humana de los Estados Unidos de América) que no se ha relacionado con enfermedad alguna. También se clasifica como un dependovirus, porque su réplica depende de la presencia de un virus auxiliar, tal como un adenovirus. Se han aislado cinco serotipos, de los cuales el AAV-2 es el mejor caracterizado. ASV tiene un ADN lineal de una sola cadena que está encapsidado en proteínas cápsidas VP1, VP2 y VP3 para formar un virión icosahédrico de 20 a 24 nanómetros de diámetro (Muzyczka y McLaughlin, 1988).

El ADN del AAV tiene una longitud aproximada de 4,7 kilobases. Contiene dos marcos de lectura abierta y está flanqueado por dos ITR. Hay dos genes mayores en el genoma AAV: rep y cap. El gen rep codifica para las proteínas responsables de las replicaciones virales, mientras que el cap codifica para la proteína cápsida VP1-3. Cada ITR forma una estructura de pasador de pelo en forma de T. Estas repeticiones terminales son los únicos componentes cis esenciales del AAV para la integración cromosómica. Por lo tanto, el AAV se puede usar como un vector con todas las secuencias de codificación viral removidas y remplazadas por casetes de genes para el suministro. Se han identificado tres promotores virales identificados y nombrados p5, p19, y p40, de acuerdo con su posición en el mapa. La transcripción de p5 y p19 da como resultado la producción de proteínas rep, y la transcripción de p40 produce las proteínas cápsidas (Hermonat y Muzyczka, 1984).

Hay varios factores que estimulan a los investigadores a estudiar la posibilidad de usar rAAV como un vector de expresión. Uno es que los requerimientos para entregar un gen para integrarlo en el cromosoma huésped son sorprendentemente pocos. Es necesario tener los ITRx de 145 pares de bases, que son solamente el 6 % del genoma de AAV. Esto deja espacio en el vector para armar una inserción de ADN de 4,5 kilobases. Aunque esta capacidad para acarreo puede evitar que el AAV suministre genes grandes, es ampliamente adecuada para suministrar construcciones antisentido.

El AAV también es una buena elección de vehículos de suministro debido a su seguridad. Hay un mecanismo relativamente complicado de rescate: no solamente el adenovirus del tipo natural, sino también los genes del AAV se requieren para movilizar el rAAV. Igualmente, el AAV no es patógeno y no se asocia con enfermedad alguna. La remoción de las secuencias de codificación virales minimiza las reacciones inmunes a la expresión de genes virales, y por lo tanto, rAAV no evoca una respuesta inflamatoria.

AAV también es una buena elección de vehículos de suministro debido a su seguridad. Hay un mecanismo de rescate relativamente complicado: no solamente el adenovirus del tipo natural sino también los genes de AAV se requieren para movilizar el rAAV. Igualmente, AAV no es patógeno y no está asociado con enfermedad alguna. La remoción de las secuencias de codificación viral minimiza las reacciones inmunes a la expresión de gen viral, y por lo tanto, rAAV no evoca una respuesta inflamatoria.

4. OTROS VECTORES VIRALES COMO CONSTRUCCIONES DE EXPRESIÓN

Se pueden emplear otros vectores virales como construcciones de expresión en la presente invención para la

suministro de secuencias de oligonucleótidos o polinucleótidos a una célula huésped. Se pueden emplear vectores derivados de virus tales como el virus de vacuna (Ridgeway, 1988; Coupar y col., 1988), lentivirus, virus de polio y virus de herpes. También se puede esperar que sean útiles otros vectores derivados de virus de viruelas, tales como los vectores derivados de viruela aviaria. Ellos ofrecen varias características atractivas para distintas células de mamíferos (Friedmann, 1989; Ridgeway, 1988; Coupar y col., 1988; Horwich y col., 1990).

Con el reciente reconocimiento del virus de hepatitis B defectuoso, se ha obtenido nuevo conocimiento de la relación estructura-función de diferentes secuencias virales. Los estudios *in vitro* mostraron que el virus puede retener la capacidad para empaque dependiente de auxiliar y a pesar de la transcripción inversa la eliminación de hasta el 80 % de su genoma (Horwich y col., 1990). Esto sugirió que grandes porciones del genoma podrían ser remplazadas con material genético extraño. El hepatotropismo y la persistencia (integración) fueron propiedades particularmente atractivas para la transferencia de genes dirigida por hígado. Chang y col. (1991) introdujeron el gen de cloranfenicol acetil-transferasa (CAT) en el genoma de virus de hepatitis B de pato en lugar de las secuencias de codificación de superficie y pre-superficie que codifican la polimerasa. Fue co-transfectado con virus del tipo natural dentro de una línea celular de hepatoma aviario. Se usaron medios de cultivo que contenían altas titulaciones del virus recombinante para infectar hepatocitos primarios de pato. Se detectó una expresión de gen CAT estable durante al menos 24 días después de la transfección (Chang y col., 1991).

Los vectores 'virales' adicionales incluyen partículas que parecen virus (VLPs) y fagos.

5. VECTORES NO VIRALES

Para efectuar la expresión de las secuencias de oligonucleótidos o polinucleótidos de la presente invención, la construcción de expresión debe ser suministrada dentro de una célula. Esta entrega puede llevarse a cabo *in vitro*, como en procedimientos de laboratorio para transformar líneas celulares, o in vivo o ex vivo, como en el tratamiento de ciertos estados de enfermedad. Como se describe anteriormente, un mecanismo preferido para la entrega es vía infección viral donde la construcción de expresión se encapsula en una partícula viral infecciosa.

En cuanto la construcción de expresión ha sido entregada en la célula el ácido nucleico que codifica las secuencias deseadas de oligonucleótidos o polinucleótidos se puede posicionar y expresar en diferentes sitios. En ciertas realizaciones, el ácido nucleico que codifica la construcción puede integrarse de manera estable en el genoma de la célula. Esta integración puede ser en el lugar y orientación específicos vía la recombinación homóloga (reemplazo de genes) o puede integrarse en un lugar aleatorio, no específico (aumento de gen). En todavía otras realizaciones, el ácido nucleico se puede mantener establemente en la célula como un segmento episomal, separado de ADN. Estos segmentos de ácido nucleico o "episomas" codifican secuencias suficientes para permitir el mantenimiento y la replicación independientes de o en sincronización con el ciclo de la célula huésped. De qué manera se entrega la construcción de expresión a una célula y dónde en la célula permanece el ácido nucleico depende del tipo de construcción de expresión empleado.

En ciertas realizaciones de la invención, la construcción de la expresión que comprende una o más secuencias de oligonucleótidos o polinucleótidos simplemente puede consistir en plásmidos o ADN recombinante no protegido. Se puede realizar la transferencia de una construcción, por ejemplo, por cualquier método que físicamente o químicamente permeabilice la membrana celular. Esto es particularmente aplicable para la transferencia *in vitro* pero se puede aplicar al uso in vivo también. Dubensky y col. (1984) inyectaron con éxito ADN de poliomavirus en forma de precipitados de fosfato de calcio en el hígado y bazo de ratones adultos y recién nacidos demostrando la replicación viral activa y la infección aguda. Benvenisty y Reshef (1986) también demostraron que la inyección intraperitoneal directa de plásmidos precipitados de fosfato de calcio da como resultado la expresión de los genes transfectados. Se anticipa que ADN que codifica un gen de interés también se puede transferir de manera similar en vivo y expresar el producto del gen.

Otra realización de la invención para transferir una construcción de expresión de ADN no protegido (naked) puede incluir bombardeo de partículas. Este método depende de la capacidad para acelerar microproyectiles recubiertos de ADN a una alta velocidad permitiéndoles perforar membranas celulares y entrar a las células sin matarlas (Klein y col., 1987). Se han desarrollado varios dispositivos para acelerar partículas pequeñas. Uno de estos dispositivos depende de una descarga de alto voltaje para generar una corriente eléctrica, la cual a su vez proporciona la fuerza motriz (Yang y col., 1990). Los microproyectiles usados han consistido de sustancias biológicamente inertes tales como tungsteno o perlas de oro.

Los órganos seleccionados que incluyen el hígado, piel y tejido muscular de ratas y ratones se han bombardeado en vivo (Yang y col., 1990; Zelenin y col., 1991). Esto puede requerir exposición quirúrgica del tejido o las células, para eliminar cualquier tejido que interviene entre el cañón y el órgano objetivo, es decir, tratamiento ex vivo. De nuevo, el ADN que codifica un gen particular se puede suministrar vía este método y todavía ser incorporado.

También se pueden usar bacterias como un método de suministro (por ejemplo listeria, véase WO2004/11048) y en particular BCG.

COMPOSICIONES POLIPEPTÍDICAS

Se pueden preparar los polipéptidos usando cualquier variedad de técnicas muy conocidas. Los polipéptidos recombinantes codificados por secuencias de ADN como se describe anteriormente se pueden preparar usando cualquiera de una variedad de vectores de expresión conocidos para los expertos en la técnica. Se puede lograr la expresión en cualquier célula huésped que ha sido transformada o transfectada con un vector de expresión que contiene una molécula que codifica un polipéptido recombinante. Los células huéspedes adecuadas incluyen procariotas, levaduras y células eucariotas superiores, tales como células de mamíferos y células vegetales. De preferencia, las células huéspedes empleadas son E. coli, levaduras o una línea celular de mamífero tal como COS o CHO. Los sobrenadantes de sistemas adecuados de huésped/vector que secretan proteína recombinante o polipéptido en medio de cultivo pueden concentrarse primero usando un filtro comercialmente disponible. Después de la concentración, se puede aplicar el concentrado a una matriz de purificación adecuada, tal como una matriz de afinidad o una resina de intercambio de iones. Finalmente, se puede emplear uno o más etapas de HPLC de fase inversa para purificar adicionalmente un polipéptido recombinante.

También se pueden generar polipéptidos más cortos mediante medios sintéticos, usando técnicas muy conocidas para los expertos en la técnica. Por ejemplo, estos polipéptidos se pueden sintetizar usando cualquier técnica de fase sólida disponible comercialmente, tales como el método de síntesis de fase sólida de Merrifield, donde los aminoácidos se añaden secuencialmente a una cadena de aminoácidos creciente. Véase Merrifield, J. Am. Chem. Soc. 85: 2149-2146 (1963). El equipo para la síntesis automatizada de polipéptidos está comercialmente disponible de proveedores como Perkin Elmer/Applied BioSystems Division (Foster City, CA), y se pueden operar de acuerdo a las instrucciones del fabricante.

Dentro de ciertas realizaciones específicas, un polipéptido puede ser una proteína de fusión que comprende múltiples proteínas modificadas Rv3616c como se describe en la presente, o que comprende al menos una proteína Rv3616c modificada como se describe en la presente y una secuencia no relacionada tal como las descritas en (i) a (xvi) y (a) a (g) en lo anterior.

Un asociado de fusión, por ejemplo, ayuda a proporcionar epítopos auxiliares T (un asociado de fusión inmunológica), de preferencia epítopos auxiliares T reconocidos por humanos, o puede ayudar a expresar la proteína (un realzador de expresión) a rendimientos superiores que la proteína recombinante natural. Ciertos asociados de fusión preferidos son asociados de fusión tanto inmunológicos como realzadores de la expresión. Otros asociados de fusión se pueden seleccionar para aumentar la solubilidad de la proteína o para permitir que la proteína sea destinada a los compartimientos intracelulares deseados. Todavía otros asociados de fusión incluyen marcadores de afinidad, que facilitan la purificación de la proteína.

Las proteínas de fusión generalmente se pueden preparar usando técnicas estándar, incluyendo conjugación química. De preferencia, una proteína de fusión se expresa como una proteína recombinante, permitiendo la producción de niveles aumentados, relativos a una proteína no fusionada, en un sistema de expresión. En resumen, las secuencias de ADN que codifican los componentes del polipéptido se pueden ensamblar por separado y ligarse a un vector de expresión apropiado. El extremo 3' de la secuencia de ADN que codifica un componente de polipéptido se liga, con o sin un enlazador péptido, al extremo 5' de una secuencia de ADN que codifica el segundo componente de polipéptido de manera que los cuadros de lectura de las secuencias están en fase. Esto permite la traducción en una sola proteína de fusión que retiene la actividad biológica en ambos polipéptidos componentes.

Se puede emplear una secuencia de enlazador de péptido para separar los asociados de la fusión por una distancia suficiente para asegurar que cada polipéptido se pliegue en sus estructuras secundarias y terciarias. Esta secuencia de enlazador de péptido se incorpora dentro de la proteína de fusión usando técnicas estándares muy conocidas en la técnica. Se pueden elegir secuencias de enlazadores de péptidos adecuadas basándose en los siguientes factores: (1) su capacidad para adoptar una conformación extendida flexible; (2) su incapacidad para adoptar una estructura secundaria que pudiera interactuar con epítopos funcionales en el primero y el segundo polipéptidos; y (3) la falta de residuos hidrofóbicos o cargados que podrían reaccionar con los epítopos funcionales de polipéptidos. Las secuencias de enlazador de péptidos preferidas contienen residuos Gly, Asn y Ser. Otros aminoácidos neutros, tales como Thr y Ala se pueden usar también en la secuencia enlazadora. Las secuencias de aminoácidos que se pueden emplear de manera útil como enlazadores incluyen las dadas a conocer en Maratea y col., Gene 40: 39-46 (1985) ; Murphy y col., Proc. Natl. Acad. Sci. USA 83: 8258-8262 (1986) ; Patente de Estados Unidos 4,935,233 y Patente de Estados Unidos 4,751,180. No se requieren secuencias enlazadoras cuando el primero y el segundo polipéptidos no tienen regiones amino N-terminales no esenciales que se puedan usar para separar los dominios funcionales y evitar la interferencia estérica.

Dentro de las realizaciones preferidas, un asociado de fusión inmunológica se deriva a partir de la proteína D, una proteína de superficie de la bacteria gram-negativa Haemophilus influenza B (WO 91/18926). De preferencia, un derivado de proteína D comprende aproximadamente el primer tercio de la proteína (por ejemplo, los primeros 100-110 aminoácidos de la terminal N), y un derivado de la proteína D puede ser lipidado. Dentro de ciertas realizaciones preferidas, los primeros 109 residuos de un asociado de fusión de la lipoproteína D se incluye en el término N para proporcionar al polipéptido epítopos de células T exógenos adicionales y para aumentar el nivel de la expresión en E. coli (funcionando de este modo como un realzador de la expresión). La cola de lípido asegura la presentación óptima de las células que presentan antígeno a antígeno. Otros asociados de fusión incluyen la proteína no estructural del virus de influenza, NS1 (hemaglutinina). Típicamente, se usan los 81 aminoácidos N-terminales,

aunque se pueden usar diferentes fragmentos que incluyen epítopos del auxiliar T.

En otra realización, el asociado de la fusión inmunológica es la proteína conocida como LYTA, o una porción de la misma (de preferencia una porción de la terminal C). LYTA se deriva a partir de Streptococcus pneumoniae, que sintetiza una N-acetil-L-alanina amidasa conocida como amidasa de LYTA (codificada por el gen LytA; Gene 43: 265-292 (1986)). LYTA es una autolisina que específicamente degrada ciertos enlaces en la estructura central del peptidoglicano. El dominio de terminal C de la proteína LYTA es responsable de la afinidad a la colina o a algunos análogos de colina tales como DEAE. Esta propiedad ha sido explotada para el desarrollo de plásmidos que expresan C-LYTA de E. coli útiles para la expresión de las proteínas de fusión. Se ha descrito la purificación de proteínas híbridas que contienen el fragmento de C-LYTA en el término amino (véase Biotechnology 10: 795-798 (1992)). Dentro de una realización preferida, una porción de repetición de LYTA se puede incorporar a una proteína de fusión. Una porción de repetición se encuentra en la región terminal C comenzando en el residuo 178. Una porción de repetición preferida incorpora los residuos 188-305.

COMPOSICIONES FARMACÉUTICAS

En las realizaciones adicionales, el polinucleótido o las composiciones de polinucleótidos dadas a conocer en la presente, se pueden formular en soluciones farmacéuticamente aceptables o fisiológicamente aceptables para su administración a una célula o a un animal, ya sea solos, o en combinación con una u otras realizaciones de terapia. Las composiciones se pueden presentar en forma de polvo (por ejemplo, secadas por congelación) para la reconstitución poco antes del uso, y estas composiciones secas generalmente son más estables durante el almacenamiento.

Las composiciones farmacéuticas pueden comprender una proteína de fusión o un polinucleótido que codifica una proteína de fusión, en combinación con un vehículo o excipiente farmacéuticamente aceptable.

Queda entendido también que, si se desea, el segmento de ácido nucleico (por ejemplo ARN o ADN) que expresa un polipéptido como se describe en la presente se puede administrar en combinación con otros agentes también, tales como, por ejemplo otras proteínas o polipéptidos o varios agentes farmacéuticamente activos, incluyendo agentes quimioterapéuticos contra la infección por *M. tuberculosis*. De hecho, virtualmente no hay límite a otros componentes que también se pueden incluir, dado que los agentes adicionales no causan un efecto adverso significativo al contacto con las células o tejidos huéspedes. Las composiciones así pueden ser suministradas junto con otros varios agentes como se requiera en el caso particular. Estas composiciones se pueden purificar de células huéspedes u otras fuentes biológicas, o alternativamente se pueden sintetizar químicamente como se describe en la presente. Igualmente, estas composiciones además pueden comprender composiciones derivadas de ARN o ADN.

La formulación de soluciones excipientes y vehículo farmacéuticamente aceptables es muy conocida para los expertos en la técnica, como lo es el desarrollo de dosis y regímenes de tratamiento adecuados para el uso de las composiciones particulares descritas en la presente en una variedad de regímenes de tratamiento, incluyendo por ejemplo, la administración y formulación oral, parenteral, intravenoso, intranasal, e intramuscular. Otras vías de administración incluyen vía las superficies de las mucosas.

Típicamente, las formulaciones que comprenden una cantidad terapéuticamente efectiva suministran aproximadamente 0,01 microgramo hasta aproximadamente 1000 microgramos de polipéptido de Rv3616c modificado por administración, más típicamente aproximadamente 0,1 microgramo hasta aproximadamente 100 microgramos de polipéptido por administración (por ejemplo de 0,5 a 50 microgramos). Con respecto a las composiciones de polinucleótidos, estas típicamente suministran aproximadamente 10 microgramos hasta aproximadamente 20 miligramos del polinucleótido de la invención por administración, más típicamente aproximadamente de 0,1 miligramo hasta aproximadamente 10 miligramos del polinucleótido de la invención por administración.

Naturalmente, la cantidad de compuesto(s) activo(s) en cada composición terapéuticamente útil puede prepararse de una manera que se obtendrá una dosis adecuada en cualquier dosis unitaria dada del compuesto. Los factores como la solubilidad, biodisponibilidad, vida media biológica, vía de administración, vida en el estante del producto, así como otras consideraciones farmacológicas serán contempladas por un experto en la técnica de preparar estas formulaciones farmacéuticas, y como tal, puede ser deseable una variedad de dosificaciones y regímenes de tratamiento.

1. ADMINISTRACIÓN ORAL

En ciertas aplicaciones, las composiciones farmacéuticas dadas a conocer en la presente se pueden administrar por vía oral a un animal. Como tal, estas composiciones se pueden formular con un diluyente inerte o con un vehículo comestible asimilable, o pueden incluirse en cápsulas de gelatina de corteza dura o blanda, o pueden comprimirse en tabletas, o pueden incorporarse directamente con la comida de la dieta.

Los compuestos activos se pueden incorporar con excipientes y ser usados en la forma de tabletas ingeribles, tabletas bucales, trociscos, cápsulas, elíxires, suspensiones, jarabes obleas, y similares (Mathiowitz y col., 1997; Hwang y col., 1998; Patente de Estados Unidos 5,641,515; Patente de Estados Unidos 5,580,579 y Patente de

Estados Unidos 5,792,451, cada una específicamente incorporada en la presente mediante referencia por entero). Las tabletas, trociscos, píldoras, cápsulas y similares también pueden contener lo siguiente: un aglutinante, como goma de tragacanto, acacia, almidón de maíz, o gelatina; excipientes, tales como di-fosfato de calcio; un agente desintegrante, tal como almidón de maíz, almidón de papa, ácido algínico y similares; un lubricante, tal como estearato de magnesio; y un agente edulcorante, tal como sacarosa, lactosa o sacarina se puede añadir o un agente saborizante, tal como menta, aceite de gaulteria o cereza. Cuando la unidad forma de dosis unitaria es una cápsula, puede contener, además de los materiales del tipo anterior, un vehículo líquido. Otros distintos materiales pueden estar presentes como recubrimientos o para modificar de otra manera la forma física de la unidad de dosificación. Por ejemplo, las tabletas, píldoras o cápsulas se pueden recubrir con goma laca, azúcar o ambas. Un jarabe o elíxir puede contener el compuesto activo, sacarosa como un agente edulcorante, metil- y propil-parabenos como conservadores, un tinte y saborizante, tal como sabor a cereza o a naranja. Desde luego, cualquier material usado para preparar cualquier forma de dosis unitaria deberá ser farmacéuticamente pura y sustancialmente no tóxico en las cantidades empleadas. Además, los compuestos activos se pueden incorporar en preparaciones y formulaciones de liberación sostenida.

Para la administración oral las composiciones de la presente invención se pueden incorporar alternativamente con uno o más excipientes de administración en formulación administrada oralmente o sublingualmente en forma de un enjuague bucal, dentífrico, tableta bucal, aspersor oral. Por ejemplo, se puede preparar un enjuague bucal incorporando el ingrediente activo en la cantidad adecuada en un solvente apropiado, tal como solución de borato de sodio (Solución de Dobell). Alternativamente, el ingrediente activo puede incorporarse en una solución oral tal como una que contenga borato de sodio, glicerina y bicarbonato de potasio, o disperso en un dentífrico, o añadido en un cantidad terapéuticamente efectiva a una composición que puede incluir agua, aglutinantes, abrasivos, agentes saborizantes, agentes formadores de espuma y humectantes. Alternativamente las composiciones se pueden modelar en una forma de tableta o solución que se pueda colocar bajo la lengua o disolverse de otro modo en la boca.

2. ADMINISTRACIÓN INYECTABLE

En general puede ser deseable suministrar las composiciones farmacéuticas dadas a conocer en la presente por vía parenteral, intravenosa, intramuscular, intradérmica, o hasta intraperitoneal. como se describe en la patente de Estados Unidos 5.543.158; patente de Estados Unidos 5.641.515 y patente de Estados Unidos 5.399.363.

Las soluciones de los compuestos activos son de bases libres o sales farmacológicamente aceptables se pueden preparar en agua convenientemente mezclada con un tensoactivo, tal como hidroxi-propil-celulosa. Las dispersiones también se pueden preparar en glicerol, glicoles de polietileno líquido, y mezclas de los mismos y en aceites. Bajo condiciones ordinarias de almacenamiento y uso, estas preparaciones contienen un conservador para evitar el crecimiento de microorganismos.

Las formas farmacéuticas adecuadas para uso inyectable incluyen soluciones acuosas o dispersiones estériles y polvos estériles para la preparación extemporánea de soluciones o dispersiones inyectables (Patente de Estados Unidos 5,466,468, específicamente incorporada en la presente mediante referencia por entero).

En todos los casos la forma debe ser estéril y debe ser fluida al grado que fácilmente exista la posibilidad de usar una jeringa. Debe ser estable bajo condiciones de fabricación y almacenamiento y debe conservarse contra la acción contaminante de los microorganismos tales como bacterias y hongos. El vehículo puede ser un solvente o medio de dispersión que contenga, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol, y polietilenglicol líquido y similares), mezclas adecuadas de los mismos, y/o aceites vegetales. La fluidez adecuada se puede mantener, por ejemplo, mediante el uso de un recubrimiento, tal como lecitina, por el mantenimiento del tamaño de partícula requerido en el caso de dispersión y por el uso de tensoactivos. La prevención de la acción de microorganismos se puede facilitar por varios agentes antibacterianos y anti hongos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico, timerosal, y similares. En muchos casos, puede ser preferible incluir agentes isotónicos, por ejemplo azúcares o cloruro de sodio. La absorción prolongada de las composiciones inyectables se puede producir mediante el uso en las composiciones de agentes que retrasen la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Para la administración parenteral en una solución acuosa, por ejemplo, la solución deberá ser convenientemente regulada, si es necesario, y el diluyente líquido primero volverse isotónico con suficiente solución salina o glucosa. Estas soluciones acuosas particulares son especialmente adecuadas para la administración intravenosa, intramuscular, subcutánea e intraperitoneal. En conexión a esto, un medio acuoso estéril que se pueda emplear será conocido por los expertos en la técnica a la luz de la presente divulgación. Por ejemplo, una dosis se puede disolver en 1 mililitro de solución de NaCl isotónica y ya sea añadirla a 1000 mililitros de fluido hipodermoclisis o inyectarla en el sitio propuesto de infusión (ver, por ejemplo, Remington's Pharmaceutical Sciences, 15ª Edición, páginas 1035-1038 y 1570-1580). Alguna variación en dosificación necesariamente ocurrirá dependiendo de la condición del sujeto que está siendo tratado. La persona responsable de la administración, en cualquier caso, determinará la dosis apropiada para el sujeto individual. Más aún, para la administración humana, las preparaciones deberán satisfacer la esterilidad, pirogenicidad, y las normas generales de seguridad y pureza requeridas por las normas de la Oficina de Productos Biológicos de la Administración de Alimentos y Fármacos (FDA) de los Estados Unidos de América.

Las soluciones inyectables estériles se preparan incorporando los compuestos activos en la cantidad requerida en el solvente apropiado con varios de los demás ingredientes enumerados anteriormente, como se requiera, seguido por esterilización filtrada. Generalmente, las dispersiones se preparan incorporando los distintos ingredientes activos esterilizados en un vehículo estéril que contiene el medio de dispersión básico y los demás ingredientes requeridos de los enumerados anteriormente. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos preferidos de preparación son las técnicas de secado al vacío y secado congelado que producen un polvo del ingrediente activo más cualquier ingrediente adicional deseado a partir de una solución previamente estéril por filtración del mismo.

Las composiciones dadas a conocer en la presente se pueden formular en una forma neutra o de sal. Las sales farmacéuticamente aceptables, incluyen las sales de adición ácida (formadas con los grupos amono libres de la proteína) y que se forman con ácidos inorgánicos tales como, por ejemplo, ácidos clorhídrico o fosfórico, o ácidos orgánicos tales como el ácido acético, oxálico, tartárico, mandélico, y similares. Las sales formadas con los grupos carboxilo libres también se pueden derivar de las bases inorgánicas tales como, por ejemplo, hidróxidos de sodio, de potasio, de amonio, de calcio, o férrico, y bases orgánicas como isopropil-amina, trimetil-amina, histidina, procaína y similares. Después de la formulación, las soluciones se administrarán de una manera compatible con la formulación de dosis y en tal cantidad como sea terapéuticamente efectiva. Las formulaciones se administran fácilmente en una variedad de formas de dosis tales como soluciones inyectables, cápsulas de liberación de fármacos y similares.

Como se usa en el presente documento, “vehículo” incluye cualquiera y todos los solventes, medios de dispersión, vehículos, recubrimientos, diluyentes, agentes antibacterianos y antifúngicas, agentes isotónicos y retardantes de la absorción, reguladores, soluciones vehículo, suspensiones, coloides y similares. El uso de estos medios y agentes para sustancias farmacéuticamente activas es muy conocido en la técnica. Siempre y cuando algún medio o agente sea incompatible con el ingrediente activo, se contempla su uso en composiciones terapéuticas. También se pueden incorporar ingredientes activos suplementarios en las composiciones.

La frase “farmacéuticamente aceptable” se refiere a entidades moleculares y composiciones que no producen una reacción alérgica o adversa similar cuando se administra a un humano. La preparación de una composición acuosa que contiene una proteína como un ingrediente activo es muy conocida en la técnica. Típicamente, estas composiciones se preparan como inyectables, ya sea como soluciones o suspensiones líquidas; también se pueden preparar formas sólidas adecuadas para solución en, o suspensión en, líquido antes de una inyección. La preparación también se puede emulsionar.

3. ADMINISTRACIÓN NASAL Y BUCAL

En ciertas realizaciones, las composiciones farmacéuticas se pueden administrar mediante aspersores intranasales, aspersores bucales, inhalación, y/u otros vehículos de administración en aerosol. Los métodos para suministrar composiciones de genes, ácidos nucleicos, y péptidos directamente a los pulmones por ejemplo vía aspersores en aerosol nasal y bucal se han descrito, por ejemplo, en la Patente de Estados Unidos 5,756,353 y Patente de Estados Unidos 5,804,212.

Igualmente, la administración de fármacos usando resinas de micropartículas intranasales (Takenaga y col., 1998) y compuestos de lisofosfatidil-glicerol (patente de Estados Unidos 5.725.871, específicamente incorporada en la presente mediante referencia por entero) también son muy conocidos en la técnica farmacéutica. Asimismo, la administración de fármacos transmucosales en forma de una matriz de soporte de poli-tetrafluoro-etileno se describe en la Patente de Estados Unidos 5,780,045 (específicamente incorporada en la presente mediante referencia por entero).

4. ADMINISTRACIÓN MEDIADA POR LIPOSOMAS, NANOCÁPSULAS, Y MICROPARTÍCULAS

En ciertas realizaciones, los inventores contemplan el uso de liposomas, nanocápsulas, micropartículas, micro esferas, partículas de lípidos, vesículas y similares para la introducción de las composiciones de la presente invención en células huéspedes adecuadas. en particular, las composiciones de la presente invención se pueden formular para su administración encapsuladas en una partícula de lípido, un liposoma, una vesícula, una nanoesfera, o una nanopartícula o similar.

Estas formulaciones se pueden preferir para la introducción de formulaciones farmacéuticamente aceptables de los ácidos nucleicos o construcciones descritos en la presente. La formación y el uso de liposomas es conocido generalmente para los expertos en la técnica (véase por ejemplo, Couvreur y col., 1977; Couvreur, 1988; Lasic, 1998; que describe el uso de liposomas y nanocápsulas de la terapia de antibiótico dirigida para infecciones y enfermedades bacterianas intracelulares). Recientemente, se descubrieron liposomas con estabilidad en suero y medios tiempos de circulación mejorados (Gabizon y Papahadjopoulos, 1988; Allen y Choun, 1987; Patente de Estados Unidos 5,741,516, específicamente incorporada en la presente mediante referencia por entero). Además, han sido revisados varios métodos de liposomas y preparaciones parecidas a liposomas como vehículos potenciales de fármacos (Takakura, 1998; Chandran y col., 1997; Margalit, 1995; Patente de Estados Unidos 5,567,434; Patente de Estados Unidos 5,552,157; Patente de Estados Unidos 5,565,213; Patente de Estados Unidos 5,738,868 y Patente de Estados Unidos 5,795,587, cada una específicamente incorporada en la presente mediante referencia

por entero).

Los liposomas se han usado con éxito con muchos tipos de células que normalmente son resistentes a la transfección mediante otros procedimientos que incluyen suspensiones con células T, cultivos de hepatocitos primarios y células PC 12 (Renneisen y col., 1990; Muller y col., 1990). En adición, los liposomas están libres de restricciones en la longitud de ADN que son típicas de los sistemas de administración basado en virus. Los liposomas se han usado con efectividad para introducir genes, fármacos (Heath y Martin, 1986; Heath y col., 1986; Balazsovits y col., 1989; Fresta y Puglisi, 1996), agentes radioterapéuticos (Pikul y col., 1987), enzimas (Imaizumi y col., 1990a; Imaizumi y col., 1990b), virus (Faller y Baltimore, 1984), factores de la transcripción y efectores aloestéricos (Nicolau y Gersonde, 1979) en una variedad de líneas celulares cultivadas y animales. En adición, varios ensayos clínicos exitosos que examinaron la efectividad de la administración de fármacos mediados por liposomas se han completado (Lopez-Berestein y col., 1985a; 1985b; Coune, 1988; Sculier y col., 1988). Además, varios estudios sugieren que el uso de liposomas no se asocia con las respuestas autoinmunes, la toxicidad o localización gonadal después de la administración sistémica (Mori y Fukatsu, 1992).

Los liposomas se forman de fosfolípidos que están dispersos en un medio acuoso y espontáneamente forman vesículas bicapa concéntricas multilamelar (también denominadas vesículas multilamelares (MLVs). Las vesículas multilamelares generalmente tienen diámetros de 25 nanómetros a 4 micras. La sonicación de las vesículas multilamelares da como resultado la formación de pequeñas vesículas unilamelares (SUVs) con diámetros en el intervalo de variación de 200 a 500 Å, que contienen una solución acuosa en el núcleo.

Los liposomas se parecen a las membranas celulares y se contemplan para su uso en relación con la presente invención como vehículos para las composiciones de péptidos. Son ampliamente adecuados como sustancias solubles tanto en agua como en lípidos pueden ser atrapadas, es decir, en los espacios acuosos y dentro de la bicapa misma, respectivamente. Es posible que los liposomas que llevan fármacos aún puedan emplearse para administración específica al sitio o agentes activos modificando selectivamente la formulación liposomal.

Además de las enseñanzas de Couvreur y col. (1977; 1988), se puede utilizar la siguiente información para generar formulaciones liposomales. Los fosfolípidos pueden formar una variedad de estructuras distintas de los liposomas cuando se dispersan en agua, dependiendo de la razón molar del lípido contra el agua. A razones bajas el liposoma es la estructura preferida. Las características físicas de los liposomas dependen en el pH, la fuerza iónica y la presencia de cationes divalentes. Los liposomas pueden mostrar baja permeabilidad a las sustancias iónicas y polares, pero a temperaturas elevadas sufren una transición de fases que altera marcadamente su permeabilidad. La transición de fases involucra un cambio de una estructura ordenada, estrechamente empaquetada, conocida como estado de gel, a una estructura menos ordenada, empaquetada sin apretar, conocida como el estado fluido. Esto ocurre a una temperatura de transición de fases característica y da como resultado un aumento en la permeabilidad a iones, azúcares y fármacos.

En adición a la temperatura, la exposición a las proteínas puede alterar la permeabilidad de los liposomas. Ciertas proteínas solubles, tales como el citocromo c, enlazan, deforman y penetran la bicapa, causando mediante esto cambios en la permeabilidad. El colesterol inhibe esta penetración de proteínas, evidentemente empaquetando los fosfolípidos más estrechamente. Esto se interpreta como que las formaciones de liposomas más útiles para la administración de antibiótico e inhibidor contendrán colesterol.

La capacidad para atrapar solutos varía entre diferentes tipos de liposomas. Por ejemplo las vesículas multilamelares son moderadamente atrapando solutos, pero las vesículas unilamelares pequeñas son extremadamente ineficientes. Las vesículas unilamelares pequeñas ofrecen la ventaja de homogeneidad y reproducibilidad en distribución de tamaño, sin embargo, y un compromiso ente la eficiencia de atrapado y tamaño es ofrecido por las vesículas unilamelares grandes (LUVs). Estas se preparan por la evaporación de éter y son de tres a cuatro veces más eficientes en atrapar solutos que las vesículas multilamelares.

En adición a las características de los liposomas, un determinante importante para atrapar compuestos es las propiedades fisicoquímicas del mismo compuesto. Los compuestos polares son atrapados en espacios acuosos y los compuestos no polares se enlazan a la bicapa de lípido de la vesícula. Los compuestos polares son atrapados en los espacios acuosos y los compuestos no polares se enlazan a la bicapa de lípido de la vesícula. Los compuestos polares son liberados a través de la permeación o cuando se rompe la bicapa, pero los compuestos no polares permanecen afiliados con la bicapa a menos que se desorganice por la temperatura o exposición a lipoproteínas. Ambos tipos muestran las tasas de eflujo máximas a la temperatura de transición de fases.

Los liposomas interactúan con las células por medio de cuatro mecanismos diferentes: endocitosis por células fagocíticas del sistema retículo endotelial tales como macrófagos y neutrófilos; adsorción a la superficie celular, ya sea por fuerzas electrostáticas o hidrofóbicas débiles no específicas, o por interacciones específicas con componentes en la superficie celular; la fusión con la membrana celular de plasma por inserción de la bicapa de lípido del liposoma en la membrana de plasma, con liberación simultánea de contenido liposomal dentro del citoplasma; y por transferencia de lípidos liposomales a las membranas celular o subcelular, o viceversa, sin asociación alguna del contenido del liposoma. Frecuentemente es difícil determinar cuál mecanismo es operativo y más de uno puede operar al mismo tiempo.

El destino y disposición de los liposomas inyectados intravenosamente dependen de sus propiedades físicas, tales como tamaño, fluidez y carga de superficie. Pueden persistir en tejidos por horas o días, dependiendo de su composición, y sus vidas medias en la sangre varían de minutos a varias horas. Los liposomas más grandes, tales como las vesículas multilamelares y las vesículas unilamelares grandes, se absorber rápidamente por las células fagocíticas del sistema retículo-endotelial, pero la fisiología del sistema circulatorio restringe la salida de estas especies grandes en la mayoría de los sitios. Pueden salir solamente en lugares donde existen grandes aberturas o poros en el endotelio apilar, tal como las sinusoides del hígado o bazo. Así, estos órganos son el sitio predominante de absorción. Por otro lado, las vesículas unilamelares pequeñas muestran una distribución de tejido más amplia, pero todavía quedan secuestradas en gran medida en el hígado y bazo. En general, este comportamiento en vivo limita el destino potencial de los liposomas solamente a esos órganos y tejidos accesibles a su tamaño grande. Estos incluyen la sangre, el hígado, el bazo, la médula ósea y los órganos linfoides.

La focalización generalmente no es una limitación en términos de la presente invención. Sin embargo, si se desea focalización específica, hay métodos disponibles para que esto se lleve a cabo. Los anticuerpos se pueden usar para enlazar la superficie del liposoma y dirigir el anticuerpo y su contenido de fármaco a los receptores antigénicos específicos en una superficie de tipo de célula particular. Los determinantes de carbohidratos (los componentes en la superficie celular de glicoproteínas o glicolípidos que representan un papel en el reconocimiento, interacción y adhesión de célula a célula) pueden ser usados como sitios de reconocimiento ya que tienen el potencial para dirigir los liposomas a tipos de células particulares. En su mayor parte, se contempla que se usaría la inyección intravenosa de las preparaciones liposomales, pero otras vías de administración también son concebibles.

Alternativamente, la invención proporciona formulaciones de nanocápsulas farmacéuticamente aceptables de las composiciones de la presente invención. Las nanocápsulas generalmente pueden atrapar compuestos de una manera estable y reproducible (Henry-Michelland y col., 1987; Quintanar-Guerrero y col., 1998; Douglas y col., 1987). Para evitar los efectos secundarios debidos a la sobrecarga polimérica intracelular, estas partículas ultrafinas (de un tamaño aproximado de 0,1 μm) deberán diseñarse usando polímeros capaces de ser degradados en vivo. Las nanopartículas de cianoacrilato de polialquilo biodegradables que satisfacen estos requerimientos se contemplan para su uso en la presente invención. Estas partículas se pueden hacer fácilmente, como se describe en (Couvreur y col., 1980; 1988; zur Muhlen y col., 1998; Zambaux y col. 1998; Pinto-Alphandry y col., 1995 y Patente de Estados Unidos 5,145,684, específicamente incorporadas en la presente mediante referencia por entero).

También se pueden usar parches dérmicos para la administración transcutánea.

COMPOSICIONES INmunogÉnicAS

En ciertas realizaciones de la presente invención, se proporcionan composiciones inmunogénicas. Las composiciones inmunogénicas comprenderán una o más secuencias de la Rv3616c modificadas (polipéptidos o polinucleótidos) como los comentados anteriormente, en combinación con un inmunoestimulante.

Las composiciones inmunogénicas también pueden comprender una proteína de fusión o un polinucleótido que codifica una proteína de fusión, en combinación con un vehículo o excipiente farmacéuticamente aceptable.

Un inmunoestimulante puede ser cualquier sustancia que aumenta o potencia una respuesta inmunitaria (anticuerpo y/o mediado pro célula) a un antígeno exógeno. Los ejemplos de los inmuno-estimulantes incluyen a los adyuvantes.

La preparación de composiciones inmunogénicas generalmente se describe en, por ejemplo, Powell y Newman, Editores, Vaccine Design (en enfoque de subunidad y adyuvante) (1995). Las composiciones farmacéuticas y las composiciones inmunogénicas dentro del alcance de la presente invención también pueden contener otros compuestos, que pueden ser biológicamente activos o inactivos. Por ejemplo, pueden estar presentes una o más porciones inmunogénicas de otros antígenos de *M. tuberculosis*, ya sea incorporadas en un polipéptido de fusión o como un componente separado, dentro de la composición farmacéutica o inmunogénica.

Las composiciones inmunogénicas ilustrativas pueden contener un polinucleótido (por ejemplo ADN) que codifica uno o más de los polipéptidos como se describe en lo anterior, de manera que el polipéptido es generado en el sitio /provocando mediante esto una respuesta inmunitaria). Como se notó anteriormente, el ADN puede estar presente dentro de cualquier variedad de sistemas de liberación conocidos para los expertos en la técnica, incluyendo sistemas de expresión de ácido nucleico, sistemas de expresión de bacterias y virus. Muchas técnicas de liberación de genes son muy conocidas en la técnica, como las descritas por Rolland, Crit. Rev. Therap. Drug Carrier Systems 15: 143-198 (1998), y las referencias citadas en la misma. Los sistemas de expresión de ácido nucleico apropiados contienen las secuencias de ADN necesarias para la expresión en el paciente (tal como un promotor y señal de terminación adecuadas). Los sistemas de liberación bacterianos incluyen la administración de una célula huésped de bacteria (por ejemplo, una cepa de *Mycobacterium*, Bacillus o Lactobacillus, incluyendo el Bacillus-Calmette-Guerin o el Lactococcus lactis) que expresa al polipéptido (por ejemplo en su superficie o secreta el polipéptido) (véase, por ejemplo, Ferreira, y col., An Acad Bras Cienc (2005) 77: 113-124; y Raha, y col., Appl Microbiol Biotechnol (2005) PubMedID 15635459). En una realización preferida, el ADN puede ser introducido usando un sistema de expresión viral (por ejemplo, vacuna u otro virus de viruela, retrovirus o adenovirus), lo cual puede incluir el uso de un virus competente de replicación no patógeno (defectuoso). Los sistemas adecuados se dan a conocer, por ejemplo, en

Fisher-Hoch y col., Proc. Natl. Acad. Sci. USA 86: 317-321 (1989) ; Flexner y col., Ann. N.Y. Acad. Sci. 569: 86-103 (1989) ; Flexner y col., Vaccine 8: 17-21 (1990) ; Patentes de los Estados Unidos de América Números 4,603,112, 4,769,330, y 5,017,487; WO 89/01973; Patentes Números US4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, Biotechniques 6: 616-627 (1988) ; Rosenfeld y col., Science 252: 431-434 (1991) ; Kolls y col., Proc. Natl. Acad. Sci. USA 91: 215-219 (1994) ; Kass-Eisler y col., Proc. Natl. Acad. Sci. USA 90: 11498-11502 (1993) ; Guzman y col., Circulation 88: 2838-2848 (1993) ; y Guzman y col., Cir. Res. 73: 1202-1207 (1993). Las técnicas para incorporar ADN en estos sistemas de expresión son muy conocidas para las personas con experiencia ordinaria en la técnica. El ADN también puede ser “no protegido”, como se describe, por ejemplo, en Ulmer y col., Science 259: 1745-1749 (1993) y revisado por Cohen, Science 259: 1691-1692 (1993). La absorción del ADN no protegido puede aumentarse recubriendo el ADN sobre cuentas biodegradables, que son eficientemente transportadas dentro de las células. Será evidente que una composición inmunogénica puede comprender tanto un componente de polinucleótido como un polipéptido. Esta composición inmunogénica puede proporcionar una respuesta inmunitaria aumentada.

Será evidente que una composición inmunogénica puede contener sales farmacéuticamente aceptables de los polinucleótidos y polipéptidos proporcionados en la presente. Estas sales se pueden preparar de bases no tóxicas farmacéuticamente aceptables, incluyendo bases orgánicas (por ejemplo, sales de aminas primarias, secundarias y terciarias y aminoácidos básicos) y bases inorgánicas (por ejemplo, sales de sodio, potasio, litio, amonio, calcio y magnesio).

Aunque se puede emplear cualquier vehículo conocido para las personas con experiencia ordinaria en la técnica en las composiciones inmunogénicas de la presente invención, el tipo de vehículo variará dependiendo del modo de administración, incluyendo por ejemplo la administración tópica, oral, nasal, intravenosa, intracraneal, intraperitoneal, subcutánea o intramuscular. Para la administración parenteral, tal como la inyección subcutánea, el vehículo de preferencia comprende agua, solución salina, alcohol, una cera o un regulador. Para la administración oral, se puede emplear cualquiera de los vehículos anteriores o un vehículo sólido tal como manitol, lactosa, almidón, estearato de magnesio, sacarina de sodio, talco, celulosa, glucosa, sacarosa, y carbonato de magnesio. También se pueden emplear microesferas biodegradables (por ejemplo polilactato, poliglicolato) como vehículos de las composiciones farmacéuticas de la presente invención. Las microesferas biodegradables adecuadas se dan a conocer, por ejemplo en las Patentes de los Estados Unidos de América Números 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 y 5,942,252. También se puede emplear un vehículo que comprende los complejos de proteína en partículas descritos en la Patente de Estados Unidos 5,928,647, los cuales son capaces de inducir respuestas de linfocitos T citotóxicos restringidos clase I en un huésped.

Estas composiciones también pueden comprender reguladores (por ejemplo, solución salina regulada neutra o solución salina regulada de fosfato), carbohidratos (por ejemplo, glucosa, manosa, sacarosa o dextranos), manitol, proteínas, polipéptidos o aminoácidos tales como glicina, antioxidantes, bacteriostatos, agentes quelantes tales como EDTA o glutatión, adyuvantes (por ejemplo, hidróxido de aluminio), solutos que vuelven la formulación isotónica, hipotónica o débilmente hipertónica con la sangre de un receptor, agentes de suspensión, agentes espesantes y/o conservadores. Alternativamente, las composiciones de la presente invención se pueden formular como liofilizados. Los compuestos pueden también ser encapsulados con liposomas usando tecnología muy conocida.

Cualquiera de una variedad de inmunoestimulantes se puede emplear en las composiciones inmunogénicas de la presente invención. Por ejemplo, se puede incluir un adyuvante. Un adyuvante se refiere a los componentes en una vacuna o composición terapéutica que aumenta la respuesta inmunitaria específica al antígeno (véase, por ejemplo, Edelman, AIDS Res. Hum Retrovirus 8: 1409-1411 (1992)). Los adyuvantes inducen las respuestas inmunitarias de la respuesta tipo Th-1 y tipo Th-2. Las citocinas del tipo Th-1 (por ejemplo, IFN-γ, IL-2, y IL-12) tienden a favorecer la inducción de la respuesta inmunitaria mediada por la célula a un antígeno administrado, mientras que las citocinas del tipo Th-2 (por ejemplo, IL-4, IL-5, Il-6, IL-10) tienden a favorecer la inducción de las respuestas inmunitarias humorales.

Dentro de las composiciones inmunogénicas proporcionadas en la presente, la composición adyuvante de preferencia se diseña para inducir una respuesta inmunitaria predominantemente del tipo Th1.

Las composiciones adyuvantes adecuadas incluyen emulsiones de aceite en agua. En particular una emulsión de aceite en agua que comprende de 0,5 a 10 miligramos de aceite metabolizable (por ejemplo, escualeno), 0,5 a 11 miligramos de tocol (por ejemplo alfa-tocoferol) y 0,1 a 4 miligramos de agente emulsionante (por ejemplo mono-oleato de polioxietilen-sorbitán) por dosis humana. Véase, por ejemplo, la Publicación Internacional Número WO2008/043774 la cual se incorpora mediante referencia por la presente.

Un adyuvante alternativo comprende una fracción de saponina inmunológicamente activa derivada de la corteza de Quillaja Saponaria Molina (tal como la fracción purificada HPLC conocida como QS21, como se describe en la Patente de Estados Unidos 5,057,540) presentada en la forma de un liposoma y un lipopolisacárido (tal como monofosforilo-lípido A 3-des-O-acilado). Estas composiciones además pueden comprender un esterol (por ejemplo colesterol), de manera que en donde la proporción de saponina: esterol es de 1: 1 a 1: 100 peso/peso (por ejemplo la proporción de saponina: esterol es de 1: 1 a 1: 10 peso/peso). Particularmente adecuados son aquellos

adyuvantes en donde el QS21 y el monofosforilo-lípido A 3-des-O-acilado están presentes en una proporción de QS21: 3D-MPL de 1: 1 peso /peso y ambos están presentes en una dosis humana a un nivel por debajo de 30 microgramos. Estas composiciones de adyuvantes se describen, por ejemplo, en la Publicación Internacional Número WO2007/068907 y en la Patente de Estados Unidos 2008279926, las cuales se incorporan mediante la presente por referencia.

Otros sistemas adyuvantes de interés incluyen los basados en las sales de aluminio junto con el lipopolisacárido monofosforilo-lípido A 3-des-O-acilado. El antígeno y el monofosforilo-lípido A 3-des-O-acilado se pueden co-adsorber a las mismas partículas de sal metálica o se pueden adsorber a distintas partículas de sal metálica. Véanse, por ejemplo, WO00/23105, US7357936 y US2008 0226672A1, las cuales se incorporan mediante la presente por referencia, las cuales describen las composiciones inmunogénicas que comprenden un enlace de antígeno a una primera partícula de sal metálica (en particular fosfato de aluminio o hidróxido de aluminio) y monofosforilo-lípido A 3-des-O-acilado el cual se enlaza a una segunda partícula de sal metálica (en particular fosfato de aluminio o hidróxido de aluminio).

Cualquier composición inmunogénica proporcionada en la presente se puede preparar usando métodos muy conocidos que dan como resultado una combinación de antígeno, aumentador de la respuesta inmunitaria y un vehículo o excipiente adecuado (conforme sea necesario).

Cualquiera de una variedad de vehículos de liberación se puede emplear dentro de las composiciones farmacéuticas y composiciones inmunogénicas para facilitar la producción de una respuesta inmunitaria específica de antígeno.

Los vehículos de liberación incluyen células que presentan antígeno (APCs), tales como células dendríticas, macrófagos, células B, monocitos y otras células que se pueden diseñar genéticamente para ser APCs eficientes. Estas células, pueden, pero no necesitan, ser modificadas genéticamente para aumentar la capacidad para presentar el antígeno, para mejorar la activación y/o mantenimiento de la respuesta de la célula T y/o ser inmunológicamente compatible con el receptor (es decir, haplotipo HLA emparejado). Las APCs generalmente se pueden aislar de cualquier variedad de fluidos y órganos biológicos y pueden ser células autólogas, alogénicas, singénicas o xenogénicas.

Ciertas realizaciones de la presente invención usan células dendríticas o progenitoras de las mismas como células que presentan antígeno. Las células dendríticas son APCs muy potentes (Banchereau y Steinman, Nature 392: 245-251 (1998)) y se ha demostrado que son efectivas como adyuvante fisiológico para provocar inmunidad profiláctica o terapéutica (véase Timmerman y Levy, Ann. Rev. Med. 50: 507-529 (1999)). En general, las células dendríticas se pueden identificar con base en su forma típica (estrellada in situ, con procesos citoplásmicos marcados (dendritas) visibles *in vitro*), su capacidad para absorber, procesar y presentar antígenos con alta eficiencia y su capacidad para activar respuestas de células T sencillas. Las células dendríticas pueden, desde luego, ser diseñadas genéticamente para expresar receptores de superficie celular específicos o ligandos que no se encuentran comúnmente en las células dendríticas in vivo o ex vivo, y estas células dendríticas modificadas son contempladas por la presente invención. Como una alternativa a las células dendríticas, se pueden usar células dendríticas cargadas de antígeno de vesículas secretadas (llamadas exosomas) dentro de una composición inmunogénica (véase Zitvogel y col., Nature Med. 4: 594-600 (1998)).

Las células dendríticas y progenitoras se pueden obtener de la sangre periférica, médula ósea, ganglios linfáticos, bazo, piel, sangre del cordón umbilical o cualquier otro ejido o fluido adecuado. Por ejemplo, las células dendríticas se pueden diferenciar ex vivo añadiendo una combinación de citocinas tales como GM-CSF, IL-4, IL-13 y/o TNFα a cultivos de monocitos cosechados de sangre periférica. De manera alternativa, las células positivas CD34 cosechadas de la sangre periférica, sangre del cordón umbilical o médula ósea se pueden diferenciar en células dendríticas añadiendo a medio de cultivo combinaciones de GM-CSF, IL-3, TNFα, ligando CD40, LPS, ligando flt3 y/o otros compuestos que inducen diferenciación, maduración y proliferación de las células dendríticas.

Las células dendríticas se categorizan convenientemente como células “inmaduras” y “maduras”, lo cual permite una manera simple de discriminar entre dos fenotipos bien caracterizados. Sin embargo, esta nomenclatura no deberá considerarse que excluya todas las etapas de diferenciación intermedias. Las células dendríticas inmaduras se caracterizan como APC con una alta capacidad para absorción y procesamiento de antígeno, lo cual se correlaciona con la alta expresión del receptor Fcγ y el receptor de manosa. El fenotipo maduro se caracteriza típicamente por una expresión más baja de estos marcadores, pero una expresión alta de las moléculas de la superficie celular responsables de la activación de las células T tales como las MHC de clase I y clase II, las moléculas de adhesión (por ejemplo, CD54 y CD11) y las moléculas co-estimulantes (por ejemplo, CD40, CD80, CD86 y 4-1BB).

Las APCs generalmente se pueden transfectar con un polinucleótido que codifica una proteína (o porción u otra variante de la misma) de manera que el polipéptido se expresa sobre la superficie de la célula. Esta transfección se puede llevar a cabo ex vivo, y se puede usar una composición farmacéutica o una composición inmunogénica que comprende estas células transfectadas, como se describe en la presente. Alternativamente, puede ser administrado a un paciente un vehículo de liberación de genes que se dirige a una célula dendrítica o que presenta antígeno, dando como resultado una transfección que ocurre en vivo. La transfección en vivo y ex vivo de las células dendríticas, por ejemplo, generalmente se puede realizar usando cualquier método de transfección conocido en la

técnica, tal como los descritos en la Publicación Internacional Número WO 97/24447, o el enfoque de cañón de genes descrito por Mahvi y col., Immunology y Cell Biology 75: 456-460 (1997). La carga de antígenos de las células dendríticas se puede lograr incubando células dendríticas o células progenitoras con el polipéptido, ADN (sin protección o con un vector plásmido) o ARN; o con bacteria o virus recombinante que expresa antígeno (por ejemplo, vectores de vacuna, viruela aviaria, adenovirus o lentivirus). Antes de la carga, el polipéptido se puede conjugar covalentemente a un asociado inmunológico que proporciona ayuda de célula T (por ejemplo, una molécula vehículoa). Alternativamente, se puede pulsar una célula dendrítica con un asociado inmunológico no conjugado, por separado o en la presencia del polipéptido.

Las composiciones inmunogénicas y las composiciones farmacéuticas se pueden presentar en contenedores de dosis unitaria o dosis múltiple, tales como ampollas o frascos sellados. Estos contenedores de preferencia se sellan herméticamente para conservar la esterilidad de la formulación hasta su uso. En general, las formulaciones se pueden almacenar como suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos. Alternativamente, una composición inmunogénica o una composición farmacéutica se pueden almacenar en una condición seca-congelada que requiere solamente la adición de un vehículo líquido estéril inmediatamente antes del uso.

En algunas realizaciones, un “cebado” o primera administración de una proteína Rv3616c modificada (incluyendo las proteínas de fusión), o polinucleótido que codifica a esa proteína, es seguido por una o más administraciones subsiguientes o de “refuerzo” de una proteína Rv3616c modificada (incluyendo las proteínas de fusión) o polinucleótido que codifica esa proteína (método de “cebado y refuerzo”). Por ejemplo, una primera administración con un polipéptido de Rv3616c modificado (incluyendo las proteínas de fusión) o polinucleótido que codifica la proteína es seguido por una o más administraciones subsiguientes de un polipéptido de Rv3616c codificado (incluyendo las proteínas de fusión) o un polinucleótido que codifica a este polipéptido.

En una realización, una primera administración con una proteína Rv3616c modificada o polinucleótido es seguida por una o más administraciones subsiguientes de una proteína Rv3616c modificada. En una realización, una primera administración con una proteína Rv3616c modificada o polinucleótido es seguida por una o más administraciones subsiguientes de un polinucleótido de Rv3616c modificado. Usualmente la primera administración o “cebado” y la segunda administración o “refuerzo” se dan con una separación de aproximadamente 2 a 12 semanas, o hasta de 4 a 6 meses. Las subsiguientes administraciones de “refuerzo” se dan aproximadamente con una separación de 6 meses, o tanto como una separación de 1, 2, 3, 4 o 5 años. El tratamiento de refuerzo convencional (por ejemplo, una administración de cebado de proteína seguida por una administración de refuerzo de proteína) también puede ser útil para prevenir o tratar la tuberculosis (por ejemplo prevenir o tratar la tuberculosis latente, en particular prevenir o retardar la reactivación de la tuberculosis).

DIAGNÓSTICOS

En otro aspecto, la presente invención proporciona métodos para usar uno o más de las proteínas Rv3616c modificadas descritas anteriormente para diagnosticar la tuberculosis, tal como la tuberculosis latente (por ejemplo usando ensayos basados en respuesta de células T o ensayos basados en anticuerpos de formato convencional).

Por ejemplo, se proporciona un método para determinar la infección por *M. tuberculosis* latente en un individuo, que comprende:

(a) obtener una muestra del individuo;
(b) poner en contacto la muestra con una proteína Rv3616c modificada;
(c) cuantificar la respuesta de la muestra.

La muestra puede ser, por ejemplo, sangre entera o células purificadas. De manera conveniente, la muestra contendrá células mononucleadas de sangre periférica (PBMC). En una realización de la invención el individuo será seropositivo. En una segunda realización de la invención el individuo será seronegativo.

Convenientemente el individuo no habrá sido previamente vacunado contra la infección por *M. tuberculosis* (por ejemplo convenientemente el individuo no habrá sido vacunado previamente con una BCG).

La respuesta de la muestra puede ser cuantificada por una variedad de medios conocidos para los expertos en la técnica, incluyendo el monitoreo de la proliferación de linfocitos o la producción de citocinas específicas o anticuerpos. Por ejemplo se puede usar el ELISPOT de células T para monitorear las citocinas tales como el interferón gamma (IFNγ), interleucina 2 (IL2) y la interleucina 5 (IL5). Se puede usar ELLISPOT de células B para monitorear la estimulación de antígenos específicos de *M. tuberculosis*. La respuesta celular puede también caracterizarse por el uso de teñido intra y extracelular y análisis mediante citómetro de flujo.

Los métodos para cuantificar una respuesta de proliferación de muestra incluyen:

(i) pulsar las células cultivadas con una radio-marca (por ejemplo timidina tritiada) y monitorear la absorción de tritio (por ejemplo, centelleo de gas);
(ii) etiquetamiento con succinimidil-éster de diacetato de carboxi-fluoresceína (CFSE) y monitoreo de fluorescencia de la división celular usando citometría de flujo.

La cuantificación de una muestra de respuesta de citocina incluye en particular monitorear la producción de interferón gamma.

Cuando se usan estos métodos de cuantificación, se puede definir una respuesta a un antígeno mediante una proporción señal a ruido (proporción S/N) de al menos 2: 1 (por ejemplo, al menos 3: 1 o al menos 5: 1).

En otro aspecto de la presente invención se proporcionan métodos para diagnosticar la infección por *M. tuberculosis* latente usando una prueba de piel. Como se usa en el presente documento, una “prueba de piel” es un ensayo realizado directamente sobre un paciente en el cual se mide una reacción de hipersensibilidad de tipo retardado (DTH) (tal como hinchazón, enrojecimiento o dermatitis) después de la inyección intradérmica de una proteína Rv153c como se describe anteriormente. Esta inyección se puede lograr usando cualquier dispositivo adecuado suficiente para poner en contacto combinaciones del antígeno con las células dérmicas del paciente, tal como una jeringa de tuberculina o una jeringa de 1 mililitro. La reacción se mide después de un periodo de tiempo, por ejemplo al menos 48 horas después de la inyección, especialmente de 48 a 72 horas.

La reacción DTH es una respuesta inmunitaria mediada por células, que es mayor en los pacientes que han sido expuestos anteriormente al antígeno de la prueba. La respuesta se puede medir visualmente, usando una regla. En general, una respuesta que es mayor que aproximadamente 0,5 centímetros de diámetro, especialmente mayor que aproximadamente 1,0 centímetros de diámetro, es una respuesta positiva, indicadora de una infección anterior de *M. tuberculosis* que puede o no manifestarse como una enfermedad activa.

Para su uso en una prueba de piel, el componente de proteína Rv3616c se formula convenientemente como una composición farmacéutica que contiene un vehículo fisiológicamente aceptable. De manera conveniente, el vehículo empleado en estas composiciones farmacéuticas es una solución salina con conservadores apropiados, tales como fenol y/o Tween 80MR.

La presente invención además proporciona kits para su uso dentro de cualquiera de los métodos de diagnóstico anteriores. Estos kits típicamente comprenden dos o más componentes necesarios para realizar una prueba de diagnóstico. Los componentes pueden ser compuestos, reactivos, contenedores y/o equipo. Por ejemplo, un contenedor dentro de un kit puede contener una proteína Rv3616c modificada. Esta proteína se puede proporcionar unida a un material de soporte. Estos kits también pueden contener un reactivo de detección que contenga un grupo reportero adecuado para la detección directa o indirecta del enlace de anticuerpo.

Otros kits de diagnóstico incluyen los diseñados para la detección de respuestas mediadas por células (que pueden por ejemplo) usarse en los métodos de diagnóstico de la presente invención). Estos kits típicamente comprenderán:

(i) un aparato para obtener una muestra de célula apropiada de un sujeto;
(ii) medios para estimular a la muestra de célula con un polipéptido de Rv3616c (o una variante del mismo, fragmentos inmunogénicos del mismo, o ADN que codifique estos polipéptidos) ;
(iii) medios para detectar o cuantificar la respuesta celular a la estimulación.

Los medios adecuados para cuantificar la respuesta celular incluyen un kit ELISPOT de células B o alternativamente un kit ELISPOT de células T, los cuales son conocidos por los expertos en la técnica.

Un kit posible comprende:

(a) un polipéptido de la invención; y
(b) un reactivo de detección adecuado para la detección directa o indirecta de enlace de anticuerpos.

De particular interés son los kits personalizados para cuantificar respuestas de células T:

Un kit de diagnóstico que comprende:

(a) un polipéptido de la invención; y
(b) un aparato suficiente para poner en contacto el polipéptido con las células dérmicas de un individuo.

Un kit de diagnóstico que comprende:

(a) un polipéptido de la invención;
(b) un aparato suficiente para poner en contacto al polipéptido con una muestra (por ejemplo sangre entera o más convenientemente PBMC) de un individuo, y
(c) medios para cuantificar la respuesta de las células T (por ejemplo la proliferación de la producción del IFN-gamma).

EJEMPLOS

Los siguientes ejemplos se proporcionan a manera de ilustración solamente y no a manera de limitación. Los expertos en la técnica fácilmente reconocerán una variedad de parámetros no críticos que podrían cambiarse o modificarse para producir esencialmente los mismos resultados.

Ejemplo 1 – Identificación de Rv3616c como un objetivo de vacuna de TB latente

El gen Rv3616c codifica para una proteína rica en alanina y glicina hipotética conservada.

Se seleccionó la Rv3616c basada en un análisis amplio de genoma de los genes de *Mycobacterium tuberculosis* asociados con el mantenimiento e inefectividad de la fase latente como en Murphy y Brown BMC.Infect. Dis. 2007 7: 84-99. Los objetivos de genes de latencia potencial en *Mycobacterium tuberculosis* fueron jerarquizados a través de un meta-análisis de bioinformática de conjuntos de bases de ADN de ámbito genoma publicados de la expresión de genes bacterianos bajo condiciones simuladas de latencia. La localización subcelular de las proteínas de *M. tuberculosis* codificados por genes, se llevó a cabo posteriormente en el genoma entero para identificar los objetivos de vacuna.

En resumen, las condiciones en los modelos de latencia fueron bastante variados de manera que se desarrolló un sistema de puntuación de cero a cinco para normalizar estos datos con base en dos criterios: 1) la relevancia de las condiciones experimentales al estado latente y 2) el orden de jerarquía de la expresión. La puntuación máxima para un conjunto de datos particular se ajustó con base en la relevancia potencial a la ocurrencia clínica de las infecciones de *M. tuberculosis* en fase latente. La Tabla 1 muestra los conjuntos de datos recolectados para la etapa 1 junto con las puntuaciones máximas ajustadas para cada conjunto de datos. Los conjuntos de datos adicionales sobre la esencialidad de los genes para el crecimiento se obtuvieron de los estudios publicados usando experimentos de eliminación genética basados en transposón (TraSH). Los genes que no tuvieron efecto sobre el crecimiento recibieron una calificación de cero.

Tabla 1 – Fuentes, modelos experimentales, y criterios de puntuación para eliminación genética en todo el genoma y la expresión de gen en el microarreglo de ADN de *M. tuberculosis* (esencialmente en fase de crecimiento).

Referencia	Modelo Experimental	Punto del Tiempo: Punto máximo[a]
Betts JC y col. Mol. Microbiol. 2002 43: 717-731	Inanición bajo O_2 controlado	96h: 3 24h: 2 4h: 1
Hampshire T y col. Tuberculosis.(Edimb.) 2004 84: 228-238	Merma de nutrientes bajo O2 controlado	62 y 75d: 5 49d: 4 18d: 2
Muttucumaru DG y col. Tuberculosis.(Edimb.) 2004 84: 239-246	Modelo de hipoxia de Wayne#	14d (NRP-2) : 4 7d (NRP-1) : 2
Voskuil MI y col. Tuberculosis.(Edimb.) 2004 84: 218-227	Modelo de hipoxia de Wayne#	30 y 80d: 5 14 y 20d: 4 10 y 12d: 3 6 y 8d: 2
Schnappinger D y col. J. Exp. Med. 2003 198: 693-704	Infección de macrófagos de ratón +/- γ-INF	24 y 48h: 5

(continuación)

Referencia	Modelo Experimental	Punto del Tiempo: Punto máximo[a]
Karakousis PC y col. J. Exp. Med. 2004 200: 647-657	Implante subcutáneo de fibra hueca en ratones	10d: 3
Talaat AM y col. Proc. Natl. Acad. Sci. EUA 2004, 101: 4602-4607	Infección de ratones. MTB cosechado de pulmónb	28d: 3
Sassetti CM y col. Mol. Microbiol. 2003 48: 77-84	Bibliotecas mutadas de TraSH cultivadas en medio sólido	14d: 5
Rengarajan J y col. Proc.Natl.Acad.Sci.U.S.A 2005, 102: 8327-8332	Infección de macrófagos de ratón, +/- γ-INF con bibliotecas mutadas de TraSH de *M. tuberculosis*	7d: 5
Sassetti CM y col. Proc.Natl.Acad.Sci.U.S.A 2003 100: 12989-12994	Ratones C57BL/6J infectados con bibliotecas mutadas de TraSH de *M. tuberculosis*	7, 14, 28 y 56d: 5

[a]Puntuación máxima basada en la relevancia como modelo de latencia; h = hora ; d = día.

[b]Proporción de *M. tuberculosis* a partir de pulmón de Balb/c contra MTB en cultivo aireado durante 28 días.

Wayne LG y Hayes LG Infect. Immun. 1996 64: 2062-2069

Etapa 2 – Para aplicar el segundo criterio, el orden de jerarquía de la expresión del gen, las puntuaciones de gen a partir de cada conjunto de datos se ordenaron del más alto al más bajo basado en la proporción de la expresión (doblar la expresión en la condición experimental contra las células en el cultivo líquido en fase log). El gen de puntuación más alta recibió la puntuación máxima para ese conjunto de datos particular (enlistado en la columna 3 de la Tabla 1, (por ejemplo. 5, 4…, 1 punto). La puntuación disminuyó en 0,005 puntos para cada gen en orden hasta cero, o se alcanzó el final del conjunto de datos. Así, cuando la puntuación máxima fue de 4 puntos, el gen en rango 100 recibiría una puntuación de 3.500. Para una puntuación máxima de 5 puntos, 1000 genes o 25 % del genoma de *M. tuberculosis* recibió una puntuación. Para experimentos donde se recolectaron los datos de múltiples puntos del tiempo, se usó como puntuación final la puntuación máxima a través de todos los puntos del tiempo.

En la etapa 3 las puntuaciones para cada gen en cada una de las condiciones experimentales se recolectaron en una base de datos de Acceso de Microsoft. Se añadieron los campos de Referencia para facilitar la jerarquización, tales como la Refseq ID, función de Genbank, nota de Genbank, clasificación de Tuberculist, y vínculos KEGG y Sanger Center. Al combinar los datos de diferentes estudios y fuentes, se logró una vista de consenso sobre los genes y sendas particulares más críticas para la supervivencia en el estado latente.

En la etapa 4, se derivó una lista jerarquizada de objetivos terapéuticos utilizando los 400 genes con mayor puntuación (~10 % del genoma) suplementados por análisis computacional y manual experto de sendas bioquímicas, enzimología, rastreabilidad de fármacos, homología con genes humanos y otros conocimientos anteriores. La gran mayoría de los genes con mayor puntuación vinieron del subconjunto donde dos o tres de los grupos se intersectan.

En la etapa 5, se llevó a cabo en el genoma entero la identificación de la localización subcelular de las proteínas de *M. tuberculosis* codificadas por genes. La heurística usada para la predicción de la membrana se describe en Chalker y col. J. Bacteriol. 2001 183: 1259-1268. Los perfiles promedio de hidropatía (H) (von Heijne G J. Mol. Biol. 1992 225: 487-494) se generaron usando valores de hidropatía GES (Engelman DM y col. Annu. Rev. Biophys.

Biophys. Chem. 1986 15: 321-353) ponderados usando una ventana trapezoide. Usando un proceso similar a las etapas iniciales del algoritmo TopPred II (Claros MG y col. Comput. Appl. Biosci. 1994 10: 685-686), se pronosticaron los segmentos de transmembrana helicoidal (TMS) para cada secuencia de péptidos seleccionando 19 aminoácidos centrados en el valor H más alto (MaxH), enmascarando estos de más consideración, y repitiendo el proceso hasta que no quedaron picos con una H de >0,5. Se asignaron los lugares subcelulares basándose en el valor MaxH pico, el número de segmentos con una H de >1,0, y distribución y valores H pico del TMS supuesto. Se eligió un corte MaxH de 1,15 para maximizar la discriminación entre dos Swiss Protein que liberaron 34 conjuntos de datos de prueba que contenían transmembrana y proteínas citoplásmicas, respectivamente (Boyd D y col. Protein Sci. 1998 7: 201-205). Las proteínas con una MaxH de <1,15 se clasificaron como citoplásmica, mientras que aquellas con una MaxH de >1,15 y al menos tres posibles TMS se clasificaron como proteínas de membrana. Se definieron proteínas ancladas como teniendo exactamente dos TMS, uno comenzando antes del aminoácido (aa) 35 y uno teniendo una H de >1,15 con la otra teniendo una H no menor que 0,5. Se usó específicamente SignalP con escenarios Gram positivos para M. bacterium para identificar las proteínas secretadas entre aquéllas clasificadas como ya sea citoplásmicas o “desconocidas” en el análisis heurístico (Nielsen H y col. Protein Eng. 1997 10: 1-6).

La Rv3616c quedó en un lugar muy alto como antígeno de vacuna de acuerdo con varios criterios:

(i) Rv3616c es regulada ascendentemente de manera consistente a través de todos los modelos de latencia. Entre el acervo entero de 3999 genes calificados en el meta-análisis, Rv3616c quedó en el cuartil superior de los genes sobre-expresados a través de todos los modelos de latencia. La puntuación regulada ascendente para Rv3616c fue de 6,52 lo cual se comparó favorablemente con la puntuación superior de genes de 22,28.

(ii) Rv3616c se clasificó como que es muy esencial para la supervivencia en el modelo de infección de bazo de ratón (calificando 4,945, de una calificación posible de 5).

(iii) La localización subcelular pronosticó que la proteína Rv3616c es una proteína de enlace de membrana y así tiene exposición extracelular significativa, indicando conveniencia como un objetivo de vacuna.

(iv) Rv3616c puede provocar una respuesta protectora contra la amenaza inicial de tuberculosis.

(v) Rv3616c es ampliamente reconocida como un antígeno.

Ejemplo 2 – Predicción de epítopo de Rv3616c

Método

La predicción del epítopo de la célula T se basó en los siguientes enfoques:

Predicción	Nombre	URL/Referencias
CD4 y CD8	Multipred	website: antigen.i2r.a-star.edu.sg/multipred/ Zhang,G.L., Khan,A.M., Srinivasan,K.N., August,J.T. y Brusic,V. (2005) “MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides” Nucleic Acids Res. 33, W172 - W179.
	SVMHC	website: www-bs.informatik.uni-tuebingen.de/SVMHC “Prediction of MHC class I binding peptides, using SVMHC.” Pierre Dönnes y Arne Elofsson in: BMC Bioinformatics 2002 3: 25

(continuación)

Predicción	Nombre	URL/Referencias
CD4	ProPred	website: www.imtech.res.in/raghava/propred/ Singh,H. y Raghava,G.P.S.(2001) "ProPred: Prediction of HLA-DR binding sites." Bioinformatics,17(12), 1236-37.
	Tepitope2	Programa de la casa basado en: H. Bian, J. Hammer (2004) "Discovery of promiscuous HLA-II-restricted T cell epitopes with TEPITOPE." Methods 34 : 468-75
CD8	nHLA	website: www.imtech.res.in/raghava/nhlapred/ Bhasin M. y Raghava G P S (2006) "A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes"; J. Biosci. 32: 31-42
	NetCTL	website: www.cbs.dtu.dk/services/NetCTL/ "An integrative approach to CTL epitope prediction. A combined algorithm integrating MHC-I binding, TAP transport efficiency, y proteasomal cleavage predictions." Larsen M.V., Lundegaard C., Kasper Lamberth, Buus S,. Brunak S., Lund O., y Nielsen M. European Journal of Immunology. 35(8) : 2295-303. 2005
	Epijen	website: www.jenner.ac.uk/EpiJen/ Doytchinova, I. A., P. Guan, D. R. Flower. "EpiJen: a server for multi-step T cell epitope prediction." BMC Bioinformatics, 2006, 7, 131.
	Syfpeithi	website: www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm Hans-Georg Rammensee, Jutta Bachmann, Niels Nikolaus Emmerich, Oskar Alexander Bachor, Stefan Stevanovic: "SYFPEITHI: database for MHC ligands and peptide motifs." Immunogenetics (1999) 50: 213-219

(continuación)

Predicción	Nombre	URL/Referencias
	PredTAP	website: antigen.i2r.a-star.edu.sg/predTAP/ Zhang,G.L., Petrovsky,N., Kwoh,C.K., August,J.T. y Brusic,V. (2006) "PREDTAP: a system for prediction of peptide binding to the human transporter associated with antigen processing." Immunome Res. 2(1), 3.
	PAPROC	http: //www.paproc2.de/paproc1/paproc1.html C. Kuttler, A.K. Nussbaum, T.P. Dick, H.-G. Rammensee, H. Schild, K.P. Hadeler, "An algorithm for the prediction of proteasomal cleavages", J. Mol. Biol. 298 (2000), 417-429 A.K. Nussbaum, C. Kuttler, K.P. Hadeler, H.-G. Rammensee, H. Schild, "PAProC: A Prediction Algorithm for Proteasomal Cleavages available on the WWW", Immunogenetics 53 (2001), 87-94

Resultados

Tabla 2 – Epítopos de células T CD4+ humanas de Rv3616c supuesta

Número de epítopo del CD4 Supuesto	Posición dl amino-ácido	Secuencia del epítopo	SEQ ID No:	Alelo HLA
1	5	FIIDPTISA	SEQ ID No: 29	DRB1_0301, DRB1_0401, DRB1_1101
2	31	ILYSSLEYF	SEQ ID No: 30	DRB1_0301
3	36	LEYFEKALE	SEQ ID No: 31	DRB1_1301
4	63	YAGKNRNHV	SEQ ID No: 32	DRB1_0801
5	87	LIHDQANAV	SEQ ID No: 33	DRB1_0301, DRB1_0401
6	111	FVRPVAVDL	SEQ ID No: 34	DRB1_0101

(continuación)

Número de epítopo del CD4 Supuesto	Posición dl amino-ácido	Secuencia del epítopo	SEQ ID No:	Alelo HLA
7	119	LTYIPVVGH	SEQ ID No: 35	DRB1_0401
8	121	YIPVVGHAL	SEQ ID No: 36	DRB1_0101
9	151	YLVVKTLIN	SEQ ID No: 37	DRB1_0401
10	152	LVVKTLINA	SEQ ID No: 38	DRB1_1301
11	154	VKTLINATQ	SEQ ID No: 39	DRB1_0401
12	164	LKLLAKLAE	SEQ ID No: 40	DRB1_0301, DRB1_0801, DRB1_1101, DRB1_1301
13	173	LVAAAIADI	SEQ ID No: 41	DRB1_0301, DRB1_1101, DRB1_1301
14	181	IISDVADII	SEQ ID No: 42	DRB1_0301
15	197	WEFITNALN	SEQ ID No: 43	DRB1_0401
16	252	LFGAAGLSA	SEQ ID No: 44	DRB1_1501
17	264	LAHADSLAS	SEQ ID No: 45	DRB1_0401
18	270	LASSASLPA	SEQ ID No: 46	DRB1_0401
19	288	FGGLPSLAQ	SEQ ID No: 47	DRB1_0401

Tabla 3 – Epítopos de células T CD8+ humanas de Rv3616c supuesta

Número del epítopo de CD8 supuesto	Posición del amino ácido	Secuencia del epítopo	SEQ ID No:	Alelo HLA
1	5	FIIDPTISA	SEQ ID No: 48	A2
2	6	IIDPTISAI	SEQ ID No: 49	A_0101, A2
3	9	PTISAIDGL	SEQ ID No: 50	A2, A_0201, B7, B8
4	10	TISAIDGLY	SEQ ID No: 51	A1, A_0101, A3, A_0301
5	12	SAIDGLYDL	SEQ ID No: 52	A2, B_3501
6	13	AIDGLYDLL	SEQ ID No: 53	A_0101, A_0201, B44
7	17	LYDLLGIGI	SEQ ID No: 54	A24
8	25	IPNQGGILY	SEQ ID No: 55	B7, A_0101, B_3501, B51
9	30	GILYSSLEY	SEQ ID No: 56	A1, A_0101, A3, A_0301
10	33	YSSLEYFEK	SEQ ID No: 57	A1, A_0301
11	35	SLEYFEKAL	SEQ ID No: 58	A_0201, B7, Cw_0401, Cw_0602
12	38	YFEKALEEL	SEQ ID No: 59	A24, A_2402, B8, Cw_0401, Cw_0602
13	39	FEKALEELA	SEQ ID No: 60	B44, B_4403
14	69	NHVNFFQEL	SEQ ID No: 61	A24, Cw_0602
15	76	ELADLDRQL	SEQ ID No: 62	A_0201
16	77	LADLDRQLI	SEQ ID No: 63	A_0101, B51
17	79	DLDRQLISL	SEQ ID No: 64	A_0101, A_0201
18	80	LDRQLISLI	SEQ ID No: 65	A24, B7, B51
19	94	AVQTTRDIL	SEQ ID No: 66	B7

(continuación)

Número del epítopo de CD8 supuesto	Posición del amino ácido	Secuencia del epítopo	SEQ ID No:	Alelo HLA
20	103	EGAKKGLEF	SEQ ID No: 67	A24, B7
21	107	KGLEFVRPV	SEQ ID No: 68	A_0201, B51
22	108	GLEFVRPVA	SEQ ID No: 69	A_0101, A_0301
23	109	LEFVRPVAV	SEQ ID No: 70	B44
24	111	FVRPVAVDL	SEQ ID No: 71	B7, B8, B_3501
25	113	RPVAVDLTY	SEQ ID No: 72	B7, A_0101, B_3501, B51
26	116	AVDLTYIPV	SEQ ID No: 73	A2, A_0201
27	120	TYIPVVGHA	SEQ ID No: 74	A24
28	121	YIPVVGHAL	SEQ ID No: 75	A_0101, A2, A_0201, B7, B8
29	129	LSAAFQAPF	SEQ ID No: 76	A1, B7, B_3501
30	130	SAAFQAPFC	SEQ ID No: 77	A_0201
31	131	AAFQAPFCA	SEQ ID No: 78	A_0301, B_3501
32	133	FQAPFCAGA	SEQ ID No: 79	A2, A_0201
33	135	APFCAGAMA	SEQ ID No: 80	B7, B_3501
34	136	PFCAGAMAV	SEQ ID No: 81	A3
35	141	AMAVVGGAL	SEQ ID No: 82	A2, A_0201, A24, B7
36	143	AVVGGALAY	SEQ ID No: 83	A1, A3, A_0301, B7
37	147	GALAYLVVK	SEQ ID No: 84	A3, A_0301
38	149	LAYLVVKTL	SEQ ID No: 85	B8, B44, B51

(continuación)

Número del epítopo de CD8 supuesto	Posición del amino ácido	Secuencia del epítopo	SEQ ID No:	Alelo HLA
39	150	AYLVVKTLI	SEQ ID No: 86	A24
40	155	KTLINATQL	SEQ ID No: 87	A_0201, A2, A_0301, A24
41	156	TLINATQLL	SEQ ID No: 88	A2, A_0201, A3, A_0101, Cw_0401
42	158	INATQLLKL	SEQ ID No: 89	B7, B8, Cw_0602
43	159	NATQLLKLL	SEQ ID No: 90	A_2402, B7, B_3501, B44, Cw_0401, Cw_0602
44	162	QLLKLLAKL	SEQ ID No: 91	A2, A_0201, A_0301, A_2402, B8, Cw_0401, Cw_0602
45	165	KLLAKLAEL	SEQ ID No: 92	A2, A_0201, A_0301, B7, B8, Cw_0602
46	166	LLAKLAELV	SEQ ID No: 93	A2, A_0201, A_0101, B8
47	169	KLAELVAAA	SEQ ID No: 94	A2
48	170	LAELVAAAI	SEQ ID No: 95	A1, A24, B51
49	173	LVAAAIADI	SEQ ID No: 96	B7, B51
50	177	AIADIISDV	SEQ ID No: 97	A2, A_0201, Cw_0602
51	178	IADIISDVA	SEQ ID No: 98	A_0101, B_3501
52	182	ISDVADIIK	SEQ ID No: 99	A1, A_0301
53	192	TLGEVWEFI	SEQ ID No: 100	A2, A_0201
54	199	FITNALNGL	SEQ ID No: 101	A2
55	202	NALNGLKEL	SEQ ID No: 102	B51, A_2402, B_3501, Cw_0602
56	213	KLTGWVTGL	SEQ ID No: 103	A2, A_0201

(continuación)

Número del epítopo de CD8 supuesto	Posición del amino ácido	Secuencia del epítopo	SEQ ID No:	Alelo HLA
57	214	LTGWVTGLF	SEQ ID No: 104	A1, A_0101, A24
58	225	GWSNLESFF	SEQ ID No: 105	A24
59	228	NLESFFAGV	SEQ ID No: 106	A2, A_0201
60	231	SFFAGVPGL	SEQ ID No: 107	A2, A_0201, A24, Cw_0401
61	238	GLTGATSGL	SEQ ID No: 108	A2, A_0201
62	246	LSQVTGLFG	SEQ ID No: 109	A1, B8
63	247	SQVTGLFGA	SEQ ID No: 110	A2
64	258	LSASSGLAH	SEQ ID No: 111	A1, A3, B7, B8
65	260	ASSGLAHAD	SEQ ID No: 112	A1, A3, A_0301
66	262	SGLAHADSL	SEQ ID No: 113	A_0201
67	263	GLAHADSLA	SEQ ID No: 114	A_0101, A_0201, A_0301
68	269	SLASSASLP	SEQ ID No: 115	A_0201, A_0301
69	271	ASSASLPAL	SEQ ID No: 116	B7
70	286	SGFGGLPSL	SEQ ID No: 117	A2, A_0201, B51
71	291	LPSLAQVHA	SEQ ID No: 118	B7, B_3501, B51
72	298	HAASTRQAL	SEQ ID No: 119	B7, B8, B_3501
73	301	STRQALRPR	SEQ ID No: 120	A3, A_0301
74	307	RPRADGPVG	SEQ ID No: 121	B7, B_0702, B8, B51
75	319	EQVGGQSQL	SEQ ID No: 122	B7, B44

(continuación)

Número del epítopo de CD8 supuesto	Posición del amino ácido	Secuencia del epítopo	SEQ ID No:	Alelo HLA
76	350	GASKGTTTK	SEQ ID No: 123	A3, A_0301
77	351	ASKGTTTKK	SEQ ID No: 124	A3, A_0301
78	353	KGTTTKKYS	SEQ ID No: 125	A_0301, B8
79	368	TEDAERAPV	SEQ ID No: 126	B44

Como puede verse a partir de las Tablas 2 y 3, la Rv3616c contiene un número pronosticado de epítopos de células T CD4+ y CD8. Además, esta información sugiere que la proteína lleva epítopos que pueden ser reconocidos por HLAs, lo cual ocurre en todo el mundo (esto es HLAs de individuos caucásicos, africanos, asiáticos o latinoamericanos, véase el sitio web en www.allelefrequencies.net).

Ejemplo 3 – Identificación del epítopo de RV3616C

Se preparó una gama de 30 péptidos traslapados que cubrían la longitud completa de la Rv3616c (Véase la Figura 1 para conocer los detalles, y las SEQ ID NOs: 127-156), y se probaron para determinar su capacidad para estimular las células mononucleares de sangre periférica (PBMC) a partir de cuatro donantes de PPD+.

Los datos, mostrados en la Figura 2, revelan que los péptidos 1-7 y 17-30 fueron inmunogénicos para estos individuos. Estos péptidos están adecuadamente presentes dentro de la secuencia de las proteínas Rv3616c modificadas de la invención.

Se debe observar que los péptidos 8-16 (residuos de aminoácidos 92-215) pueden ser inmunogénicos en otros individuos de diferente tipo de HLA.

Ejemplo 4 – homólogos de Rv3616c H37Rv

Las secuencias de Rv3616c a partir de un número de cepas de *M. tuberculosis* y BCG se identificaron utilizando las búsquedas BLASTP de GenBank (número de acceso de secuencia de referencia H37Rv: NP_218133.1) :

Cepa	Número de acceso	% de identidad
CDC1551	NP_338263.1	99
F11	YP_001289574.1	99
Haarlem	ZP_02248979.1	99
C	ZP_00877472.1	99
BCG	YP_979759.1	99

La alineación de las secuencias homólogas indica un alto nivel de identidad.

ENSAYOS BIOLÓGICOS

Cuantificación de las respuestas de células T a Rv3616c

Los polipéptidos se pueden rastrear para determinar su capacidad para activar las células T (inducción de proliferación y/o producción de citocinas) en las células mononucleares de sangre periférica (PBMC) o en preparaciones de sangre entera a partir de los individuos infectados (tal como latentemente infectados).

Los individuos latentemente infectados usualmente se identifican mediante una prueba de piel que tiene un diámetro mayor de 10 milímetros, y sin síntomas, sin cultivo positivo para Mtb, con un esputo negativo, y sin lesiones (como se detectan mediante rayos-X del pecho).

Se pueden utilizar un número de ensayos *in vitro* basados en muestras de células mononucleares de sangre periférica (PBMC) o en sangre entera: después de la reestimulación en la presencia del antígeno (o del fragmento variante/inmunogénico del mismo, como sea apropiado), se puede determinar la proliferación de las células (como se mide mediante CFSE/citometría de flujo), o se puede cuantificar la producción de citocinas (presentes en las células del sobrenadante de cultivo y medidas mediante ELISA, o; después del teñido intracelular de las células T CD4 y CD8 y el análisis mediante citometría de flujo).

Por ejemplo, las muestras de células mononucleares de sangre periférica (PBMC) se pueden obtener a partir de sangre entera heparinizada mediante centrifugación en gradiente de densidad Ficoll-Hypaque siguiendo los procedimientos convencionales. Las células entonces se pueden lavar y se pueden crioconservar en nitrógeno líquido hasta la prueba (para otros detalles, véase Lalvani A y col., J. Infect. Dis., 1999, 180: 1656-1664).

Proliferación de células T

La respuesta inmunitaria específica se puede caracterizar llevando a cabo un análisis de proliferación de linfocitos utilizando la timidina tritiada. Esta técnica evalúa la expansión celular después de la estimulación *in vitro* a un antígeno. En la práctica, la proliferación celular se determina mediante la estimación de la incorporación de timidina tritiada en el ADN, un proceso estrechamente relacionado con los cambios subyacentes en el número de células.

De una manera más adecuada, la proliferación de linfocitos se puede llevar a cabo utilizando el succinimidil-éster de diacetato de carboxi-fluoresceína (CFSE). El succinimidil-éster de diacetato de carboxi-fluoresceína (CFSE) se acopla de una manera espontánea e irreversible con las proteínas tanto intracelulares como de superficie celular, mediante la reacción con las cadenas laterales de lisina y otros grupos amina disponibles. Cuando se dividen las células de linfocitos, el marcado con succinimidil-éster de diacetato de carboxi-fluoresceína (CFSE) se distribuye igualmente entre las células hijas, las cuales, por consiguiente, son la mitad de fluorescentes que las progenitoras. Como un resultado, el llevar a la mitad la intensidad de fluorescencia celular marca a cada generación sucesiva de una población de células proliferantes, y es fácilmente seguida por la citometría de flujo (para otros detalles, véase Hodgkins, PD y col., J. Exp. Med., 1996, 184: 277-281).

Prácticamente, después de la descongelación, las células mononucleares de sangre periférica (PBMC) se pueden lavar y teñir con el succinimidil-éster de diacetato de carboxi-fluoresceína (CFSE) antes de cultivarse (2 x 105 células) durante 72 horas con 10 microgramos/mililitro de antígeno en el medio de cultivo (RPMI-1640 complementado con glutamina, aminoácido no esencial, piruvato, y suero AB humano inactivado por calor). Las células entonces se pueden cosechar, y se puede caracterizar su fenotipo utilizando teñido superficial para identificar las células T CD8 y CD4+ de memoria. Subsiguientemente, se puede utilizar el análisis de citometría de flujo para indicar el grado de proliferación de linfocitos en respuesta a cada antígeno (proporción de células con una disminución en la intensidad del succinimidil-éster de diacetato de carboxi-fluoresceína (CFSE) después de la estimulación *in vitro*).

Producción de citocina

La producción de IFN-γ (o la producción de otras citocinas, tales como, por ejemplo, IL2, TNF-alfa, IL5, IL12, etc.) se puede medir utilizando un análisis de inmunoabsorción ligado a enzimas (ELISA). Las placas ELISA se pueden recubrir con un anticuerpo monoclonal de ratón dirigido hacia el IFN-γ humano (PharMingen, San Diego, CA) en suero regulado con fosfato (PBS) durante cuatro horas a temperatura ambiente. Entonces los pozos se bloquean con suero regulado con fosfato (PBS) que contiene leche deshidratada y descremada al 5 % (peso/volumen) durante 1 hora a temperatura ambiente. Las placas se lavan entonces, por ejemplo, seis veces en suero regulado con fosfato (PBS) /TWEEN-20 al 0,2 por ciento, y las muestras se diluyen a 1: 2 en el medio de cultivo, y las placas ELISA se incuban durante la noche a temperatura ambiente. Las placas se lavan nuevamente, y se puede agregar a cada pozo, un suero con IFN-γ policlonal de conejo anti-humano, por ejemplo, diluido a 1: 3,000 en suero regulado con fosfato (PBS) /suero de cabra normal al 10 por ciento. Las placas se incuban entonces durante dos horas a temperatura ambiente, se lavan, y se puede agregar IgG anti-conejo acoplada con peroxidasa de radícula roja (Sigma Chemical So., St. Louis, MO), por ejemplo, a una dilución de 1: 2,000 en suero regulado con fosfato (PBS) /leche deshidratada y no descremada al 5 por ciento. Después de una incubación durante dos horas adicionales a temperatura ambiente, las placas se lavan, y se agrega el sustrato de TMB. La reacción se puede interrumpir después de 20 minutos con ácido sulfúrico 1 N. Entonces se puede determinar la densidad óptica a 450 nanómetros utilizando 570 nanómetros como una longitud de onda de referencia. Típicamente, se pueden considerar positivas las fracciones que den como resultado que ambas réplicas den una OD dos veces mayor que la OD promedio a

partir de las células cultivadas en el medio solo.

Ejemplo 5 – Inmunogenicidad de la Rv3616c en ratones CB6F1

La inmunogenicidad del antígeno se evaluó en ratones CB6F1 (cruza de la primera generación de ratones BALB/c y C57BL/6).

Los ratones CB6F1 se inmunizaron intramuscularmente tres veces (en el día 0, en el día 14, y en el día 28) con 0,5 microgramos de antígeno de proteína en combinación con el Sistema Adyuvante AS01E (una formulación adyuvante liposomal que comprende 3D-MPL y QS21).

El diseño experimental fue como sigue:

Grupo	**Día 0**	**Día 14**	**Día 28**
1	0,5 µg Rv3616c/AS01E	0,5 µg Rv3616c/AS01E	0,5 µg Rv3616c/AS01E

Se utilizó un total de 24 ratones en el grupo del protocolo.

Se recolectaron linfocitos de sangre periférica (PBL) y se reservaron en el día 21 (es decir, 7 días después de la segunda inmunización), y en el día 35 (es decir, 7 días después de la tercera inmunización), y se midieron las respuestas de células T CD4 y CD8 específicas del antígeno (como se determinaron mediante las células T CD4 o CD8 que produjeron IL-2 y/o IFN-gamma y/o TNF-alfa) mediante citometría de flujo después de la reestimulación *in vitro* durante la noche con reservas de los péptidos 15mer que cubrían las secuencias de interés. La detección de las células T de ratón que expresan IL-2 y/o IFN-gamma y/o TNF-alfa se hizo utilizando amplificación de la expresión de citocina *in vitro* impulsada por el antígeno a corto plazo.

Dicho de una manera breve, se agregó la solución PharmLyse (BD-Pharmingen) a la sangre periférica de ratón heparinizada, con el objeto de lisar los glóbulos rojos sanguíneos. Los PBLs (linfocitos de sangre periférica) obtenidos se lavaron, y entonces se incubaron en la presencia de una reserva de péptidos 15-mer - traslapándose por 11 aminoácidos - que cubrían la secuencia del antígeno de interés, y de 1 microgramo/mililitro de anticuerpos para CD28 y CD49d (BD-Pharmingen). Cada péptido 15-mer se utilizó en una concentración final de 1 microgramo/mililitro. Los controles con el medio también se estimularon con anticuerpos para CD28 y CD49d.

El compuesto bloqueador de la secreción de citocina de brefeldina-A (BD-Pharmingen) se agregó 2 horas después del establecimiento de los cultivos a 37°C, con CO2 al 5 por ciento, y las células se mantuvieron a 37°C, con CO2 al 5 % durante 4 horas adicionales, seguido por la incubación durante la noche a +4°C.

Las células entonces se cosecharon y se tiñeron con anticuerpos anti-CD4 acoplados con Pacific Blue (BD – clon RM4-5, BD-Pharmingen), y anticuerpos anti-CD8 alfa acoplados con cianina5.5 (Cy5.5) de proteína de peridinin-clorofila A (PerCp) (clon 53-6.7, BD-Pharmingen).

Las células luego se lavaron, se fijaron, se permeabilizaron (kit Cytofix-cytoperm, BD-Pharmingen), y se tiñeron con anticuerpos anti-IFN-g acoplados con aloficocianina (clon XMG1.2, BDPharmingen), anticuerpos anti-IL-2 acoplados con isotiocianato de fluoresceína (FITC) (clon JES 6-5H4, Beckman Coulter), y anticuerpos anti-TNF alfa acoplados con ficoeritrina (PE) (clon MP6-XT22, BD-Pharmingen). Después de los lavados finales, las células teñidas se analizaron en un citómetro de flujo LSR II (Beckton-Dickinson). Se adquirió un número mínimo de 10,000 células en el subconjunto CD8+. Para mayores antecedentes, véase Walzer T y col., Cell Immunol., 2000, 206(1) : 16-25 y Maecker HT y col., J. Immunol. Methods, 2001, 255(1-2) : 27-40.

Como controles negativos, algunas células también se cultivaron durante la noche *in vitro* en el medio de cultivo (no estimulado). Las respuestas específicas del antígeno se calcularon sustrayendo la respuesta de citocina promedio producida por las células no estimuladas a partir de la respuesta de citocina promedio producida por las células estimuladas por el péptido.

En cada punto del tiempo y para cada grupo, los datos se recolectaron a partir de 4 reservas de 6 ratones cada una. Los datos que se encuentran a continuación se presentan como el porcentaje de células T CD4 o CD8 productoras de IL-2 y/o IFN-gamma y/o TNF-alfa. Se grafica cada reserva individual de ratones (triángulos), así como el valor promedio del grupo (barra).

La Figura 3 muestra que, en el día 21 (es decir, 7 días después de la segunda inmunización), se detectan las respuestas de células T CD4 y CD8 específicas de Rv3616c en los ratones inmunizados con 0,5 microgramos de la Rv3616c/AS01E.

La Figura 4 muestra el perfil de citocinas de la respuesta de células T CD4 a partir de los linfocitos de sangre

periférica (PBL) estimulados por la reserva del péptido de Rv3616c (sin remover el medio) en el día 21 (es decir, 7 días después de la segunda inmunización).

La Figura 5 muestra el perfil de citocinas de la respuesta de células T CD8 a partir de los linfocitos de sangre periférica (PBL) estimulados por la reserva del péptido de Rv3616c (sin remover el medio) en el día 21 (es decir, 7 días después de la segunda inmunización).

La Figura 6 muestra que en el día 35 (es decir, 7 días después de la tercera inmunización), se detectan las respuestas de células T CD4 y CD8 específicas de Rv3616c en los ratones inmunizados con 0,5 microgramos de la Rv3616c/AS01E. La tercera dosis aumenta la respuesta de células T CD4 pero no la respuesta de células T CD8. Debido a las dificultades técnicas, solamente hubo datos disponibles para una sola reserva.

La Figura 7 muestra el perfil de citocinas de la respuesta de células T CD4 a partir de los linfocitos de sangre periférica (PBL) estimulados por la reserva del péptido de Rv3616c (sin remover el medio) en el día 35 (es decir, 7 días después de la tercera inmunización). Debido a las dificultades técnicas, solamente hubo datos disponibles para una sola reserva.

La Figura 8 muestra el perfil de citocinas de la respuesta de células T CD8 a partir de los linfocitos de sangre periférica (PBL) estimulados por la reserva del péptido de Rv3616c (sin remover el medio) en el día 35 (es decir, 7 días después de la tercera inmunización). Debido a las dificultades técnicas, solamente hubo datos disponibles para una sola reserva.

Ejemplo 6 – Inmunogenicidad de la Rv3616c EN RATONES C57BL/6

La inmunogenicidad del antígeno también se evaluó en ratones C57BL/6.

Los ratones C57BL/6 se inmunizaron intramuscularmente tres veces (en el día 0, en el día 14, y en el día 28) con 1 microgramo de antígeno de proteína en combinación con un Sistema Adyuvante AS01E (una formulación adyuvante liposomal que comprende 3D-MPL y QS21).

El diseño experimental fue el siguiente:

Grupo	Día 0	Día 14	Día 28
1	1 µg Rv3616c/AS01E	1 µg Rv3616c/AS01E	1 µg Rv3616c/AS01E

Se recolectaron linfocitos de sangre periférica (PBL) y se reservaron en el día 21 (es decir, 7 días después de la segunda inmunización), y en el día 35 (es decir, 7 días después de la tercera inmunización), y se midieron las respuestas de células T CD4 y CD8 específicas del antígeno (como se determinan mediante las células T CD4 o CD8 productoras de IL-2 y/o IFN-gamma y/o TNF-alfa) mediante citometría de flujo después de la reestimulación *in vitro* durante la noche con reservas de péptidos 15mer que cubrían las secuencias de interés. El procedimiento seguido fue como se describe anteriormente.

Como controles negativos, algunas células también se cultivaron durante la noche *in vitro* en el medio de cultivo (no estimulado). Las respuestas específicas del antígeno se calcularon sustrayendo la respuesta de citocina promedio producida por las células no estimuladas a partir de la respuesta de citocina promedio producida por las células estimuladas por el péptido.

En cada punto del tiempo y para cada grupo, los datos se recolectaron a partir de 4 reservas de 6 ratones cada una. Los datos que se encuentran a continuación, se presentan como el porcentaje de células T CD4 o CD8 productoras de IL-2 y/o IFN-gamma y/o TNF-alfa. Se grafica cada reserva individual de ratones (triángulos), así como el valor promedio del grupo (barra).

La Figura 9 muestra que en el día 21 (es decir, 7 días después de la segunda inmunización), se detectan las respuestas de células T CD4 y CD8 específicas de Rv3616c en los ratones inmunizados con 1 microgramo de la Rv3616c/AS01E, aunque la respuesta de células T CD8 específicas del antígeno es muy baja (por consiguiente, no se muestran los datos del perfil de citocinas).

La Figura 10 muestra el perfil de citocinas de la respuesta de células T CD4 a partir de los linfocitos de sangre periférica (PBL) estimulados por la reserva del péptido de Rv3616c (sin remover el medio) en el día 21 (es decir, 7 días después de la segunda inmunización).

La Figura 11 muestra que en el día 35 (es decir, 7 días después de la tercera inmunización), se detectan las respuestas de células T CD4 y CD8 específicas de Rv3616c en los ratones inmunizados con 1 microgramo de la Rv3616c/AS01E. Una tercera dosis de inmunización aumenta las respuestas de células T CD4, pero solo

ligeramente la respuesta de células T CD8.

La Figura 12 muestra el perfil de citocinas de la respuesta de células T CD4 a partir de los linfocitos de sangre periférica (PBL) estimulados por la reserva del péptido de Rv3616c (sin remover el medio) en el día 35 (es decir, 7 días después de la tercera inmunización).

La Figura 13 muestra el perfil de citocinas de la respuesta de células T CD8 a partir de los linfocitos de sangre periférica (PBL) estimulados por la reserva del péptido de Rv3616c (sin remover el medio) en el día 35 (es decir, 7 días después de la tercera inmunización).

Ejemplo 7 – reconocimiento *in vitro* de la Rv3616c por PBMC a partir de seres humanos con TB latente

Los experimentos se llevaron a cabo con el objeto de evaluar la respuesta de células T periféricas específica para el antígeno de la invención en 4 adultos saludables limpios de tuberculosis (TB) (prueba de piel PPD = 0 milímetros), y 8 adultos saludables latentemente infectados por tuberculosis (TB) (prueba de piel PPD = 15 milímetros o más) de Sudáfrica.

Datos de la prueba cutánea PPD

Número ID del Individuo	Diámetro de Induración (mm)
4	0
5	0
33	0
38	0
36	15
46	15
13	15
7	16
58	25
74	26
8	53
60	55

La respuesta inmunitaria mediada por células (CMI) se evaluó mediante la medición de citocinas sobre las células mononucleares de sangre periférica (PBMCs) aisladas mediante un ensayo de teñido de citocina intracelular (ICS).

El teñido de citocina intracelular (ICS) llevado a cabo fue una adaptación de la metodología anteriormente descrita (véase Von Eschen y col., Hum. Vaccin., 2009, 5(7)). Las células mononucleares de sangre periférica (PBMCs) se estimularon *in vitro* mediante una reserva de péptidos 15-mer - traslapados por 11 aminoácidos - que cubrían la secuencia entera del antígeno de interés. Las células se estimularon con los péptidos durante 2 horas, se cultivaron adicionalmente durante la noche en la presencia de Brefeldina A, se procesaron para el teñido de citocina intracelular (ICS), y se analizaron utilizando citometría de flujo. Se midieron las frecuencias de las células T CD3+CD4+ o CD3+CD8+ específicas del antígeno que expresan IFN-gamma y/o TNF-alfa y/o IL-17. Se sustrajeron las respuestas de las células estimuladas por el medio a partir de las respuestas obtenidas en las células estimuladas por las reservas de péptidos.

ICS: anticuerpos

PO Anti-CD3 (Invitrogen – cat CD0330)
PB Anti-CD4 (BD - cat 558116)
APC-H7 Anti-CD8 (BD – cat 641400)
AF700 Anti-IFNg (BD-Pharmingen – cat 557995)
PE-Cy7 Anti-TNF (BD-Pharmingen – cat 557647)
AF647 Anti-IL17 (BD-Pharmingen – cat 51-7178-71)

Los resultados se presentan como el número de células T CD3+CD4+ específicas del antígeno que expresan TNF-alfa e IFN-gamma, por millón de células T CD3+CD4+, debido a que estas células representan la principal población de las células T CD4 específicas del antígeno (se remueve el nivel de respuesta del fondo debido al medio). No se detectaron células T CD3+CD8+ específicas del antígeno. La Figura 14 muestra que se mide una respuesta de células T CD4 específicas del antígeno en 6 de los 8 individuos latentemente infectados (y no en los individuos números 7 y 74) cuando se compara con la respuesta de células T CD4 no específicas, medida en los individuos limpios.

Ejemplo 8 – producción de secuencias de rv3616c modificadas

(i) Clonación

La secuencia de nucleótidos de Rv3616 de *Mycobacterium tuberculosis* H37Rv se optimizó en sus codones para la expresión en E. coli y se sintetizó el gen. El inserto obtenido en seguida de la subclonación se clonó en pET21b+ (Novagen) utilizando un sitio de restricción NdeI en el extremo N y un sitio de restricción XhoI en el término C. Para generar las construcciones de Rv3616c modificadas, se llevaron a cabo una serie de amplificaciones con reacción en cadena de la polimerasa (PCR) utilizando diferentes cebadores, con el objeto de suprimir los residuos de nucleótidos específicos dentro de la Rv3616c. Los insertos modificados se clonaron entonces en pET26b+ y/o pET19b (Novagen).

Clon que se va a generar	Cebadores utilizados
pET26_Rv3616Δ136-183His	CAN1001/1004 CAN1003/1002
pET26_Rv3616Δ150-160His	CAN1001/1006 CAN1005/1002
pET26_Rv3616Δ136-154His	CAN1001/1008 CAN1007/1002
pET26_Rv3616Δ166-182His	CAN1001/1010 CAN1009/1002
pET19_Rv3616Δ136-183His	CAN1001/1004 CAN1003/1002
pET19_Rv3616Δ150-160His	CAN1001/1006 CAN1005/1002
pET19_Rv3616Δ136-154His	CAN1001/1008 CAN1007/1002
pET19_Rv3616Δ166-182His	CAN1001/1010 CAN1009/1002
pET26_Rv3616Δ135-139His	CAN1001/1065 CAN1064/1002
pET26_Rv3616Δ142-145His	CAN1001/1067 CAN1066/1002
pET26_Rv3616Δ145-152His	CAN1001/1069 CAN1068/1002

(continuación)

Clon que se va a generar	Cebadores utilizados
pET26_Rv3616Δ138-145His	CAN1001/1071 CAN1070/1002
pET26_Rv3616Δ149-154His	CAN1001/1073 CAN1072/1002

Cebador	Secuencia de Cebador	Sitio de Restricción
CAN1001	ggaattccatatgagccgtgcctttattattgatccgac	Nde1
CAN1002	ccg ctc gag cac cac att gcg aac cag aac	Xho1
CAN1003	ctg agc gca gca ttt cag gca ccg atg tgg ccg ata tta tta aag	ninguno
CAN1004	ctttaataatatcggccacatcggtgcctgaaatgctgcgctcag	ninguno
CAN1005	gttgtgggtggtgctctgacccagctgctgaaactg	ninguno
CAN1006	cagtttcagcagctgggtcagagcaccacccacaac	ninguno
CAN1007	ctgagcgcagcatttcaggcgaaaaccctgattaatgcaac	ninguno
CAN1008	gttgcattaatcagggttttcgcctgaaatgctgcgctcag	ninguno
CAN1009	gcaacccagctgctgaaatccgatgtggccgatattattaaag	ninguno
CAN1010	ctttaataatatcggccacatcggatttcagcagctgggttgc	ninguno
CAN1064	ctgagcgcagcatttcagggtgcaatggcagttgtg	ninguno
CAN1065	cacaactgccattgcaccctgaaatgctgcgctcag	ninguno
CAN1066	caatggcagttgtgggtggtgctaaaaccctgattaatgcaac	ninguno
CAN1067	gttgcattaatcagggttttagcaccacccacaactgccattg	ninguno
CAN1068	ccgttttgtgccggtgcaggtggtgctctggcatatc	ninguno
CAN1069	gatatgccagagcaccacctgcaccggcacaaaacgg	ninguno

(continuación)

Cebador	Secuencia de Cebador	Sitio de Restricción
CAN1070	gccggtgcaatggcagttgttgtgaaaaccctgattaatg	ninguno
CAN1071	cattaatcagggttttcacaacaactgccattgcaccggc	ninguno
CAN1072	gcatttcaggcaccgtttggtggtgctctggcatatc	ninguno
CAN1073	gatatgccagagcaccaccaaacggtgcctgaaatgc	ninguno

(ii) Expresión de las proteínas recombinantes

Cepa huésped: E.coli competente en T7 Express (New England Biolabs : derivado BL21 mejorado.

La transformación de Escherichia coli T7 Express con el ADN del plásmido se llevó a cabo mediante los métodos convencionales con células tratadas con CaCl2 (Hanahan D. « Plasmid transformation by Simanis. » en Glover, D. M. (Editor), DNA Cloning. IRL Press, Londres. (1985) : páginas 109-135).

ID de Plásmidos Recombinantes	Cepa Huésped	Agar de Placa
pET21_Rv3616His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml CarbenicilinaB
pET26_Rv3616Δ136-183His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
pET26_Rv3616Δ150-160His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
pET26_Rv3616Δ136-154His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
pET26_Rv3616Δ166-182His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
pET19_Rv3616Δ136-183His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml CarbenicilinaB
pET19_Rv3616Δ150-160His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml CarbenicilinaB
pET19_Rv3616Δ136-154His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml CarbenicilinaB
pET19_Rv3616Δ166-182His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml CarbenicilinaB

(continuación)

ID de Plásmidos Recombinantes	Cepa Huésped	Agar de Placa
pET26_Rv3616Δ135-139His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
pET26_Rv3616Δ142-145His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
pET26_Rv3616Δ145-152His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
pET26_Rv3616Δ138-145His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
pET26_Rv3616Δ149-154His	T7 ExpressA	Placa de agar LB con Phytone y 100 µg/ml KanamicinaC
A: NEB (número de catálogo: C2566H) B: Teknova, CA, EUA (número de catálogo L1092) C: Teknova, CA, EUA (número de catálogo L1096)		

Se utilizó una placa de agar confluente inoculada con T7 Express de E. coli transformado + plásmido, con el fin de inocular 800 mililitros de caldo LB APS + 50 µg/ml de antibiótico, para obtener la O.D.600nm de entre 0,05 y 0,1. Los cultivos se incubaron a 37°C, a 250 revoluciones por minuto hasta una O.D.600nm de alrededor de 0,8.

La expresión de la proteína recombinante se indujo mediante la adición de 1 mM final de isopropil-ß-D-1-tiogalactopiranosida (IPTG; EMD Chemicals Inc) al medio de cultivo en crecimiento. La inducción se mantuvo durante 3 horas a 37 ºC (o durante la noche a 16°C).

(iii) Purificación

El cultivo bacteriano se centrifugó durante 15 minutos, a 4°C, a 8000g. Los aglomerados del cultivo bacteriano se volvieron a suspender en regulador de lisis (regulador Tris 20 mM (pH de 8,0), y una mezcla de cóctel de inhibidores de proteasa (Complete sin EDTA). Las bacterias se lisaron con el sistema de alteración Constant Cell (Constant System). Los componentes solubles (sobrenadante) e insolubles (aglomerado) se separaron mediante centrifugación a 20,000g durante 20 minutos a 4°C.

Los componentes insolubles (aglomerados) se volvieron a solubilizar en regulador HEPES 20 mM que contenía guanidina HCl 6M, NaCl 500 mM, imidazol 10 mM, pH de 8,0. El sobrenadante se cargó entonces sobre una columna IMAC de 5 mililitros (BioRad). Después de los lavados, se llevó a cabo la elución utilizando un regulador HEPES 20 mM (pH de 8,0) que contenía Guanidina-HCl 6M, NaCl 500 mM, e imidazol 250 mM.

Se llevaron a cabo dos etapas de diálisis en membrana 12-14000 MWCO (SpectraPor) : primaria en un regulador de urea 8M que contenía HEPES 20mM, NaCl 150 mM a un pH de 8,0, seguida por una segunda diálisis en suero regulado con fosfato (PBS), urea 4M, pH de 7,4.

(iv) SDS-PAGE

Se recolectaron muestras a partir de cultivos no inducidos e inducidos, para determinar el perfil de expresión, y se analizaron mediante SDS-PAGE.

Dicho de una manera breve, las muestras se trataron con regulador de muestreo NUPAGE 4X LDS (Invitrogen), se redujeron utilizando DTT 0,05M, y se calentaron a 70°C durante 10 minutos. Las muestras entonces se centrifugaron a la máxima velocidad durante 2 minutos, y se cargaron en gel Bis-Tris al 4-12 % NUPAGE Novex (Invitrogen). La migración se llevó a cabo a 200V durante 35 minutos en Regulador de Ejecución 1X NUPAGE MES (Invitrogen), y el gel se tiñó para permitir la visualización de las proteínas separadas, cuyos resultados se muestran en las Figuras 17 y 18.

Cuando se compara con la expresión del tipo silvestre H37Rv, las construcciones Rv3616Δ138-145, Rv3616Δ136-154, Rv3616Δ150-160, Rv3616Δ166-182, Rv3616Δ149-154 y Rv3616Δ135-139 mejoraron notoriamente.

La construcción Rv3616Δ136-183 contuvo un codón de PARO erróneo dentro de la secuencia, y en consecuencia, la expresión de la secuencia no procedió como se pretendía.

Ejemplo 9 –Producción adicional de secuencias Rv3616c modificadas

Empleando una metodología análoga a la descrita en el Ejemplo 8, en donde se utilizó la cepa BL21 (DE3) en lugar de T7 Express, y se utilizó la placa de agar confluente para inocular 25 mililitros de caldo LB APS con antibiótico, se llevaron a cabo tres ejecuciones de expresión (empezando a partir de la misma placa de transformación) con respecto a una gama de construcciones de Rv3616c modificadas.

Los productos de las ejecuciones de expresión se analizaron mediante SDS-PAGE, y se proporciona un gel representativo a partir de una de las ejecuciones de expresión en la Figura 19. Se encontró que Rv3616Δ138-145 ofrece la mejor expresión de proteína, seguida cercanamente por Rv3616Δ149-154 y Rv3616Δ136-154, mostrando la Rv3616Δ135-139 también una buena expresión.

La cuantificación de la banda correspondiente a la proteína objetivo se llevó a cabo utilizando el software ImageQuant TL. Dicho de una manera breve, los geles SDS-PAGE se tiñeron utilizando el teñido InstantBlue (Novexin), y se escanearon con un UVP BioImaging System en un formato de archivos TIFF. Las bandas entonces se analizaron utilizando el software ImageQuantTL 7,0 de GE Healthcare. Se utilizó la proteína no inducida Rv3616 como el control para la expresión negativa, debido a que no se observó ninguna reactividad con el anticuerpo (Ab) anti-marca His.

Construcción	% Banda: Gel 1	% Banda: Gel 2	% Banda: Gel 3	% Banda promedio
Rv3616 no inducida	9	8	7	8
Rv3616	8	8	8	8
Rv3616Δ150-160	8	10	12	10
Rv3616Δ136-154	22	28	29	26
Rv3616Δ166-182	10	10	9	10
Rv3616Δ135-139	15	16	17	16
Rv3616Δ142-145	9	9	8	9
Rv3616Δ145-152	10	9	10	10
Rv3616Δ149-154	26	28	31	28
Rv3616Δ138-145	23	25	21	23

(continuación)

Construcción	% Banda: Gel 1	% Banda: Gel 2	% Banda: Gel 3	% Banda promedio
% Banda: medida del volumen de la banda dividido entre el volumen total de todas las bandas en la pista.				

En el porcentaje de banda, las proteínas Rv3616Δ149-154, Rv3616Δ138-145 y Rv3616Δ136-154 se expresan todas en niveles notoriamente más altos, comparándose con la secuencia natural o con la construcción de Rv3616Δ150-160 conocida. Rv3616Δ135-139 también se expresó en un alto nivel.

Ejemplo 10 – Inmunogenicidad de rv3616δ138-145 en ratones CB6F1

Se evaluó la inmunogenicidad de Rv3616Δ138-145 en ratones CB6F1.

Los ratones CB6F1 se inmunizaron intramuscularmente tres veces (en el día 0, en el día 14, y en el día 28) con 50 microlitros de la vacuna de prueba que contenía un intervalo de dosis (8 microgramos, 2 microgramos, y 0,5 microgramos) de Rv3616Δ138-145 en combinación con el Sistema Adyuvante AS01E (una formulación adyuvante liposomal que comprende 3D-MPL y QS21). Las formulaciones también contenían urea (4M) y arginina (500 mM).

El diseño experimental fue el siguiente:

Grupo	Día 0	Día 14	Día 28
1	8 µg Rv3616Δ138-145 /AS01E	8 µg Rv3616Δ138-145 /AS01E	8 µg Rv3616Δ138-145 /AS01E
2	2 µg Rv3616Δ138-145 /AS01E	2 µg Rv3616Δ138-145 /AS01E	2 µg Rv3616Δ138-145 /AS01E
3	0,5 µg Rv3616Δ138-145 /AS01E	0,5 µg Rv3616Δ138-145 /AS01E	0,5 µg Rv3616Δ138-145 /AS01E

Se utilizaron un total de 20 ratones en cada grupo de inmunización. 10 ratones recibieron solución salina como un grupo de control negativo (no se muestran los datos).

Se recolectaron linfocitos de sangre periférica (PBL) y se reservaron en los días 21 (es decir, 7 días después de la segunda inmunización), y 35 (es decir, 7 días después de la tercera inmunización), y se midieron las respuestas de células T CD4 y CD8 específicas del antígeno (como se determinaron por las células T CD4 o CD8 productoras de IL-2 y/o IFN-gamma y/o TNF-alfa) mediante citometría de flujo después de una reestimulación *in vitro* durante 6 horas con las reservas de péptidos 15mer que cubrían la secuencia completa del antígeno de Rv3616c. La detección de las células T de ratón que expresaron IL-2 y/o IFN-gamma y/o TNF-alfa se hizo utilizando una amplificación de la expresión de citocina *in vitro* impulsada por el antígeno a corto plazo.

Dicho de una manera breve, se agregó la solución PharmLyse (BD-Pharmingen) a la sangre periférica de ratón heparinizada, con el objeto de lisar los glóbulos rojos sanguíneos. Los PBLs (linfocitos de sangre periférica) obtenidos se lavaron, y entonces se incubaron en la presencia de una reserva de péptidos 15-mer - traslapados por 11 aminoácidos - que cubrían la secuencia del antígeno de interés, y de 1 microgramo/mililitro de anticuerpos para CD28 y CD49d (BD-Pharmingen). Cada péptido 15-mer se utilizó en una concentración final de 1 microgramo/mililitro. Los pozos de control con el medio también se estimularon con anticuerpos para CD28 y CD49d.

El compuesto bloqueador de la secreción de citocina Brefeldina-A (BD-Pharmingen) se agregó 2 horas después del establecimiento de los cultivos a 37°C, con CO2 al 5 por ciento, y las células se mantuvieron a 37°C, con CO2 al 5 % durante 4 horas adicionales, seguidas por su almacenamiento durante la noche a +4°C.

Las células entonces se cosecharon y se tiñeron con anticuerpos anti-CD4 acoplados con Pacific Blue (clon RM4-5, BD-Pharmingen), y anticuerpos anti-CD8 alfa acoplados con cianina5.5 (Cy5.5) de proteína de peridinina-clorofila (PerCp) (clon 53-6.7, BD-Pharmingen).

Las células luego se lavaron, se fijaron, se permeabilizaron (kit Cytofix-cytoperm, BD-Pharmingen), y se tiñeron con

anticuerpos anti-IFN-gamma acoplados con aloficocianina (clon XMG1.2, BD-Pharmingen), anticuerpos anti-IL-2 acoplados con isotiocianato de fluoresceína (FITC) (clon JES 6-5H4, BD-Pharmingen), y anticuerpos anti-TNF-alfa acoplados con ficoeritrina (PE) (clon MP6-XT22, BD-Pharmingen). Después de los lavados finales, las células teñidas se analizaron en un citómetro de flujo LSRII (Becton-Dickinson). Se adquirieron un mínimo de 10,000 células en el subconjunto de CD8+.

Como los controles negativos, algunas células también se cultivaron durante 6 horas *in vitro* en el medio de cultivo (no estimulado). Las respuestas específicas del antígeno se calcularon sustrayendo la respuesta de citocina promedio producida por las células no estimuladas a partir de la respuesta de citocina promedio producida por las células estimuladas por el péptido.

En cada punto del tiempo y para cada grupo, los datos se recolectaron a partir de 4 reservas de 5 ratones cada una, y los datos se presentan como el porcentaje de células T CD4 o CD8 productoras de IL-2 y/o IFN-gamma y/o TNF-alfa. Se grafica cada reserva individual de ratones (rombos negros), así como el valor promedio del grupo (barra).

Los resultados se muestran en las Figuras 20 a 25.

La Figura 20 muestra que, en ambos puntos del tiempo (7dPII y 7dPIII), se detectan las respuestas de células T CD4 específicas de Rv3616c en los ratones inmunizados con cualquier dosis de Rv3616 Δ138-145/AS01E. Los niveles de las respuestas de células T específicas de Rv3616c son más altos en el punto del tiempo 7dPIII, cuando se compara con el punto del tiempo 7dPII. Los perfiles de citocina de la respuesta de células T CD4 a partir de los linfocitos de sangre periférica (PBL) estimulados por la reserva del péptido de Rv3616c (medio removido) se muestran en las Figuras 21 (7dPII) y 22 (7dPIII).

La Figura 23 muestra que en ambos puntos del tiempo (7dPII y 7dPIII), se detectan respuestas de células T CD8 específicas de Rv3616c en los ratones inmunizados con cualquier dosis de Rv3616 Δ138-145/AS01E. Los niveles de respuestas de células T específicas de Rv3616c son más altos en el punto del tiempo 7dPII cuando se compara con el punto del tiempo 7dPIII. Los perfiles de citocina de la respuesta de células T CD8 a partir de los linfocitos de sangre periférica (PBL) estimulados por la reserva del péptido de Rv3616c (medio removido) se muestran en las Figuras 24 (7dPII) y 25 (7dPIII).

En conclusión, se puede observar que el antígeno de Rv3616c es capaz de provocar una respuesta inmunitaria en los ratones tanto CB6F1 como C57BL/6. Adicionalmente, el perfil de producción de citocina indica que una gran proporción de las células T específicas del antígeno expresan una pluralidad de citocinas asociadas con Th1 (es decir, se provoca una respuesta de células T polifuncional). Es importante que están presentes ambas células T específicas del antígeno CD4 y CD8 después de la inmunización, y las células CD8 pueden ser particularmente importantes en un escenario de tuberculosis (TB) latente. La relevancia de la Rv3616c para la infección humana se confirma por el alto nivel de reconocimiento en los individuos latentemente infectados de Sudáfrica, y por la ausencia de respuestas en los sujetos limpios. Por consiguiente, se puede esperar que la Rv3616c sea de un valor sustancial en la prevención, el tratamiento, y el diagnóstico de la infección por tuberculosis (en especial, en la infección por tuberculosis latente).

Se han preparado un número de proteínas Rv3616c modificadas que demuestran claramente una expresión igual o mejor que la secuencia de tipo silvestre H37Rv correspondiente, o que la secuencia de Rv3616Δ150-160 de la técnica anterior. La inmunogenicidad de Rv3616 Δ138-145/AS01E se confirmó en los ratones CB6F1.

Las construcciones que demuestran buenas características de expresión, mientras que se mantiene la inmunogenicidad de la secuencia de tipo silvestre, son la clave para la producción de productos de vacuna comercialmente viables. Las nuevas proteínas Rv3616c modificadas pueden ser de un gran valor en la producción comercial de las composiciones de Rv3616c, tales como las vacunas.

Aunque la invención anterior se ha descrito con algún detalle a manera de ilustración y ejemplo a efectos de claridad de la comprensión, será fácilmente evidente para un experto ordinario en este campo, a la luz de las enseñanzas de la presente invención, que se pueden hacer a la misma ciertos cambios y modificaciones.

A través de toda la memoria descriptiva y en las reivindicaciones siguientes, a menos que el contexto lo requiera de otra manera, la palabra ‘comprenden’, y las variaciones tales como ‘comprende’ y ‘comprendiendo’, se entenderán para implicar la inclusión de un entero, paso, grupo de enteros, o grupo de pasos, pero no la exclusión de cualquier otro entero, paso, grupo de enteros, o grupo de pasos.

LISTADO DE SECUENCIAS

<110> GlaxoSmithKline Biologicals s.a
Glaxo Group Limited
Blais, Normand
Gelinas, Anne-Marie
Brown, James R
Murphy, Dennis

Mettens, Pascal

<120> ANTÍGENOS MODIFICADOS

<130> VB64094PCT

<150> US61/298.710
<151> 27-01-2010

<160> 180

<170> PatentIn versión 3.5

<210> 1
<211> 392
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> MISC_FEATURE
<223> Cepa H37Rv

<400> 1

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125
```

```
Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145                 150                 155                 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
                165                 170                 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Thr
            180                 185                 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
        195                 200                 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
    210                 215                 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225                 230                 235                 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
                245                 250                 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
            260                 265                 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
        275                 280                 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
    290                 295                 300

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

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
                325                 330                 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
            340                 345                 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
        355                 360                 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
    370                 375                 380

Lys Val Leu Val Arg Asn Val Val
```

385 390

<210> 2
<211> 1179
<212> ADN
<213> *Mycobacterium tuberculosis*

<220>
<221> misc_feature
<223> Cepa H37Rv

<400> 2

atgagcagag cgttcatcat cgatccaacg atcagtgcca ttgacggctt gtacgacctt 60

ctggggattg gaatacccaa ccaagggggt atcctttact cctcactaga gtacttcgaa 120

aaagccctgg aggagctggc agcagcgttt ccgggtgatg gctggttagg ttcggccgcg 180

gacaaatacg ccggcaaaaa ccgcaaccac gtgaattttt tccaggaact ggcagacctc 240

gatcgtcagc tcatcagcct gatccacgac caggccaacg cggtccagac gacccgcgac 300

atcctggagg gcgccaagaa aggtctcgag ttcgtgcgcc cggtggctgt ggacctgacc 360

tacatcccgg tcgtcgggca cgccctatcg gccgccttcc aggcgccgtt ttgcgcgggc 420

gcgatggccg tagtgggcgg cgcgcttgcc tacttggtcg tgaaaacgct gatcaacgcg 480

actcaactcc tcaaattgct tgccaaattg gcggagttgg tcgcggccgc cattgcggac 540

atcatttcgg atgtggcgga catcatcaag ggcaccctcg gagaagtgtg ggagttcatc 600

acaaacgcgc tcaacggcct gaaagagctt tgggacaagc tcacggggtg ggtgaccgga 660

ctgttctctc gagggtggtc gaacctggag tccttctttg cgggcgtccc cggcttgacc 720

ggcgcgacca gcggcttgtc gcaagtgact ggcttgttcg gtgcggccgg tctgtccgca 780

tcgtcgggct tggctcacgc ggatagcctg gcgagctcag ccagcttgcc cgccctggcc 840

ggcattgggg gcgggtccgg ttttgggggc ttgccgagcc tggctcaggt ccatgccgcc 900

tcaactcggc aggcgctacg gccccgagct gatggcccgg tcggcgccgc tgccgagcag 960

gtcggcgggc agtcgcagct ggtctccgcg cagggttccc aaggtatggg cggacccgta 1020

ggcatgggcg gcatgcaccc ctcttcgggg gcgtcgaaag ggacgacgac gaagaagtac 1080

tcggaaggcg cggcggcggg cactgaagac gccgagcgcg cgccagtcga agctgacgcg 1140

ggcggtgggc aaaaggtgct ggtacgaaac gtcgtctaa 1179

<210> 3
<211> 392
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> MISC_FEATURE
<223> Cepa CDC1551

<400> 3

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145                 150                 155                 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
                165                 170                 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
            180                 185                 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
        195                 200                 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
    210                 215                 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225                 230                 235                 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
                245                 250                 255
```

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
260 265 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
275 280 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
290 295 300

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

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
325 330 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
340 345 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
355 360 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
370 375 380

Lys Val Leu Val Arg Asn Val Val
385 390

<210> 4
<211> 392
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> MISC_FEATURE
<223> Cepa F11

<400> 4

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1 5 10 15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
20 25 30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
35 40 45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
50 55 60

```
Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145                 150                 155                 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
                165                 170                 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
            180                 185                 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
        195                 200                 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
    210                 215                 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225                 230                 235                 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
                245                 250                 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
            260                 265                 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
        275                 280                 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
    290                 295                 300

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

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
```

325 330 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
340 345 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
355 360 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
370 375 380

Lys Val Leu Val Arg Asn Val Val
385 390

<210> 5
<211> 392
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> MISC_FEATURE
<223> Cepa Haarlem A

<400> 5

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140
```

```
Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145                 150                 155                 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
                165                 170                 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
            180                 185                 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
        195                 200                 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
    210                 215                 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225                 230                 235                 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
                245                 250                 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
            260                 265                 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
        275                 280                 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
    290                 295                 300

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

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
                325                 330                 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
            340                 345                 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
        355                 360                 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
    370                 375                 380

Lys Val Leu Val Arg Asn Val Val
385                 390
```

<210> 6
<211> 392
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> MISC_FEATURE
<223> Cepa C

<400> 6

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145                 150                 155                 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
                165                 170                 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
            180                 185                 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
        195                 200                 205
```

```
Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
    210                 215                 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225                 230                 235                 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
                245                 250                 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
            260                 265                 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
        275                 280                 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
    290                 295                 300

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

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
                325                 330                 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
            340                 345                 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
        355                 360                 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
    370                 375                 380

Lys Val Leu Val Arg Asn Val Val
385                 390
```

<210> 7
<211> 392
<212> PRT
<213> *Mycobacterium bovis*

<220>
<221> MISC_FEATURE
<223> Cepa BCG

<400> 7

```
Met Ser Arg Val Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                   10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
```

```
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145                 150                 155                 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
                165                 170                 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
            180                 185                 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
        195                 200                 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
    210                 215                 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225                 230                 235                 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
                245                 250                 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
            260                 265                 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
        275                 280                 285
```

```
Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
    290                 295                 300

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

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
                325                 330                 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
            340                 345                 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
        355                 360                 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
    370                 375                 380

Lys Val Leu Val Arg Asn Val Val
385                 390
```

<210> 8
<211> 110
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> mat_péptido
<222> (29)..(110)

<400> 8

```
Met Arg Leu Ser Leu Thr Ala Leu Ser Ala Gly Val Gly Ala Val Ala
            -25                 -20                 -15

Met Ser Leu Thr Val Gly Ala Gly Val Ala Ser Ala Asp Pro Val Asp
        -10                 -5                  -1  1

Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala Ala Leu
5                   10                  15                  20

Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser Pro Val
                25                  30                  35

Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro Gln Arg
            40                  45                  50

Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly Ala Ala Gln Tyr
        55                  60                  65
```

Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr
70 75 80

<210> 9
<211> 97
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 9

Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser
1 5 10 15

Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His Thr Ile Gly Gln Ala
20 25 30

Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His Gln Gly Glu Ser Ser
35 40 45

Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val Ala Ala Ala Ala Lys
50 55 60

Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala
65 70 75 80

Gly Thr Tyr Val Ala Ala Asp Ala Ala Ala Ala Ser Thr Tyr Thr Gly
85 90 95

Phe

<210> 10
<211> 94
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 10

```
Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met
1               5                  10                  15

Ile Arg Ala Gln Ala Ala Ser Leu Glu Ala Glu His Gln Ala Ile Val
            20                  25                  30

Arg Asp Val Leu Ala Ala Gly Asp Phe Trp Gly Gly Ala Gly Ser Val
        35                  40                  45

Ala Cys Gln Glu Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile
    50                  55                  60

Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn
65                  70                  75                  80

    Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala
                85                  90
```

<210> 11
<211> 132
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 11

Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
1 5 10 15

Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
20 25 30

Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
35 40 45

Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
50 55 60

Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
65 70 75 80

Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
85 90 95

Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp
100 105 110

Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
115 120 125

Gly Pro Pro Ala
130

<210> 12
<211> 195
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 12

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
1 5 10 15

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
20 25 30

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr

```
        35                  40                  45

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
    50                  55                  60

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

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
                85                  90                  95

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
            100                 105                 110

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
        115                 120                 125

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
    130                 135                 140

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

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
                165                 170                 175

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
            180                 185                 190

Ala Ala Ser
        195
```

<210> 13
<211> 391
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 13

```
Met Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
1               5                  10                  15

Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp
            20                  25                  30

Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
        35                  40                  45

Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
    50                  55                  60
```

```
Leu Met Val Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
65                  70                  75                  80

Ala Gly Gln Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala
                85                  90                  95

Ala Tyr Glu Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala
            100                 105                 110

Glu Asn Arg Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly
        115                 120                 125

Gln Asn Thr Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met
    130                 135                 140

Trp Ala Gln Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala
145                 150                 155                 160

Thr Ala Thr Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr
                165                 170                 175

Ser Ala Gly Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser
            180                 185                 190

Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu
        195                 200                 205

Gln Gln Leu Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu
    210                 215                 220

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

Met Val Ser Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val
                245                 250                 255

Ser Met Thr Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala
            260                 265                 270

Ala Ala Ala Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala
        275                 280                 285

Met Ser Ser Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly
    290                 295                 300

Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val
305                 310                 315                 320
```

```
Pro Gln Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg
                325                 330                 335

Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly
            340                 345                 350

Gln Met Leu Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly
        355                 360                 365

Gly Gly Leu Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met
    370                 375                 380

Pro His Ser Pro Ala Ala Gly
385                 390
```

<210> 14
<211> 423
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 14

```
Met Asp Phe Gly Leu Leu Pro Pro Glu Val Asn Ser Ser Arg Met Tyr
1               5                  10                  15

Ser Gly Pro Gly Pro Glu Ser Met Leu Ala Ala Ala Ala Ala Trp Asp
            20                  25                  30

Gly Val Ala Ala Glu Leu Thr Ser Ala Ala Val Ser Tyr Gly Ser Val
        35                  40                  45

Val Ser Thr Leu Ile Val Glu Pro Trp Met Gly Pro Ala Ala Ala Ala
    50                  55                  60

Met Ala Ala Ala Ala Thr Pro Tyr Val Gly Trp Leu Ala Ala Thr Ala
65                  70                  75                  80

Ala Leu Ala Lys Glu Thr Ala Thr Gln Ala Arg Ala Ala Ala Glu Ala
                85                  90                  95

Phe Gly Thr Ala Phe Ala Met Thr Val Pro Pro Ser Leu Val Ala Ala
            100                 105                 110

Asn Arg Ser Arg Leu Met Ser Leu Val Ala Ala Asn Ile Leu Gly Gln
        115                 120                 125

Asn Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu Tyr Ala Glu Met Trp
    130                 135                 140

Ala Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu Gly Ala Ser Ala Ala
145                 150                 155                 160
```

Ala Ser Ala Leu Pro Pro Phe Thr Pro Pro Val Gln Gly Thr Gly Pro
165 170 175

Ala Gly Pro Ala Ala Ala Ala Ala Ala Thr Gln Ala Ala Gly Ala Gly
180 185 190

Ala Val Ala Asp Ala Gln Ala Thr Leu Ala Gln Leu Pro Pro Gly Ile
195 200 205

Leu Ser Asp Ile Leu Ser Ala Leu Ala Ala Asn Ala Asp Pro Leu Thr
210 215 220

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

Ala Gln Pro Ile Val Ile Pro Thr Pro Ile Gly Glu Leu Asp Val Ile
245 250 255

Ala Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser Ile Ala Leu Ala Ile
260 265 270

Thr Asn Thr Ala Arg Pro Trp His Ile Gly Leu Tyr Gly Asn Ala Gly
275 280 285

Gly Leu Gly Pro Thr Gln Gly His Pro Leu Ser Ser Ala Thr Asp Glu
290 295 300

Pro Glu Pro His Trp Gly Pro Phe Gly Gly Ala Ala Pro Val Ser Ala
305 310 315 320

Gly Val Gly His Ala Ala Leu Val Gly Ala Leu Ser Val Pro His Ser
325 330 335

Trp Thr Thr Ala Ala Pro Glu Ile Gln Leu Ala Val Gln Ala Thr Pro
340 345 350

Thr Phe Ser Ser Ser Ala Gly Ala Asp Pro Thr Ala Leu Asn Gly Met
355 360 365

Pro Ala Gly Leu Leu Ser Gly Met Ala Leu Ala Ser Leu Ala Ala Arg
370 375 380

Gly Thr Thr Gly Gly Gly Gly Thr Arg Ser Gly Thr Ser Thr Asp Gly
385 390 395 400

Gln Glu Asp Gly Arg Lys Pro Pro Val Val Val Ile Arg Glu Gln Pro
405 410 415

Pro Pro Gly Asn Pro Pro Arg
420

<210> 15
<211> 95
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> INIT_MET
<222> (1)..(1)

<220>
<221> mat_péptido
<222> (2)..(95)

<400> 15

```
Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile Glu Ala Ala Ala Ser
-1  1               5                   10                  15

Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser Leu Leu Asp Glu Gly
                20                  25                  30

Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp Gly Gly Ser Gly Ser
            35                  40                  45

Glu Ala Tyr Gln Gly Val Gln Gln Lys Trp Asp Ala Thr Ala Thr Glu
        50                  55                  60

Leu Asn Asn Ala Leu Gln Asn Leu Ala Arg Thr Ile Ser Glu Ala Gly
    65                  70                  75

Gln Ala Met Ala Ser Thr Glu Gly Asn Val Thr Gly Met Phe Ala
80                  85                  90
```

<210> 16
<211> 338
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> mat_péptido
<222> (43)..(338)

<400> 16

```
Met Gln Leu Val Asp Arg Val Arg Gly Ala Val Thr Gly Met Ser Arg
        -40                 -35                 -30

Arg Leu Val Val Gly Ala Val Gly Ala Ala Leu Val Ser Gly Leu Val
    -25                 -20                 -15
```

```
Gly Ala Val Gly Gly Thr Ala Thr Ala Gly Ala Phe Ser Arg Pro Gly
-10                 -5                  -1  1                   5

Leu Pro Val Glu Tyr Leu Gln Val Pro Ser Pro Ser Met Gly Arg Asp
            10                  15                  20

Ile Lys Val Gln Phe Gln Ser Gly Gly Ala Asn Ser Pro Ala Leu Tyr
        25                  30                  35

Leu Leu Asp Gly Leu Arg Ala Gln Asp Asp Phe Ser Gly Trp Asp Ile
    40                  45                  50

Asn Thr Pro Ala Phe Glu Trp Tyr Asp Gln Ser Gly Leu Ser Val Val
55                  60                  65                  70

Met Pro Val Gly Gly Gln Ser Ser Phe Tyr Ser Asp Trp Tyr Gln Pro
                75                  80                  85

Ala Cys Gly Lys Ala Gly Cys Gln Thr Tyr Lys Trp Glu Thr Phe Leu
            90                  95                  100

Thr Ser Glu Leu Pro Gly Trp Leu Gln Ala Asn Arg His Val Lys Pro
        105                 110                 115

Thr Gly Ser Ala Val Val Gly Leu Ser Met Ala Ala Ser Ser Ala Leu
    120                 125                 130

Thr Leu Ala Ile Tyr His Pro Gln Gln Phe Val Tyr Ala Gly Ala Met
135                 140                 145                 150

Ser Gly Leu Leu Asp Pro Ser Gln Ala Met Gly Pro Thr Leu Ile Gly
                155                 160                 165

Leu Ala Met Gly Asp Ala Gly Gly Tyr Lys Ala Ser Asp Met Trp Gly
            170                 175                 180

Pro Lys Glu Asp Pro Ala Trp Gln Arg Asn Asp Pro Leu Leu Asn Val
        185                 190                 195

Gly Lys Leu Ile Ala Asn Asn Thr Arg Val Trp Val Tyr Cys Gly Asn
    200                 205                 210

Gly Lys Pro Ser Asp Leu Gly Gly Asn Asn Leu Pro Ala Lys Phe Leu
215                 220                 225                 230

Glu Gly Phe Val Arg Thr Ser Asn Ile Lys Phe Gln Asp Ala Tyr Asn
                235                 240                 245

Ala Gly Gly Gly His Asn Gly Val Phe Asp Phe Pro Asp Ser Gly Thr
```

250 255 260

His Ser Trp Glu Tyr Trp Gly Ala Gln Leu Asn Ala Met Lys Pro Asp
265 270 275

Leu Gln Arg Ala Leu Gly Ala Thr Pro Asn Thr Gly Pro Ala Pro Gln
280 285 290

Gly Ala
295

<210> 17
<211> 325
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> mat_péptido
<222> (41)..(325)

<400> 17

```
Met Thr Asp Val Ser Arg Lys Ile Arg Ala Trp Gly Arg Arg Leu Met
-40                 -35                 -30                 -25

Ile Gly Thr Ala Ala Ala Val Val Leu Pro Gly Leu Val Gly Leu Ala
                -20                 -15                 -10

Gly Gly Ala Ala Thr Ala Gly Ala Phe Ser Arg Pro Gly Leu Pro Val
            -5                  -1  1                   5

Glu Tyr Leu Gln Val Pro Ser Pro Ser Met Gly Arg Asp Ile Lys Val
    10                  15                  20

Gln Phe Gln Ser Gly Gly Asn Asn Ser Pro Ala Val Tyr Leu Leu Asp
25                  30                  35                  40

Gly Leu Arg Ala Gln Asp Asp Tyr Asn Gly Trp Asp Ile Asn Thr Pro
                45                  50                  55

Ala Phe Glu Trp Tyr Tyr Gln Ser Gly Leu Ser Ile Val Met Pro Val
            60                  65                  70

Gly Gly Gln Ser Ser Phe Tyr Ser Asp Trp Tyr Ser Pro Ala Cys Gly
        75                  80                  85

Lys Ala Gly Cys Gln Thr Tyr Lys Trp Glu Thr Phe Leu Thr Ser Glu
    90                  95                  100

Leu Pro Gln Trp Leu Ser Ala Asn Arg Ala Val Lys Pro Thr Gly Ser
105                 110                 115                 120
```

```
Ala Ala Ile Gly Leu Ser Met Ala Gly Ser Ser Ala Met Ile Leu Ala
                125                 130                 135

Ala Tyr His Pro Gln Gln Phe Ile Tyr Ala Gly Ser Leu Ser Ala Leu
            140                 145                 150

Leu Asp Pro Ser Gln Gly Met Gly Pro Ser Leu Ile Gly Leu Ala Met
        155                 160                 165

Gly Asp Ala Gly Gly Tyr Lys Ala Ala Asp Met Trp Gly Pro Ser Ser
    170                 175                 180

Asp Pro Ala Trp Glu Arg Asn Asp Pro Thr Gln Gln Ile Pro Lys Leu
185                 190                 195                 200

Val Ala Asn Asn Thr Arg Leu Trp Val Tyr Cys Gly Asn Gly Thr Pro
                205                 210                 215

Asn Glu Leu Gly Gly Ala Asn Ile Pro Ala Glu Phe Leu Glu Asn Phe
            220                 225                 230

Val Arg Ser Ser Asn Leu Lys Phe Gln Asp Ala Tyr Asn Ala Ala Gly
        235                 240                 245

Gly His Asn Ala Val Phe Asn Phe Pro Pro Asn Gly Thr His Ser Trp
    250                 255                 260

Glu Tyr Trp Gly Ala Gln Leu Asn Ala Met Lys Gly Asp Leu Gln Ser
265                 270                 275                 280

Ser Leu Gly Ala Gly
                285
```

<210> 18
<211> 144
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> INIT_MET
<222> (1)..(1)

<220>
<221> mat_péptido
<222> (2)..(144)

<400> 18

```
Met Ala Thr Thr Leu Pro Val Gln Arg His Pro Arg Ser Leu Phe Pro
-1  1               5                  10                  15
```

Glu Phe Ser Glu Leu Phe Ala Ala Phe Pro Ser Phe Ala Gly Leu Arg
20 25 30

Pro Thr Phe Asp Thr Arg Leu Met Arg Leu Glu Asp Glu Met Lys Glu
35 40 45

Gly Arg Tyr Glu Val Arg Ala Glu Leu Pro Gly Val Asp Pro Asp Lys
50 55 60

Asp Val Asp Ile Met Val Arg Asp Gly Gln Leu Thr Ile Lys Ala Glu
65 70 75

Arg Thr Glu Gln Lys Asp Phe Asp Gly Arg Ser Glu Phe Ala Tyr Gly
80 85 90 95

Ser Phe Val Arg Thr Val Ser Leu Pro Val Gly Ala Asp Glu Asp Asp
100 105 110

Ile Lys Ala Thr Tyr Asp Lys Gly Ile Leu Thr Val Ser Val Ala Val
115 120 125

Ser Glu Gly Lys Pro Thr Glu Lys His Ile Gln Ile Arg Ser Thr Asn
130 135 140

<210> 19
<211> 228
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> mat_péptido
<222> (24)..(228)

<400> 19

Met Arg Ile Lys Ile Phe Met Leu Val Thr Ala Val Val Leu Leu Cys
-20 -15 -10

Cys Ser Gly Val Ala Thr Ala Ala Pro Lys Thr Tyr Cys Glu Glu Leu
-5 -1 1 5

Lys Gly Thr Asp Thr Gly Gln Ala Cys Gln Ile Gln Met Ser Asp Pro
10 15 20 25

Ala Tyr Asn Ile Asn Ile Ser Leu Pro Ser Tyr Tyr Pro Asp Gln Lys
30 35 40

Ser Leu Glu Asn Tyr Ile Ala Gln Thr Arg Asp Lys Phe Leu Ser Ala
45 50 55

```
Ala Thr Ser Ser Thr Pro Arg Glu Ala Pro Tyr Glu Leu Asn Ile Thr
        60                  65                  70

Ser Ala Thr Tyr Gln Ser Ala Ile Pro Pro Arg Gly Thr Gln Ala Val
    75                  80                  85

Val Leu Lys Val Tyr Gln Asn Ala Gly Gly Thr His Pro Thr Thr Thr
90                  95                  100                 105

Tyr Lys Ala Phe Asp Trp Asp Gln Ala Tyr Arg Lys Pro Ile Thr Tyr
                110                 115                 120

Asp Thr Leu Trp Gln Ala Asp Thr Asp Pro Leu Pro Val Val Phe Pro
            125                 130                 135

Ile Val Gln Gly Glu Leu Ser Lys Gln Thr Gly Gln Gln Val Ser Ile
        140                 145                 150

Ala Pro Asn Ala Gly Leu Asp Pro Val Asn Tyr Gln Asn Phe Ala Val
    155                 160                 165

Thr Asn Asp Gly Val Ile Phe Phe Phe Asn Pro Gly Glu Leu Leu Pro
170                 175                 180                 185

Glu Ala Ala Gly Pro Thr Gln Val Leu Val Pro Arg Ser Ala Ile Asp
                190                 195                 200

Ser Met Leu Ala
            205
```

<210> 20
<211> 355
<212> PRT
<213> *Mycobacterium tuberculosis*

<220>
<221> mat_péptido
<222> (33)..(355)

<400> 20

```
Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
        -30                 -25                 -20

Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
    -15                 -10                 -5                  -1

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
1               5                   10                  15

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
```

```
            20                  25                  30

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
        35                  40                  45

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
    50                  55                  60

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

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
                85                  90                  95

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
            100                 105                 110

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
        115                 120                 125

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
    130                 135                 140

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

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
                165                 170                 175

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
            180                 185                 190

Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala
        195                 200                 205

Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly
    210                 215                 220

Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu
225                 230                 235                 240

Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val
                245                 250                 255

Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile
            260                 265                 270

Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp
        275                 280                 285
```

```
Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp Gln
    290                 295                 300

Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
305                 310                 315                 320

Pro Pro Ala
```

<210> 21
<211> 323
<212> PRT
<213> Secuencia artificial

<220>
<223> Mtb32A madura mutada en Ser/Ala.

<400> 21

```
Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
1               5                   10                  15

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
            20                  25                  30

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
        35                  40                  45

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
    50                  55                  60

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

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
                85                  90                  95

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
            100                 105                 110

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
        115                 120                 125

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
    130                 135                 140

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

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala
```

```
                165                 170                 175

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
            180                 185                 190

Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala
        195                 200                 205

Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly
    210                 215                 220

Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu
225                 230                 235                 240

Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val
                245                 250                 255

Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile
            260                 265                 270

Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp
        275                 280                 285

Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp Gln
    290                 295                 300

Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
305                 310                 315                 320

Pro Pro Ala
```

<210> 22
<211> 96
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 22

Met Ser Gln Ile Met Tyr Asn Tyr Pro Ala Met Leu Gly His Ala Gly
1 5 10 15

Asp Met Ala Gly Tyr Ala Gly Thr Leu Gln Ser Leu Gly Ala Glu Ile
20 25 30

Ala Val Glu Gln Ala Ala Leu Gln Ser Ala Trp Gln Gly Asp Thr Gly
35 40 45

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

Leu Val Arg Ala Tyr His Ala Met Ser Ser Thr His Glu Ala Asn Thr
65 70 75 80

Met Ala Met Met Ala Arg Asp Thr Ala Glu Ala Ala Lys Trp Gly Gly
85 90 95

<210> 23
<211> 723
<212> PRT
<213> Secuencia artificial

<220>
<223> Mtb72f

<400> 23

```
Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1               5                  10                  15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
            20                  25                  30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
        35                  40                  45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
    50                  55                  60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65                  70                  75                  80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
                85                  90                  95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
            100                 105                 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
        115                 120                 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
    130                 135                 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145                 150                 155                 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
                165                 170                 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
            180                 185                 190
```

```
Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
        195                 200                 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
    210                 215                 220

Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
225                 230                 235                 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
                245                 250                 255

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
            260                 265                 270

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
        275                 280                 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
    290                 295                 300

Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
305                 310                 315                 320

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
                325                 330                 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly
            340                 345                 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
        355                 360                 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
    370                 375                 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
385                 390                 395                 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
                405                 410                 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
            420                 425                 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
        435                 440                 445
```

```
Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
    450                 455                 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465                 470                 475                 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
                485                 490                 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
            500                 505                 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
        515                 520                 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
    530                 535                 540

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
545                 550                 555                 560

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
                565                 570                 575

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
            580                 585                 590

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

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
    610                 615                 620

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
625                 630                 635                 640

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
                645                 650                 655

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
            660                 665                 670

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
        675                 680                 685

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
    690                 695                 700
```

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
705 710 715 720

Ala Ala Ser

<210> 24
<211> 723
<212> PRT
<213> Secuencia artificial

<220>
<223> M72

<400> 24

```
Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1               5                  10                  15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
            20                  25                  30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
        35                  40                  45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
    50                  55                  60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65                  70                  75                  80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
                85                  90                  95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
            100                 105                 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
        115                 120                 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
    130                 135                 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145                 150                 155                 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
                165                 170                 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
            180                 185                 190
```

```
Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
        195                 200                 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
    210                 215                 220

Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
225                 230                 235                 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
                245                 250                 255

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
            260                 265                 270

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
        275                 280                 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
    290                 295                 300

Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
305                 310                 315                 320

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
                325                 330                 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly
            340                 345                 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
        355                 360                 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
    370                 375                 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
385                 390                 395                 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
                405                 410                 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
            420                 425                 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
        435                 440                 445
```

```
Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
    450                 455                 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465                 470                 475                 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
                485                 490                 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
            500                 505                 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
        515                 520                 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
    530                 535                 540

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
545                 550                 555                 560

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
                565                 570                 575

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
            580                 585                 590

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

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
    610                 615                 620

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
625                 630                 635                 640

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
                645                 650                 655

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
            660                 665                 670

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
        675                 680                 685

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala
    690                 695                 700

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
```

705 710 715 720

Ala Ala Ser

<210> 25
<211> 702
<212> PRT
<213> Secuencia artificial

<220>
<223> Mtb71f

<400> 25

Asp Pro Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val
1 5 10 15

Val Ala Ala Leu Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn
20 25 30

Ala Ser Pro Val Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro
35 40 45

Pro Pro Gln Arg Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly
50 55 60

Ala Ala Gln Tyr Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn
65 70 75 80

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

His Gly Ala Met Ile Arg Ala Gln Ala Ala Ser Leu Glu Ala Glu His
100 105 110

Gln Ala Ile Val Arg Asp Val Leu Ala Ala Gly Asp Phe Trp Gly Gly
115 120 125

Ala Gly Ser Val Ala Cys Gln Glu Phe Ile Thr Gln Leu Gly Arg Asn
130 135 140

Phe Gln Val Ile Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln
145 150 155 160

Ala Ala Gly Asn Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser
165 170 175

Trp Ala Thr Ser Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu Val
180 185 190

```
Ala Ser Gln Ser Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His Thr
        195                 200                 205

Ile Gly Gln Ala Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His Gln
    210                 215                 220

Gly Glu Ser Ser Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val Ala
225                 230                 235                 240

Ala Ala Ala Lys Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn Leu
                245                 250                 255

Gly Glu Ala Ala Gly Thr Tyr Val Ala Ala Asp Ala Ala Ala Ala Ser
            260                 265                 270

Thr Tyr Thr Gly Phe Asp Ile Met Asp Phe Gly Leu Leu Pro Pro Glu
        275                 280                 285

Val Asn Ser Ser Arg Met Tyr Ser Gly Pro Gly Pro Glu Ser Met Leu
    290                 295                 300

Ala Ala Ala Ala Ala Trp Asp Gly Val Ala Ala Glu Leu Thr Ser Ala
305                 310                 315                 320

Ala Val Ser Tyr Gly Ser Val Val Ser Thr Leu Ile Val Glu Pro Trp
                325                 330                 335

Met Gly Pro Ala Ala Ala Ala Met Ala Ala Ala Ala Thr Pro Tyr Val
            340                 345                 350

Gly Trp Leu Ala Ala Thr Ala Ala Leu Ala Lys Glu Thr Ala Thr Gln
        355                 360                 365

Ala Arg Ala Ala Ala Glu Ala Phe Gly Thr Ala Phe Ala Met Thr Val
    370                 375                 380

Pro Pro Ser Leu Val Ala Ala Asn Arg Ser Arg Leu Met Ser Leu Val
385                 390                 395                 400

Ala Ala Asn Ile Leu Gly Gln Asn Ser Ala Ala Ile Ala Ala Thr Gln
                405                 410                 415

Ala Glu Tyr Ala Glu Met Trp Ala Gln Asp Ala Ala Val Met Tyr Ser
            420                 425                 430

Tyr Glu Gly Ala Ser Ala Ala Ala Ser Ala Leu Pro Pro Phe Thr Pro
        435                 440                 445

Pro Val Gln Gly Thr Gly Pro Ala Gly Pro Ala Ala Ala Ala Ala Ala
```

```
    450                 455                 460

Thr Gln Ala Ala Gly Ala Gly Ala Val Ala Asp Ala Gln Ala Thr Leu
465                 470                 475                 480

Ala Gln Leu Pro Pro Gly Ile Leu Ser Asp Ile Leu Ser Ala Leu Ala
                485                 490                 495

Ala Asn Ala Asp Pro Leu Thr Ser Gly Leu Leu Gly Ile Ala Ser Thr
            500                 505                 510

Leu Asn Pro Gln Val Gly Ser Ala Gln Pro Ile Val Ile Pro Thr Pro
        515                 520                 525

Ile Gly Glu Leu Asp Val Ile Ala Leu Tyr Ile Ala Ser Ile Ala Thr
    530                 535                 540

Gly Ser Ile Ala Leu Ala Ile Thr Asn Thr Ala Arg Pro Trp His Ile
545                 550                 555                 560

Gly Leu Tyr Gly Asn Ala Gly Gly Leu Gly Pro Thr Gln Gly His Pro
                565                 570                 575

Leu Ser Ser Ala Thr Asp Glu Pro Glu Pro His Trp Gly Pro Phe Gly
            580                 585                 590

Gly Ala Ala Pro Val Ser Ala Gly Val Gly His Ala Ala Leu Val Gly
        595                 600                 605

Ala Leu Ser Val Pro His Ser Trp Thr Thr Ala Ala Pro Glu Ile Gln
    610                 615                 620

Leu Ala Val Gln Ala Thr Pro Thr Phe Ser Ser Ser Ala Gly Ala Asp
625                 630                 635                 640

Pro Thr Ala Leu Asn Gly Met Pro Ala Gly Leu Leu Ser Gly Met Ala
                645                 650                 655

Leu Ala Ser Leu Ala Ala Arg Gly Thr Thr Gly Gly Gly Gly Thr Arg
            660                 665                 670

Ser Gly Thr Ser Thr Asp Gly Gln Glu Asp Gly Arg Lys Pro Pro Val
        675                 680                 685

Val Val Ile Arg Glu Gln Pro Pro Pro Gly Asn Pro Pro Arg
    690                 695                 700
```

<210> 26
<211> 920
<212> PRT
<213> Secuencia artificial

<220>
<223> M72-Mtb9.9-Mtb9.8

<400> 26

```
Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1               5                  10                  15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
            20                  25                  30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
        35                  40                  45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
    50                  55                  60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65                  70                  75                  80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
                85                  90                  95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
            100                 105                 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
        115                 120                 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
    130                 135                 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145                 150                 155                 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
                165                 170                 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
            180                 185                 190

Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
        195                 200                 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
    210                 215                 220

Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
```

225 230 235 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
245 250 255

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
260 265 270

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
275 280 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
290 295 300

Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
305 310 315 320

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
325 330 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly
340 345 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
355 360 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
370 375 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
385 390 395 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
405 410 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
420 425 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
435 440 445

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
450 455 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465 470 475 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
485 490 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
500 505 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
515 520 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
530 535 540

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
545 550 555 560

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
565 570 575

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
580 585 590

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

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
610 615 620

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
625 630 635 640

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
645 650 655

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
660 665 670

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
675 680 685

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala
690 695 700

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
705 710 715 720

Ala Ala Ser Ser Thr Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp
725 730 735

Ala His Gly Ala Met Ile Arg Ala Gln Ala Ala Ser Leu Glu Ala Glu
740 745 750

```
His Gln Ala Ile Val Arg Asp Val Leu Ala Ala Gly Asp Phe Trp Gly
        755                 760                 765

Gly Ala Gly Ser Val Ala Cys Gln Glu Phe Ile Thr Gln Leu Gly Arg
    770                 775                 780

Asn Phe Gln Val Ile Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val
785                 790                 795                 800

Gln Ala Ala Gly Asn Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser
                805                 810                 815

Ser Trp Ala Thr Ser Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu
            820                 825                 830

Val Ala Ser Gln Ser Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His
        835                 840                 845

Thr Ile Gly Gln Ala Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His
    850                 855                 860

Gln Gly Glu Ser Ser Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val
865                 870                 875                 880

Ala Ala Ala Ala Lys Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn
                885                 890                 895

Leu Gly Glu Ala Ala Gly Thr Tyr Val Ala Ala Asp Ala Ala Ala Ala
            900                 905                 910

Ser Thr Tyr Thr Gly Phe Pro Trp
        915                 920
```

<210> 27
<211> 1010
<212> PRT
<213> Secuencia artificial

<220>
<223> M103

<400> 27

```
Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1               5                   10                  15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
            20                  25                  30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
        35                  40                  45
```

```
Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
    50                  55                  60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65                  70                  75                  80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
                85                  90                  95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
            100                 105                 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
        115                 120                 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
    130                 135                 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145                 150                 155                 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
                165                 170                 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
            180                 185                 190

Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
        195                 200                 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
    210                 215                 220

Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
225                 230                 235                 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
                245                 250                 255

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
            260                 265                 270

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
        275                 280                 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
    290                 295                 300
```

```
Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
305                 310                 315                 320

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
                325                 330                 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly
            340                 345                 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
        355                 360                 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
    370                 375                 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
385                 390                 395                 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
                405                 410                 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
            420                 425                 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
        435                 440                 445

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
    450                 455                 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465                 470                 475                 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
                485                 490                 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
            500                 505                 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
        515                 520                 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
    530                 535                 540

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
545                 550                 555                 560
```

```
Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
                565                 570                 575

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
            580                 585                 590

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

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
    610                 615                 620

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
625                 630                 635                 640

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
                645                 650                 655

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
            660                 665                 670

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
        675                 680                 685

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala
    690                 695                 700

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
705                 710                 715                 720

Ala Ala Ser Ser Gly Phe Ser Arg Pro Gly Leu Pro Val Glu Tyr Leu
                725                 730                 735

Gln Val Pro Ser Pro Ser Met Gly Arg Asp Ile Lys Val Gln Phe Gln
            740                 745                 750

Ser Gly Gly Asn Asn Ser Pro Ala Val Tyr Leu Leu Asp Gly Leu Arg
        755                 760                 765

Ala Gln Asp Asp Tyr Asn Gly Trp Asp Ile Asn Thr Pro Ala Phe Glu
    770                 775                 780

Trp Tyr Tyr Gln Ser Gly Leu Ser Ile Val Met Pro Val Gly Gly Gln
785                 790                 795                 800

Ser Ser Phe Tyr Ser Asp Trp Tyr Ser Pro Ala Cys Gly Lys Ala Gly
                805                 810                 815

Cys Gln Thr Tyr Lys Trp Glu Thr Phe Leu Thr Ser Glu Leu Pro Gln
```

820 825 830

Trp Leu Ser Ala Asn Arg Ala Val Lys Pro Thr Gly Ser Ala Ala Ile
835 840 845

Gly Leu Ser Met Ala Gly Ser Ser Ala Met Ile Leu Ala Ala Tyr His
850 855 860

Pro Gln Gln Phe Ile Tyr Ala Gly Ser Leu Ser Ala Leu Leu Asp Pro
865 870 875 880

Ser Gln Gly Met Gly Pro Ser Leu Ile Gly Leu Ala Met Gly Asp Ala
885 890 895

Gly Gly Tyr Lys Ala Ala Asp Met Trp Gly Pro Ser Ser Asp Pro Ala
900 905 910

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

Asn Thr Arg Leu Trp Val Tyr Cys Gly Asn Gly Thr Pro Asn Glu Leu
930 935 940

Gly Gly Ala Asn Ile Pro Ala Glu Phe Leu Glu Asn Phe Val Arg Ser
945 950 955 960

Ser Asn Leu Lys Phe Gln Asp Ala Tyr Asn Ala Ala Gly Gly His Asn
965 970 975

Ala Val Phe Asn Phe Pro Pro Asn Gly Thr His Ser Trp Glu Tyr Trp
980 985 990

Gly Ala Gln Leu Asn Ala Met Lys Gly Asp Leu Gln Ser Ser Leu Gly
995 1000 1005

Ala Gly
1010

<210> 28
<211> 1148
<212> PRT
<213> Secuencia artificial

<220>
<223> M114

<400> 28

Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1 5 10 15

```
Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
            20                  25                  30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
        35                  40                  45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
    50                  55                  60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65                  70                  75                  80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
                85                  90                  95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
            100                 105                 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
        115                 120                 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
    130                 135                 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145                 150                 155                 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
                165                 170                 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
            180                 185                 190

Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
        195                 200                 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
    210                 215                 220

Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
225                 230                 235                 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
                245                 250                 255

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
            260                 265                 270

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
```

275 280 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
290 295 300

Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
305 310 315 320

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
325 330 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly
340 345 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
355 360 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
370 375 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
385 390 395 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
405 410 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
420 425 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
435 440 445

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
450 455 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465 470 475 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
485 490 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
500 505 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
515 520 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
530 535 540

```
Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
545                 550                 555                 560

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
                565                 570                 575

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
            580                 585                 590

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

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
    610                 615                 620

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
625                 630                 635                 640

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
                645                 650                 655

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
            660                 665                 670

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
        675                 680                 685

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala
    690                 695                 700

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
705                 710                 715                 720

Ala Ala Ser Ser Thr Met Asp Phe Gly Leu Leu Pro Pro Glu Val Asn
                725                 730                 735

Ser Ser Arg Met Tyr Ser Gly Pro Gly Pro Glu Ser Met Leu Ala Ala
            740                 745                 750

Ala Ala Ala Trp Asp Gly Val Ala Ala Glu Leu Thr Ser Ala Ala Val
        755                 760                 765

Ser Tyr Gly Ser Val Val Ser Thr Leu Ile Val Glu Pro Trp Met Gly
    770                 775                 780

Pro Ala Ala Ala Ala Met Ala Ala Ala Ala Thr Pro Tyr Val Gly Trp
785                 790                 795                 800
```

```
Leu Ala Ala Thr Ala Ala Leu Ala Lys Glu Thr Ala Thr Gln Ala Arg
                805                 810                 815

Ala Ala Ala Glu Ala Phe Gly Thr Ala Phe Ala Met Thr Val Pro Pro
            820                 825                 830

Ser Leu Val Ala Ala Asn Arg Ser Arg Leu Met Ser Leu Val Ala Ala
        835                 840                 845

Asn Ile Leu Gly Gln Asn Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu
    850                 855                 860

Tyr Ala Glu Met Trp Ala Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu
865                 870                 875                 880

Gly Ala Ser Ala Ala Ala Ser Ala Leu Pro Pro Phe Thr Pro Pro Val
                885                 890                 895

Gln Gly Thr Gly Pro Ala Gly Pro Ala Ala Ala Ala Ala Ala Thr Gln
            900                 905                 910

Ala Ala Gly Ala Gly Ala Val Ala Asp Ala Gln Ala Thr Leu Ala Gln
        915                 920                 925

Leu Pro Pro Gly Ile Leu Ser Asp Ile Leu Ser Ala Leu Ala Ala Asn
    930                 935                 940

Ala Asp Pro Leu Thr Ser Gly Leu Leu Gly Ile Ala Ser Thr Leu Asn
945                 950                 955                 960

Pro Gln Val Gly Ser Ala Gln Pro Ile Val Ile Pro Thr Pro Ile Gly
                965                 970                 975

Glu Leu Asp Val Ile Ala Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser
            980                 985                 990

Ile Ala Leu Ala Ile Thr Asn Thr  Ala Arg Pro Trp His  Ile Gly Leu
        995                 1000                1005

Tyr Gly  Asn Ala Gly Gly Leu  Gly Pro Thr Gln Gly  His Pro Leu
    1010                1015                1020

Ser Ser  Ala Thr Asp Glu Pro  Glu Pro His Trp Gly  Pro Phe Gly
    1025                1030                1035

Gly Ala  Ala Pro Val Ser Ala  Gly Val Gly His Ala  Ala Leu Val
    1040                1045                1050
```

```
Gly Ala  Leu Ser Val Pro His  Ser Trp Thr Thr Ala  Ala Pro Glu
    1055                 1060                 1065

Ile Gln  Leu Ala Val Gln Ala  Thr Pro Thr Phe Ser  Ser Ser Ala
    1070                 1075                 1080

Gly Ala  Asp Pro Thr Ala Leu  Asn Gly Met Pro Ala  Gly Leu Leu
    1085                 1090                 1095

Ser Gly  Met Ala Leu Ala Ser  Leu Ala Ala Arg Gly  Thr Thr Gly
    1100                 1105                 1110

Gly Gly  Gly Thr Arg Ser Gly  Thr Ser Thr Asp Gly  Gln Glu Asp
    1115                 1120                 1125

Gly Arg  Lys Pro Pro Val Val  Val Ile Arg Glu Gln  Pro Pro Pro
    1130                 1135                 1140

Gly Asn  Pro Pro Arg
    1145
```

<210> 29
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 29

```
Phe Ile Ile Asp Pro Thr Ile Ser Ala
1               5
```

<210> 30
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 30

```
Ile Leu Tyr Ser Ser Leu Glu Tyr Phe
1               5
```

<210> 31
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 31

```
Leu Glu Tyr Phe Glu Lys Ala Leu Glu
1               5
```

<210> 32
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 32

Tyr Ala Gly Lys Asn Arg Asn His Val
1 5

<210> 33
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 33

Leu Ile His Asp Gln Ala Asn Ala Val
1 5

<210> 34
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 34

Phe Val Arg Pro Val Ala Val Asp Leu
1 5

<210> 35
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 35

Leu Thr Tyr Ile Pro Val Val Gly His
1 5

<210> 36
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 36

Tyr Ile Pro Val Val Gly His Ala Leu
1 5

<210> 37
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 37

Tyr Leu Val Val Lys Thr Leu Ile Asn
1 5

<210> 38
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 38

Leu Val Val Lys Thr Leu Ile Asn Ala
1 5

<210> 39
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 39

Val Lys Thr Leu Ile Asn Ala Thr Gln
1 5

<210> 40
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 40

Leu Lys Leu Leu Ala Lys Leu Ala Glu
1 5

<210> 41
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 41

Leu Val Ala Ala Ala Ile Ala Asp Ile
1 5

<210> 42
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 42

Ile Ile Ser Asp Val Ala Asp Ile Ile
1 5

<210> 43
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 43

Trp Glu Phe Ile Thr Asn Ala Leu Asn
1 5

<210> 44
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 44

Leu Phe Gly Ala Ala Gly Leu Ser Ala
1 5

<210> 45
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 45

Leu Ala His Ala Asp Ser Leu Ala Ser
1 5

<210> 46
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 46

Leu Ala Ser Ser Ala Ser Leu Pro Ala
1 5

<210> 47
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 47

Phe Gly Gly Leu Pro Ser Leu Ala Gln
1 5

<210> 48
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 48

Phe Ile Ile Asp Pro Thr Ile Ser Ala
1 5

<210> 49
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 49

Ile Ile Asp Pro Thr Ile Ser Ala Ile
1 5

<210> 50
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 50

Pro Thr Ile Ser Ala Ile Asp Gly Leu
1 5

<210> 51
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 51

Thr Ile Ser Ala Ile Asp Gly Leu Tyr
1 5

<210> 52
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 52

Ser Ala Ile Asp Gly Leu Tyr Asp Leu
1 5

<210> 53
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 53

Ala Ile Asp Gly Leu Tyr Asp Leu Leu
1 5

<210> 54
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 54

Leu Tyr Asp Leu Leu Gly Ile Gly Ile
1 5

<210> 55
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 55

Ile Pro Asn Gln Gly Gly Ile Leu Tyr
1 5

<210> 56
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 56

Gly Ile Leu Tyr Ser Ser Leu Glu Tyr
1 5

<210> 57
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 57

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys
1 5

<210> 58
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 58

Ser Leu Glu Tyr Phe Glu Lys Ala Leu
1 5

<210> 59
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 59

Tyr Phe Glu Lys Ala Leu Glu Glu Leu
1 5

<210> 60
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 60

Phe Glu Lys Ala Leu Glu Glu Leu Ala
1 5

<210> 61
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 61

Asn His Val Asn Phe Phe Gln Glu Leu
1 5

<210> 62
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 62

Glu Leu Ala Asp Leu Asp Arg Gln Leu
1 5

<210> 63

<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 63

Leu Ala Asp Leu Asp Arg Gln Leu Ile
1 5

<210> 64
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 64

Asp Leu Asp Arg Gln Leu Ile Ser Leu
1 5

<210> 65
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 65

Leu Asp Arg Gln Leu Ile Ser Leu Ile
1 5

<210> 66
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 66

Ala Val Gln Thr Thr Arg Asp Ile Leu
1 5

<210> 67
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 67

Glu Gly Ala Lys Lys Gly Leu Glu Phe
1 5

<210> 68
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 68

Lys Gly Leu Glu Phe Val Arg Pro Val
1 5

<210> 69
<211> 9

<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 69

Gly Leu Glu Phe Val Arg Pro Val Ala
1 5

<210> 70
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 70

Leu Glu Phe Val Arg Pro Val Ala Val
1 5

<210> 71
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 71

Phe Val Arg Pro Val Ala Val Asp Leu
1 5

<210> 72
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 72

Arg Pro Val Ala Val Asp Leu Thr Tyr
1 5

<210> 73
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 73

Ala Val Asp Leu Thr Tyr Ile Pro Val

1 5

<210> 74
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 74

Thr Tyr Ile Pro Val Val Gly His Ala
1 5

<210> 75

<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 75

Tyr Ile Pro Val Val Gly His Ala Leu
1 5

<210> 76
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 76

Leu Ser Ala Ala Phe Gln Ala Pro Phe
1 5

<210> 77
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 77

Ser Ala Ala Phe Gln Ala Pro Phe Cys
1 5

<210> 78
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 78

Ala Ala Phe Gln Ala Pro Phe Cys Ala
1 5

<210> 79
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 79

Phe Gln Ala Pro Phe Cys Ala Gly Ala
1 5

<210> 80
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 80

Ala Pro Phe Cys Ala Gly Ala Met Ala
1 5

<210> 81
<211> 9
<212> PRT

<213> *Mycobacterium tuberculosis*

<400> 81

Pro Phe Cys Ala Gly Ala Met Ala Val
1 5

<210> 82
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 82

Ala Met Ala Val Val Gly Gly Ala Leu
1 5

<210> 83
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 83

Ala Val Val Gly Gly Ala Leu Ala Tyr
1 5

<210> 84
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 84

Gly Ala Leu Ala Tyr Leu Val Val Lys
1 5

<210> 85
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 85

Leu Ala Tyr Leu Val Val Lys Thr Leu
1 5

<210> 86
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 86

Ala Tyr Leu Val Val Lys Thr Leu Ile
1 5

<210> 87
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 87

Lys Thr Leu Ile Asn Ala Thr Gln Leu
1 5

<210> 88
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 88

Thr Leu Ile Asn Ala Thr Gln Leu Leu
1 5

<210> 89
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 89

Ile Asn Ala Thr Gln Leu Leu Lys Leu
1 5

<210> 90
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 90

Asn Ala Thr Gln Leu Leu Lys Leu Leu
1 5

<210> 91
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 91

Gln Leu Leu Lys Leu Leu Ala Lys Leu
1 5

<210> 92
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 92

Lys Leu Leu Ala Lys Leu Ala Glu Leu
1 5

<210> 93
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 93

Leu Leu Ala Lys Leu Ala Glu Leu Val
1 5

<210> 94
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 94

Lys Leu Ala Glu Leu Val Ala Ala Ala
1 5

<210> 95
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 95

Leu Ala Glu Leu Val Ala Ala Ala Ile
1 5

<210> 96
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 96

Leu Val Ala Ala Ala Ile Ala Asp Ile
1 5

<210> 97
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 97

Ala Ile Ala Asp Ile Ile Ser Asp Val
1 5

<210> 98
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 98

Ile Ala Asp Ile Ile Ser Asp Val Ala
1 5

<210> 99
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 99

Ile Ser Asp Val Ala Asp Ile Ile Lys
1 5

<210> 100
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 100

Thr Leu Gly Glu Val Trp Glu Phe Ile
1 5

<210> 101
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 101

Phe Ile Thr Asn Ala Leu Asn Gly Leu
1 5

<210> 102
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 102

Asn Ala Leu Asn Gly Leu Lys Glu Leu
1 5

<210> 103
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 103

Lys Leu Thr Gly Trp Val Thr Gly Leu
1 5

<210> 104
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 104

Leu Thr Gly Trp Val Thr Gly Leu Phe
1 5

<210> 105
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 105

Gly Trp Ser Asn Leu Glu Ser Phe Phe
1 5

<210> 106

<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 106

Asn Leu Glu Ser Phe Phe Ala Gly Val
1 5

<210> 107
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 107

Ser Phe Phe Ala Gly Val Pro Gly Leu
1 5

<210> 108
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 108

Gly Leu Thr Gly Ala Thr Ser Gly Leu
1 5

<210> 109
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 109

Leu Ser Gln Val Thr Gly Leu Phe Gly
1 5

<210> 110
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 110

Ser Gln Val Thr Gly Leu Phe Gly Ala
1 5

<210> 111
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 111

Leu Ser Ala Ser Ser Gly Leu Ala His
1 5

<210> 112
<211> 9
<212> PRT

<213> *Mycobacterium tuberculosis*

<400> 112

Ala Ser Ser Gly Leu Ala His Ala Asp
1 5

<210> 113
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 113

Ser Gly Leu Ala His Ala Asp Ser Leu
1 5

<210> 114
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 114

Gly Leu Ala His Ala Asp Ser Leu Ala
1 5

<210> 115
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 115

Ser Leu Ala Ser Ser Ala Ser Leu Pro
1 5

<210> 116
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 116

Ala Ser Ser Ala Ser Leu Pro Ala Leu
1 5

<210> 117
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 117

Ser Gly Phe Gly Gly Leu Pro Ser Leu
1 5

<210> 118
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 118

Leu Pro Ser Leu Ala Gln Val His Ala
1 5

<210> 119
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 119

His Ala Ala Ser Thr Arg Gln Ala Leu
1 5

<210> 120
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 120

Ser Thr Arg Gln Ala Leu Arg Pro Arg
1 5

<210> 121
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 121

Arg Pro Arg Ala Asp Gly Pro Val Gly
1 5

<210> 122
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 122

Glu Gln Val Gly Gly Gln Ser Gln Leu
1 5

<210> 123
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 123

Gly Ala Ser Lys Gly Thr Thr Thr Lys
1 5

<210> 124
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 124

Ala Ser Lys Gly Thr Thr Thr Lys Lys
1 5

<210> 125
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 125

Lys Gly Thr Thr Thr Lys Lys Tyr Ser
1 5

<210> 126
<211> 9
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 126

Thr Glu Asp Ala Glu Arg Ala Pro Val
1 5

<210> 127
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 127

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1 5 10 15

Leu Tyr Asp Leu
20

<210> 128
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 128

Ile Asp Gly Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly
1 5 10 15

Gly Ile Leu Tyr
20

<210> 129
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 129

Asn Gln Gly Gly Ile Leu Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala
1 5 10 15

Leu Glu Glu Leu
20

<210> 130
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 130

Glu Lys Ala Leu Glu Glu Leu Ala Ala Ala Phe Pro Gly Asp Gly Trp
1 5 10 15

Leu Gly Ser Ala
20

<210> 131
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 131

Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala Gly Lys Asn Arg
1 5 10 15

Asn His Val Asn
20

<210> 132
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 132

Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu Asp
1 5 10 15

Arg Gln Leu Ile
20

<210> 133
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 133

Asp Leu Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala
1 5 10 15

Val Gln Thr Thr
20

<210> 134
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 134

Ala Asn Ala Val Gln Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys
1 5 10 15

Gly Leu Glu Phe
20

<210> 135
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 135

Ala Lys Lys Gly Leu Glu Phe Val Arg Pro Val Ala Val Asp Leu Thr
1 5 10 15

Tyr Ile Pro Val
20

<210> 136
<211> 27
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 136

Phe Val Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly
1 5 10 15

His Ala Leu Ser Ala Ala Phe Gln Ala Pro Phe
20 25

<210> 137
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 137

Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val Val Gly
1 5 10 15

Gly Ala Leu Ala
20

<210> 138
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 138

Val Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn
1 5 10 15

Ala Thr Gln Leu
20

<210> 139
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 139

Leu Ile Asn Ala Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu
1 5 10 15

Leu Val Ala Ala
20

<210> 140
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 140

Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp Ile Ile Ser Asp Val
1 5 10 15

Ala Asp Ile Ile
20

<210> 141
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 141

Ser Asp Val Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu Val Trp Glu
1 5 10 15

Phe Ile Thr Asn
20

<210> 142
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 142

Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp
1 5 10 15

Asp Lys Leu Thr
20

<210> 143
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 143

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
1 5 10 15

Gly Trp Ser Asn
20

<210> 144
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 144

Phe Ser Arg Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro
1 5 10 15

Gly Leu Thr Gly
20

<210> 145
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 145

Gly Val Pro Gly Leu Thr Gly Ala Thr Ser Gly Leu Ser Gln Val Thr
1 5 10 15

Gly Leu Phe Gly
20

<210> 146
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 146

Gln Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser Ser Gly
1 5 10 15

Leu Ala His Ala
20

<210> 147
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 147

Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu
1 5 10 15

Pro Ala Leu Ala
20

<210> 148
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 148

Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe Gly
1 5 10 15

Gly Leu Pro Ser
20

<210> 149
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 149

Gly Phe Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr
1 5 10 15

Arg Gln Ala Leu
20

<210> 150
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 150

Ala Ser Thr Arg Gln Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly
1 5 10 15

Ala Ala Ala Glu
20

<210> 151
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 151

Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln Ser Gln Leu Val
1 5 10 15

Ser Ala Gln Gly
20

<210> 152
<211> 19
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 152

Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val Gly
1 5 10 15

Met Gly Gly

<210> 153
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 153

Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser Lys Gly
1 5 10 15

Thr Thr Thr Lys
20

<210> 154
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 154

Ser Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly
1 5 10 15

Thr Glu Asp Ala
20

<210> 155
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 155

Ala Ala Gly Thr Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala
1 5 10 15

Gly Gly Gly Gln
20

<210> 156
<211> 20
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 156

Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val
1 5 10 15

Arg Asn Val Val
20

<210> 157
<211> 1053
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 157

```
Met Asn Phe Ser Val Leu Pro Pro Glu Ile Asn Ser Ala Leu Ile Phe
1               5                  10                  15

Ala Gly Ala Gly Pro Glu Pro Met Ala Ala Ala Ala Thr Ala Trp Asp
            20                  25                  30

Gly Leu Ala Met Glu Leu Ala Ser Ala Ala Ala Ser Phe Gly Ser Val
        35                  40                  45

Thr Ser Gly Leu Val Gly Gly Ala Trp Gln Gly Ala Ser Ser Ser Ala
    50                  55                  60

Met Ala Ala Ala Ala Ala Pro Tyr Ala Ala Trp Leu Ala Ala Ala Ala
65                  70                  75                  80

Val Gln Ala Glu Gln Thr Ala Ala Gln Ala Ala Ala Met Ile Ala Glu
                85                  90                  95

Phe Glu Ala Val Lys Thr Ala Val Val Gln Pro Met Leu Val Ala Ala
            100                 105                 110

Asn Arg Ala Asp Leu Val Ser Leu Val Met Ser Asn Leu Phe Gly Gln
        115                 120                 125

Asn Ala Pro Ala Ile Ala Ala Ile Glu Ala Thr Tyr Glu Gln Met Trp
    130                 135                 140

Ala Ala Asp Val Ser Ala Met Ser Ala Tyr His Ala Gly Ala Ser Ala
145                 150                 155                 160

Ile Ala Ser Ala Leu Ser Pro Phe Ser Lys Pro Leu Gln Asn Leu Ala
                165                 170                 175

Gly Leu Pro Ala Trp Leu Ala Ser Gly Ala Pro Ala Ala Ala Met Thr
            180                 185                 190

Ala Ala Ala Gly Ile Pro Ala Leu Ala Gly Gly Pro Thr Ala Ile Asn
        195                 200                 205

Leu Gly Ile Ala Asn Val Gly Gly Gly Asn Val Gly Asn Ala Asn Asn
    210                 215                 220

Gly Leu Ala Asn Ile Gly Asn Ala Asn Leu Gly Asn Tyr Asn Phe Gly
225                 230                 235                 240

Ser Gly Asn Phe Gly Asn Ser Asn Ile Gly Ser Ala Ser Leu Gly Asn
                245                 250                 255
```

```
Asn Asn Ile Gly Phe Gly Asn Leu Gly Ser Asn Asn Val Gly Val Gly
            260                 265                 270

Asn Leu Gly Asn Leu Asn Thr Gly Phe Ala Asn Thr Gly Leu Gly Asn
        275                 280                 285

Phe Gly Phe Gly Asn Thr Gly Asn Asn Asn Ile Gly Ile Gly Leu Thr
    290                 295                 300

Gly Asn Asn Gln Ile Gly Ile Gly Gly Leu Asn Ser Gly Thr Gly Asn
305                 310                 315                 320

Phe Gly Leu Phe Asn Ser Gly Ser Gly Asn Val Gly Phe Phe Asn Ser
                325                 330                 335

Gly Asn Gly Asn Phe Gly Ile Gly Asn Ser Gly Asn Phe Asn Thr Gly
            340                 345                 350

Gly Trp Asn Ser Gly His Gly Asn Thr Gly Phe Phe Asn Ala Gly Ser
        355                 360                 365

Phe Asn Thr Gly Met Leu Asp Val Gly Asn Ala Asn Thr Gly Ser Leu
    370                 375                 380

Asn Thr Gly Ser Tyr Asn Met Gly Asp Phe Asn Pro Gly Ser Ser Asn
385                 390                 395                 400

Thr Gly Thr Phe Asn Thr Gly Asn Ala Asn Thr Gly Phe Leu Asn Ala
                405                 410                 415

Gly Asn Ile Asn Thr Gly Val Phe Asn Ile Gly His Met Asn Asn Gly
            420                 425                 430

Leu Phe Asn Thr Gly Asp Met Asn Asn Gly Val Phe Tyr Arg Gly Val
        435                 440                 445

Gly Gln Gly Ser Leu Gln Phe Ser Ile Thr Thr Pro Asp Leu Thr Leu
    450                 455                 460

Pro Pro Leu Gln Ile Pro Gly Ile Ser Val Pro Ala Phe Ser Leu Pro
465                 470                 475                 480

Ala Ile Thr Leu Pro Ser Leu Asn Ile Pro Ala Ala Thr Thr Pro Ala
                485                 490                 495

Asn Ile Thr Val Gly Ala Phe Ser Leu Pro Gly Leu Thr Leu Pro Ser
            500                 505                 510
```

Leu Asn Ile Pro Ala Ala Thr Thr Pro Ala Asn Ile Thr Val Gly Ala
515 520 525

Phe Ser Leu Pro Gly Leu Thr Leu Pro Ser Leu Asn Ile Pro Ala Ala
530 535 540

Thr Thr Pro Ala Asn Ile Thr Val Gly Ala Phe Ser Leu Pro Gly Leu
545 550 555 560

Thr Leu Pro Ser Leu Asn Ile Pro Ala Ala Thr Thr Pro Ala Asn Ile
565 570 575

Thr Val Gly Ala Phe Ser Leu Pro Gly Leu Thr Leu Pro Ser Leu Asn
580 585 590

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

Leu Pro Gly Leu Thr Leu Pro Ser Leu Asn Ile Pro Ala Ala Thr Thr
610 615 620

Pro Ala Asn Ile Thr Val Ser Gly Phe Gln Leu Pro Pro Leu Ser Ile
625 630 635 640

Pro Ser Val Ala Ile Pro Pro Val Thr Val Pro Pro Ile Thr Val Gly
645 650 655

Ala Phe Asn Leu Pro Pro Leu Gln Ile Pro Glu Val Thr Ile Pro Gln
660 665 670

Leu Thr Ile Pro Ala Gly Ile Thr Ile Gly Gly Phe Ser Leu Pro Ala
675 680 685

Ile His Thr Gln Pro Ile Thr Val Gly Gln Ile Gly Val Gly Gln Phe
690 695 700

Gly Leu Pro Ser Ile Gly Trp Asp Val Phe Leu Ser Thr Pro Arg Ile
705 710 715 720

Thr Val Pro Ala Phe Gly Ile Pro Phe Thr Leu Gln Phe Gln Thr Asn
725 730 735

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

Gly Ala Leu Ile Phe Gly Glu Phe Asp Leu Pro Gln Leu Val Val His
755 760 765

```
Pro Tyr Thr Leu Thr Gly Pro Ile Val Ile Gly Ser Phe Phe Leu Pro
    770                 775                 780

Ala Phe Asn Ile Pro Gly Ile Asp Val Pro Ala Ile Asn Val Asp Gly
785                 790                 795                 800

Phe Thr Leu Pro Gln Ile Thr Thr Pro Ala Ile Thr Thr Pro Glu Phe
                805                 810                 815

Ala Ile Pro Pro Ile Gly Val Gly Gly Phe Thr Leu Pro Gln Ile Thr
            820                 825                 830

Thr Gln Glu Ile Ile Thr Pro Glu Leu Thr Ile Asn Ser Ile Gly Val
        835                 840                 845

Gly Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro
    850                 855                 860

Pro Leu Thr Ile Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu Pro Gln
865                 870                 875                 880

Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile Asp Pro Ile
                885                 890                 895

Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr
            900                 905                 910

Thr Pro Pro Leu Thr Ile Glu Pro Ile Gly Val Gly Gly Phe Thr Thr
        915                 920                 925

Pro Pro Leu Thr Val Pro Gly Ile His Leu Pro Ser Thr Thr Ile Gly
    930                 935                 940

Ala Phe Ala Ile Pro Gly Gly Pro Gly Tyr Phe Asn Ser Ser Thr Ala
945                 950                 955                 960

Pro Ser Ser Gly Phe Phe Asn Ser Gly Ala Gly Gly Asn Ser Gly Phe
                965                 970                 975

Gly Asn Asn Gly Ser Gly Leu Ser Gly Trp Phe Asn Thr Asn Pro Ala
            980                 985                 990

Gly Leu Leu Gly Gly Ser Gly Tyr  Gln Asn Phe Gly Gly  Leu Ser Ser
        995                 1000                1005

Gly Phe  Ser Asn Leu Gly Ser  Gly Val Ser Gly Phe  Ala Asn Arg
    1010                1015                1020

Gly Ile  Leu Pro Phe Ser Val  Ala Ser Val Val Ser  Gly Phe Ala
```

1025 1030 1035

Asn Ile Gly Thr Asn Leu Ala Gly Phe Phe Gln Gly Thr Thr Ser
1040 1045 1050

<210> 158
<211> 450
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 158

```
Met Ser Glu Leu Ser Val Ala Thr Gly Ala Val Ser Thr Ala Ser Ser
1               5                  10                  15

Ser Ile Pro Met Pro Ala Gly Val Asn Pro Ala Asp Leu Ala Ala Glu
            20                  25                  30

Leu Ala Ala Val Val Thr Glu Ser Val Asp Glu Asp Tyr Leu Leu Tyr
        35                  40                  45

Glu Cys Asp Gly Gln Trp Val Leu Ala Ala Gly Val Gln Ala Met Val
    50                  55                  60

Glu Leu Asp Ser Asp Glu Leu Arg Val Ile Arg Asp Gly Val Thr Arg
65                  70                  75                  80

Arg Gln Gln Trp Ser Gly Arg Pro Gly Ala Ala Leu Gly Glu Ala Val
                85                  90                  95

Asp Arg Leu Leu Leu Glu Thr Asp Gln Ala Phe Gly Trp Val Ala Phe
            100                 105                 110

Glu Phe Gly Val His Arg Tyr Gly Leu Gln Gln Arg Leu Ala Pro His
        115                 120                 125

Thr Pro Leu Ala Arg Val Phe Ser Pro Arg Thr Arg Ile Met Val Ser
    130                 135                 140

Glu Lys Glu Ile Arg Leu Phe Asp Ala Gly Ile Arg His Arg Glu Ala
145                 150                 155                 160

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

Ser Val Asp Val Ser Asp Asp Pro Ser Gly Phe Arg Arg Arg Val Ala
            180                 185                 190

Val Ala Val Asp Glu Ile Ala Ala Gly Arg Tyr His Lys Val Ile Leu
        195                 200                 205
```

```
Ser Arg Cys Val Glu Val Pro Phe Ala Ile Asp Phe Pro Leu Thr Tyr
    210                 215                 220

Arg Leu Gly Arg Arg His Asn Thr Pro Val Arg Ser Phe Leu Leu Gln
225                 230                 235                 240

Leu Gly Gly Ile Arg Ala Leu Gly Tyr Ser Pro Glu Leu Val Thr Ala
                245                 250                 255

Val Arg Ala Asp Gly Val Val Ile Thr Glu Pro Leu Ala Gly Thr Arg
            260                 265                 270

Ala Leu Gly Arg Gly Pro Ala Ile Asp Arg Leu Ala Arg Asp Asp Leu
        275                 280                 285

Glu Ser Asn Ser Lys Glu Ile Val Glu His Ala Ile Ser Val Arg Ser
    290                 295                 300

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

Ile Asp Phe Met Thr Val Arg Glu Arg Gly Ser Val Gln His Leu Gly
                325                 330                 335

Ser Thr Ile Arg Ala Arg Leu Asp Pro Ser Ser Asp Arg Met Ala Ala
            340                 345                 350

Leu Glu Ala Leu Phe Pro Ala Val Thr Ala Ser Gly Ile Pro Lys Ala
        355                 360                 365

Ala Gly Val Glu Ala Ile Phe Arg Leu Asp Glu Cys Pro Arg Gly Leu
    370                 375                 380

Tyr Ser Gly Ala Val Val Met Leu Ser Ala Asp Gly Gly Leu Asp Ala
385                 390                 395                 400

Ala Leu Thr Leu Arg Ala Ala Tyr Gln Val Gly Gly Arg Thr Trp Leu
                405                 410                 415

Arg Ala Gly Ala Gly Ile Ile Glu Glu Ser Glu Pro Glu Arg Glu Phe
            420                 425                 430

Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala
        435                 440                 445

Arg Gln
    450
```

<210> 159
<211> 324
<212> PRT
<213> *Mycobacterium tuberculosis*

<400> 159

```
Met Ser Asp Gln Val Pro Lys Pro His Arg His His Ile Trp Arg Ile
1               5                  10                  15

Thr Arg Arg Thr Leu Ser Lys Ser Trp Asp Asp Ser Ile Phe Ser Glu
            20                  25                  30

Ser Ala Gln Ala Ala Phe Trp Ser Ala Leu Ser Leu Pro Pro Leu Leu
        35                  40                  45

Leu Gly Met Leu Gly Ser Leu Ala Tyr Val Ala Pro Leu Phe Gly Pro
    50                  55                  60

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

Phe Phe Ser Pro Ser Val Val Asn Glu Ile Ile Glu Pro Thr Ile Gly
                85                  90                  95

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

Ile Ser Leu Trp Ala Gly Ser Ser Ala Ile Ser Ala Phe Val Asp Ala
        115                 120                 125

Val Val Glu Ala His Asp Gln Thr Pro Leu Arg His Pro Val Arg Gln
    130                 135                 140

Arg Phe Phe Ala Leu Phe Leu Tyr Val Val Met Leu Val Phe Leu Val
145                 150                 155                 160

Ala Thr Ala Pro Val Met Val Val Gly Pro Arg Lys Val Ser Glu His
                165                 170                 175

Ile Pro Glu Ser Leu Ala Asn Leu Leu Arg Tyr Gly Tyr Tyr Pro Ala
            180                 185                 190

Leu Ile Leu Gly Leu Thr Val Gly Val Ile Leu Leu Tyr Arg Val Ala
        195                 200                 205

Leu Pro Val Pro Leu Pro Thr His Arg Leu Val Leu Gly Ala Val Leu
    210                 215                 220

Ala Ile Ala Val Phe Leu Ile Ala Thr Leu Gly Leu Arg Val Tyr Leu
225                 230                 235                 240
```

```
Ala Trp Ile Thr Arg Thr Gly Tyr Thr Tyr Gly Ala Leu Ala Thr Pro
                245                 250                 255

Ile Ala Phe Leu Leu Phe Ala Phe Phe Gly Gly Phe Ala Ile Met Leu
            260                 265                 270

Gly Ala Glu Leu Asn Ala Ala Val Gln Glu Glu Trp Pro Ala Pro Ala
        275                 280                 285

Thr His Ala His Arg Leu Gly Asn Trp Leu Lys Ala Arg Ile Gly Val
    290                 295                 300

Gly Thr Thr Thr Tyr Ser Ser Thr Ala Gln His Ser Ala Val Ala Ala
305                 310                 315                 320

Glu Pro Pro Ser
```

<210> 160
<211> 1176
<212> ADN
<213> Secuencia artificial

<220>
<223> E. coli polinucleótido de H37Rv R3616c con optimización de codones

<400> 160

atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg 60

ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa 120

aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca 180

gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg 240

gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat 300

attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc 360

tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcaccgtt ttgtgccggt 420

gcaatggcag ttgtgggtgg tgctctggca tatctggttg tgaaaaccct gattaatgca 480

acccagctgc tgaaactgct ggcaaaactg gcagaactgg ttgcagcagc aattgcagat 540

attatttccg atgtggccga tattattaaa ggcaccctgg gcgaagtttg ggaatttatt 600

accaatgccc tgaatggtct gaaagaactg tgggataaac tgaccggttg ggttaccggt 660

ctgtttagcc gtggttggag caatctggaa tctttttttg ccggtgttcc gggtctgacc 720

ggtgcaacca gcggtctgag ccaggtgaca ggtctgtttg gagcagctgg tctgagtgct 780

agtagcggtc tggctcatgc agatagcctg gcaagcagcg catctctgcc tgcactggca 840

ggcattggtg gtggatccgg ttttggtggt ctgccgagcc tggcacaggt tcatgcagca 900

agcacccgtc aggcactgcg tccgcgtgca gatggaccgg ttggagcagc agcagaacag 960

gttggtggtc agagccagct ggttagcgca cagggtagcc agggtatggg tggtccggtg 1020

ggcatgggtg gtatgcatcc gagcagcggt gcaagcaaag gcaccaccac caaaaaatat 1080

agcgaaggag cagctgctgg caccgaagat gcagaacgtg caccggttga agcagatgcc 1140

ggtggaggtc agaaagttct ggttcgcaat gtggtg 1176

<210> 161
<211> 344
<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 136-183 suprimidos

<400> 161

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

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

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
145                 150                 155                 160

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
                165                 170                 175
```

```
Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
            180                 185                 190

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
        195                 200                 205

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
    210                 215                 220

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
225                 230                 235                 240

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
                245                 250                 255

Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln
            260                 265                 270

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
        275                 280                 285

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
    290                 295                 300

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
305                 310                 315                 320

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
                325                 330                 335

Lys Val Leu Val Arg Asn Val Val
            340
```

<210> 162
<211> 381
<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 150-160 suprimidos

<400> 162

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45
```

```
Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140

Val Gly Gly Ala Leu Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala
145                 150                 155                 160

Glu Leu Val Ala Ala Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp
                165                 170                 175

Ile Ile Lys Gly Thr Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala
            180                 185                 190

Leu Asn Gly Leu Lys Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr
        195                 200                 205

Gly Leu Phe Ser Arg Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly
    210                 215                 220

Val Pro Gly Leu Thr Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly
225                 230                 235                 240

Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala
                245                 250                 255

Asp Ser Leu Ala Ser Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly
            260                 265                 270

Gly Gly Ser Gly Phe Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala
        275                 280                 285

Ala Ser Thr Arg Gln Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly
    290                 295                 300
```

```
Ala Ala Ala Glu Gln Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln
305                 310                 315                 320

Gly Ser Gln Gly Met Gly Gly Pro Val Gly Met Gly Gly Met His Pro
                325                 330                 335

Ser Ser Gly Ala Ser Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly
            340                 345                 350

Ala Ala Ala Gly Thr Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp
        355                 360                 365

Ala Gly Gly Gly Gln Lys Val Leu Val Arg Asn Val Val
    370                 375                 380
```

<210> 163
<211> 373
<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 136-154 suprimidos

<400> 163

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Lys Thr Leu Ile Asn Ala Thr Gln Leu
    130                 135                 140
```

```
Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala Ala Ile Ala
145                 150                 155                 160

Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu
                165                 170                 175

Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp
            180                 185                 190

Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg Gly Trp Ser
        195                 200                 205

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

Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser
225                 230                 235                 240

Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser
                245                 250                 255

Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe Gly Gly Leu
            260                 265                 270

Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln Ala Leu Arg
        275                 280                 285

Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly
    290                 295                 300

Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro
305                 310                 315                 320

Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser Lys Gly Thr
                325                 330                 335

Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr Glu Asp Ala
            340                 345                 350

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

Val Arg Asn Val Val
    370
```

<210> 164
<211> 375

<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 166-182 suprimidos

<400> 164

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145                 150                 155                 160

Thr Gln Leu Leu Lys Ser Asp Val Ala Asp Ile Ile Lys Gly Thr Leu
                165                 170                 175

Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys Glu
            180                 185                 190

Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg Gly
        195                 200                 205

Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr Gly
    210                 215                 220

Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala Gly
225                 230                 235                 240
```

```
Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser Ser
                245                 250                 255

Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe Gly
            260                 265                 270

Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln Ala
        275                 280                 285

Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln Val
    290                 295                 300

Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met Gly
305                 310                 315                 320

Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser Lys
                325                 330                 335

Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr Glu
            340                 345                 350

Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln Lys
        355                 360                 365

Val Leu Val Arg Asn Val Val
    370                 375
```

<210> 165
<211> 387
<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 135-139 suprimidos

<400> 165

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80
```

```
Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Gly Ala Met Ala Val Val Gly Gly Ala Leu
    130                 135                 140

Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala Thr Gln Leu Leu Lys
145                 150                 155                 160

Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp Ile
                165                 170                 175

Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu Val Trp
            180                 185                 190

Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp Asp Lys
        195                 200                 205

Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg Gly Trp Ser Asn Leu
    210                 215                 220

Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr Gly Ala Thr Ser Gly
225                 230                 235                 240

Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser
                245                 250                 255

Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu Pro
            260                 265                 270

Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe Gly Gly Leu Pro Ser
        275                 280                 285

Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln Ala Leu Arg Pro Arg
    290                 295                 300

Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln Ser
305                 310                 315                 320

Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val Gly
                325                 330                 335
```

```
Met Gly Gly Met His Pro Ser Ser Gly Ala Ser Lys Gly Thr Thr Thr
            340                 345                 350

Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr Glu Asp Ala Glu Arg
        355                 360                 365

Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val Arg
    370                 375                 380

Asn Val Val
385
```

<210> 166
<211> 388
<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 142-145 suprimidos

<400> 166

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Gly Gly Ala
    130                 135                 140

Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala Thr Gln Leu Leu
145                 150                 155                 160
```

```
Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp
                165                 170                 175

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

Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp Asp
        195                 200                 205

Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg Gly Trp Ser Asn
    210                 215                 220

Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr Gly Ala Thr Ser
225                 230                 235                 240

Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala
                245                 250                 255

Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu
            260                 265                 270

Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe Gly Gly Leu Pro
        275                 280                 285

Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln Ala Leu Arg Pro
    290                 295                 300

Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln
305                 310                 315                 320

Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val
                325                 330                 335

Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser Lys Gly Thr Thr
            340                 345                 350

Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr Glu Asp Ala Glu
        355                 360                 365

Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val
    370                 375                 380

Arg Asn Val Val
385
```

<210> 167
<211> 384

<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 138-145 suprimido

<400> 167

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Gly Gly Ala Leu Ala Tyr Leu
    130                 135                 140

Val Val Lys Thr Leu Ile Asn Ala Thr Gln Leu Leu Lys Leu Leu Ala
145                 150                 155                 160

Lys Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp Ile Ile Ser Asp
                165                 170                 175

Val Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu Val Trp Glu Phe Ile
            180                 185                 190

Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp Asp Lys Leu Thr Gly
        195                 200                 205

Trp Val Thr Gly Leu Phe Ser Arg Gly Trp Ser Asn Leu Glu Ser Phe
    210                 215                 220

Phe Ala Gly Val Pro Gly Leu Thr Gly Ala Thr Ser Gly Leu Ser Gln
225                 230                 235                 240
```

```
Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser Ser Gly Leu
                245                 250                 255

Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu Pro Ala Leu Ala
            260                 265                 270

Gly Ile Gly Gly Gly Ser Gly Phe Gly Gly Leu Pro Ser Leu Ala Gln
        275                 280                 285

Val His Ala Ala Ser Thr Arg Gln Ala Leu Arg Pro Arg Ala Asp Gly
    290                 295                 300

Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln Ser Gln Leu Val
305                 310                 315                 320

Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val Gly Met Gly Gly
                325                 330                 335

Met His Pro Ser Ser Gly Ala Ser Lys Gly Thr Thr Thr Lys Lys Tyr
            340                 345                 350

Ser Glu Gly Ala Ala Ala Gly Thr Glu Asp Ala Glu Arg Ala Pro Val
        355                 360                 365

Glu Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val Arg Asn Val Val
    370                 375                 380
```

<210> 168
<211> 384
<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 145-152 suprimidos

<400> 168

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
```

65 70 75 80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
85 90 95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
100 105 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
115 120 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
130 135 140

Val Val Lys Thr Leu Ile Asn Ala Thr Gln Leu Leu Lys Leu Leu Ala
145 150 155 160

Lys Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp Ile Ile Ser Asp
165 170 175

Val Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu Val Trp Glu Phe Ile
180 185 190

Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp Asp Lys Leu Thr Gly
195 200 205

Trp Val Thr Gly Leu Phe Ser Arg Gly Trp Ser Asn Leu Glu Ser Phe
210 215 220

Phe Ala Gly Val Pro Gly Leu Thr Gly Ala Thr Ser Gly Leu Ser Gln
225 230 235 240

Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser Ser Gly Leu
245 250 255

Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu Pro Ala Leu Ala
260 265 270

Gly Ile Gly Gly Gly Ser Gly Phe Gly Gly Leu Pro Ser Leu Ala Gln
275 280 285

Val His Ala Ala Ser Thr Arg Gln Ala Leu Arg Pro Arg Ala Asp Gly
290 295 300

Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln Ser Gln Leu Val
305 310 315 320

Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val Gly Met Gly Gly
325 330 335

Met His Pro Ser Ser Gly Ala Ser Lys Gly Thr Thr Thr Lys Lys Tyr
340 345 350

Ser Glu Gly Ala Ala Ala Gly Thr Glu Asp Ala Glu Arg Ala Pro Val
355 360 365

Glu Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val Arg Asn Val Val
370 375 380

<210> 169
<211> 386
<212> PRT
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los residuos 149-154 suprimidos

<400> 169

```
Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1               5                  10                  15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            20                  25                  30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        35                  40                  45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    50                  55                  60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65                  70                  75                  80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                85                  90                  95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            100                 105                 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        115                 120                 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    130                 135                 140

Val Gly Gly Ala Lys Thr Leu Ile Asn Ala Thr Gln Leu Leu Lys Leu
145                 150                 155                 160

Leu Ala Lys Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp Ile Ile
```

165 170 175

Ser Asp Val Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu Val Trp Glu
180 185 190

Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp Asp Lys Leu
195 200 205

Thr Gly Trp Val Thr Gly Leu Phe Ser Arg Gly Trp Ser Asn Leu Glu
210 215 220

Ser Phe Phe Ala Gly Val Pro Gly Leu Thr Gly Ala Thr Ser Gly Leu
225 230 235 240

Ser Gln Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser Ser
245 250 255

Gly Leu Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu Pro Ala
260 265 270

Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe Gly Gly Leu Pro Ser Leu
275 280 285

Ala Gln Val His Ala Ala Ser Thr Arg Gln Ala Leu Arg Pro Arg Ala
290 295 300

Asp Gly Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln Ser Gln
305 310 315 320

Leu Val Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val Gly Met
325 330 335

Gly Gly Met His Pro Ser Ser Gly Ala Ser Lys Gly Thr Thr Thr Lys
340 345 350

Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr Glu Asp Ala Glu Arg Ala
355 360 365

Pro Val Glu Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val Arg Asn
370 375 380

Val Val
385

<210> 170
<211> 1032
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 136-183

<400> 170

atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg 60

ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa 120

aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca 180

gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg 240

gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat 300

attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc 360

tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcagatgt ggccgatatt 420

attaaaggca ccctgggcga agtttgggaa tttattacca atgccctgaa tggtctgaaa 480

gaactgtggg ataaactgac cggttgggtt accggtctgt ttagccgtgg ttggagcaat 540

ctggaatctt tttttgccgg tgttccgggt ctgaccggtg caaccagcgg tctgagccag 600

gtgacaggtc tgtttggagc agctggtctg agtgctagta gcggtctggc tcatgcagat 660

agcctggcaa gcagcgcatc tctgcctgca ctggcaggca ttggtggtgg atccggtttt 720

ggtggtctgc cgagcctggc acaggttcat gcagcaagca cccgtcaggc actgcgtccg 780

cgtgcagatg gaccggttgg agcagcagca gaacaggttg gtggtcagag ccagctggtt 840

agcgcacagg gtagccaggg tatgggtggt ccggtgggca tgggtggtat gcatccgagc 900

agcggtgcaa gcaaaggcac caccaccaaa aaatatagcg aaggagcagc tgctggcacc 960

gaagatgcag aacgtgcacc ggttgaagca gatgccggtg gaggtcagaa agttctggtt 1020

cgcaatgtgg tg 1032

<210> 171
<211> 1143
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 150-160

<400> 171

```
atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg       60
ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa      120
aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca      180
gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg      240
gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat      300
attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc      360
tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcaccgtt ttgtgccggt      420
gcaatggcag ttgtgggtgg tgctctgacc cagctgctga aactgctggc aaaactggca      480

gaactggttg cagcagcaat tgcagatatt atttccgatg tggccgatat tattaaaggc      540
accctgggcg aagtttggga atttattacc aatgccctga atggtctgaa agaactgtgg      600
gataaactga ccggttgggt taccggtctg tttagccgtg gttggagcaa tctggaatct      660
ttttttgccg gtgttccggg tctgaccggt gcaaccagcg gtctgagcca ggtgacaggt      720
ctgtttggag cagctggtct gagtgctagt agcggtctgg ctcatgcaga tagcctggca      780
agcagcgcat ctctgcctgc actggcaggc attggtggtg gatccggttt tggtggtctg      840
ccgagcctgg cacaggttca tgcagcaagc acccgtcagg cactgcgtcc gcgtgcagat      900
ggaccggttg gagcagcagc agaacaggtt ggtggtcaga gccagctggt tagcgcacag      960
ggtagccagg gtatgggtgg tccggtgggc atgggtggta tgcatccgag cagcggtgca     1020
agcaaaggca ccaccaccaa aaaatatagc gaaggagcag ctgctggcac cgaagatgca     1080
gaacgtgcac cggttgaagc agatgccggt ggaggtcaga aagttctggt tcgcaatgtg     1140
gtg                                                                  1143
```

<210> 172
<211> 1119
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 136-154

<400> 172

atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg 60

ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa 120

aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca 180

gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg 240

gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat 300

attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc 360

tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcaaaaac cctgattaat 420

gcaacccagc tgctgaaact gctggcaaaa ctggcagaac tggttgcagc agcaattgca 480

gatattattt ccgatgtggc cgatattatt aaaggcaccc tgggcgaagt ttgggaattt 540

attaccaatg ccctgaatgg tctgaaagaa ctgtgggata aactgaccgg ttgggttacc 600

ggtctgttta gccgtggttg gagcaatctg gaatcttttt ttgccggtgt tccgggtctg 660

accggtgcaa ccagcggtct gagccaggtg acaggtctgt ttggagcagc tggtctgagt 720

gctagtagcg gtctggctca tgcagatagc ctggcaagca gcgcatctct gcctgcactg 780

gcaggcattg gtggtggatc cggttttggt ggtctgccga gcctggcaca ggttcatgca 840

gcaagcaccc gtcaggcact gcgtccgcgt gcagatggac cggttggagc agcagcagaa 900

caggttggtg gtcagagcca gctggttagc gcacagggta gccagggtat gggtggtccg 960

gtgggcatgg gtggtatgca tccgagcagc ggtgcaagca aaggcaccac caccaaaaaa 1020

tatagcgaag gagcagctgc tggcaccgaa gatgcagaac gtgcaccggt tgaagcagat 1080

gccggtggag gtcagaaagt tctggttcgc aatgtggtg 1119

<210> 173
<211> 1125
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 166-182

<400> 173

atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg 60

ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa 120

aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca 180

gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg 240

gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat 300

attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc 360

tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcaccgtt ttgtgccggt 420

gcaatggcag ttgtgggtgg tgctctggca tatctggttg tgaaaaccct gattaatgca 480

acccagctgc tgaaatccga tgtggccgat attattaaag gcaccctggg cgaagtttgg 540

gaatttatta ccaatgccct gaatggtctg aaagaactgt gggataaact gaccggttgg 600

gttaccggtc tgtttagccg tggttggagc aatctggaat ctttttttgc cggtgttccg 660

ggtctgaccg gtgcaaccag cggtctgagc caggtgacag gtctgtttgg agcagctggt 720

ctgagtgcta gtagcggtct ggctcatgca gatagcctgg caagcagcgc atctctgcct 780

gcactggcag gcattggtgg tggatccggt tttggtggtc tgccgagcct ggcacaggtt 840

catgcagcaa gcacccgtca ggcactgcgt ccgcgtgcag atggaccggt tggagcagca 900

gcagaacagg ttggtggtca gagccagctg gttagcgcac agggtagcca gggtatgggt 960

ggtccggtgg gcatgggtgg tatgcatccg agcagcggtg caagcaaagg caccaccacc 1020

aaaaaatata gcgaaggagc agctgctggc accgaagatg cagaacgtgc accggttgaa 1080

gcagatgccg gtggaggtca gaaagttctg gttcgcaatg tggtg 1125

<210> 174
<211> 1161
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 135-139

<400> 174

```
atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg      60
ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa     120
aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca     180
gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg     240
gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat     300
attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc     360
tatattccgg ttgttggtca tgcactgagc gcagcatttc agggtgcaat ggcagttgtg     420
ggtggtgctc tggcatatct ggttgtgaaa accctgatta atgcaaccca gctgctgaaa     480
ctgctggcaa aactggcaga actggttgca gcagcaattg cagatattat ttccgatgtg     540
gccgatatta ttaaaggcac cctgggcgaa gtttgggaat ttattaccaa tgccctgaat     600
ggtctgaaag aactgtggga taaactgacc ggttgggtta ccggtctgtt tagccgtggt     660
tggagcaatc tggaatcttt ttttgccggt gttccgggtc tgaccggtgc aaccagcggt     720
ctgagccagg tgacaggtct gtttggagca gctggtctga gtgctagtag cggtctggct     780
catgcagata gcctggcaag cagcgcatct ctgcctgcac tggcaggcat tggtggtgga     840
tccggttttg gtggtctgcc gagcctggca caggttcatg cagcaagcac ccgtcaggca     900
ctgcgtccgc gtgcagatgg accggttgga gcagcagcag aacaggttgg tggtcagagc     960
cagctggtta gcgcacaggg tagccagggt atgggtggtc cggtgggcat gggtggtatg    1020
catccgagca gcggtgcaag caaaggcacc accaccaaaa aatatagcga aggagcagct    1080
gctggcaccg aagatgcaga acgtgcaccg gttgaagcag atgccggtgg aggtcagaaa    1140
gttctggttc gcaatgtggt g                                              1161
```

<210> 175
<211> 1164
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 142-145

<400> 175

```
atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg      60

ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa     120

aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca     180

gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg     240

gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat     300

attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc     360

tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcaccgtt ttgtgccggt     420

gcaggtggtg ctctggcata tctggttgtg aaaaccctga ttaatgcaac ccagctgctg     480

aaactgctgg caaaactggc agaactggtt gcagcagcaa ttgcagatat tatttccgat     540

gtggccgata ttattaaagg caccctgggc gaagtttggg aatttattac caatgccctg     600

aatggtctga aagaactgtg ggataaactg accggttggg ttaccggtct gtttagccgt     660

ggttggagca atctggaatc ttttttttgcc ggtgttccgg gtctgaccgg tgcaaccagc     720

ggtctgagcc aggtgacagg tctgtttgga gcagctggtc tgagtgctag tagcggtctg     780

gctcatgcag atagcctggc aagcagcgca tctctgcctg cactggcagg cattggtggt     840

ggatccggtt ttggtggtct gccgagcctg gcacaggttc atgcagcaag cacccgtcag     900

gcactgcgtc cgcgtgcaga tggaccggtt ggagcagcag cagaacaggt tggtggtcag     960

agccagctgg ttagcgcaca gggtagccag ggtatgggtg gtccggtggg catgggtggt    1020

atgcatccga gcagcggtgc aagcaaaggc accaccacca aaaaatatag cgaaggagca    1080

gctgctggca ccgaagatgc agaacgtgca ccggttgaag cagatgccgg tggaggtcag    1140

aaagttctgg ttcgcaatgt ggtg                                           1164
```

<210> 176
<211> 1152
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 138-145

<400> 176

atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg 60

ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa 120

aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca 180

gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg 240

gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat 300

attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc 360

tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcaccgtt tggtggtgct 420

ctggcatatc tggttgtgaa aaccctgatt aatgcaaccc agctgctgaa actgctggca 480

aaactggcag aactggttgc agcagcaatt gcagatatta tttccgatgt ggccgatatt 540

attaaaggca ccctgggcga agtttgggaa tttattacca atgccctgaa tggtctgaaa 600

gaactgtggg ataaactgac cggttgggtt accggtctgt ttagccgtgg ttggagcaat 660

ctggaatctt tttttgccgg tgttccgggt ctgaccggtg caaccagcgg tctgagccag 720

gtgacaggtc tgtttggagc agctggtctg agtgctagta gcggtctggc tcatgcagat 780

agcctggcaa gcagcgcatc tctgcctgca ctggcaggca ttggtggtgg atccggtttt 840

ggtggtctgc cgagcctggc acaggttcat gcagcaagca cccgtcaggc actgcgtccg 900

cgtgcagatg gaccggttgg agcagcagca gaacaggttg gtggtcagag ccagctggtt 960

agcgcacagg gtagccaggg tatgggtggt ccggtgggca tgggtggtat gcatccgagc 1020

agcggtgcaa gcaaaggcac caccaccaaa aaatatagcg aaggagcagc tgctggcacc 1080

gaagatgcag aacgtgcacc ggttgaagca gatgccggtg gaggtcagaa agttctggtt 1140

cgcaatgtgg tg 1152

<210> 177
<211> 1152
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 145-152

<400> 177

```
atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg      60
ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa     120
aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca     180
gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg     240
gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat     300
attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc     360
tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcaccgtt ttgtgccggt     420
gcaatggcag ttgttgtgaa aaccctgatt aatgcaaccc agctgctgaa actgctggca     480
aaactggcag aactggttgc agcagcaatt gcagatatta tttccgatgt ggccgatatt     540
attaaaggca ccctgggcga agtttgggaa tttattacca atgccctgaa tggtctgaaa     600
gaactgtggg ataaactgac cggttgggtt accggtctgt ttagccgtgg ttggagcaat     660
ctggaatctt tttttgccgg tgttccgggt ctgaccggtg caaccagcgg tctgagccag     720
gtgacaggtc tgtttggagc agctggtctg agtgctagta gcggtctggc tcatgcagat     780
agcctggcaa gcagcgcatc tctgcctgca ctggcaggca ttggtggtgg atccggtttt     840
ggtggtctgc cgagcctggc acaggttcat gcagcaagca cccgtcaggc actgcgtccg     900
cgtgcagatg gaccggttgg agcagcagca gaacaggttg gtggtcagag ccagctggtt     960
agcgcacagg gtagccaggg tatgggtggt ccggtgggca tgggtggtat gcatccgagc    1020
agcggtgcaa gcaaaggcac caccaccaaa aaatatagcg aaggagcagc tgctggcacc    1080
gaagatgcag aacgtgcacc ggttgaagca gatgccggtg gaggtcagaa agttctggtt    1140
cgcaatgtgg tg                                                        1152
```

<210> 178
<211> 1158
<212> ADN
<213> Secuencia artificial

<220>
<223> H37Rv Rv3616c con los codones suprimidos para los residuos 149-154

<400> 178

```
atgagccgtg cctttattat tgatccgacc attagcgcaa ttgatggtct gtatgatctg       60
ctgggtattg gtattccgaa tcagggtggt attctgtata gcagcctgga atattttgaa      120
aaagccctgg aagaactggc agcagcattt ccgggtgatg gttggctggg tagcgcagca      180
gataaatatg ccggtaaaaa tcgcaatcat gtgaattttt ttcaggaact ggccgatctg      240
gatcgtcagc tgattagcct gattcatgac caggcaaatg cagttcagac cacccgtgat      300
attctggaag gtgccaaaaa aggtctggaa tttgttcgtc cggtggcagt tgatctgacc      360
tatattccgg ttgttggtca tgcactgagc gcagcatttc aggcaccgtt ttgtgccggt      420
gcaatggcag ttgtgggtgg tgctaaaacc ctgattaatg caacccagct gctgaaactg      480
ctggcaaaac tggcagaact ggttgcagca gcaattgcag atattatttc cgatgtggcc      540
gatattatta aaggcaccct gggcgaagtt tgggaattta ttaccaatgc cctgaatggt      600
ctgaaagaac tgtgggataa actgaccggt tgggttaccg gtctgtttag ccgtggttgg      660
agcaatctgg aatctttttt tgccggtgtt ccgggtctga ccggtgcaac cagcggtctg      720
agccaggtga caggtctgtt tggagcagct ggtctgagtg ctagtagcgg tctggctcat      780
gcagatagcc tggcaagcag cgcatctctg cctgcactgg caggcattgg tggtggatcc      840
ggttttggtg gtctgccgag cctggcacag gttcatgcag caagcacccg tcaggcactg      900
cgtccgcgtg cagatggacc ggttggagca gcagcagaac aggttggtgg tcagagccag      960
ctggttagcg cacagggtag ccagggtatg ggtggtccgg tgggcatggg tggtatgcat     1020
ccgagcagcg gtgcaagcaa aggcaccacc accaaaaaat atagcgaagg agcagctgct     1080
ggcaccgaag atgcagaacg tgcaccggtt gaagcagatg ccggtggagg tcagaaagtt     1140
ctggttcgca atgtggtg                                                  1158
```

<210> 179
<211> 391
<212> PRT
<213> Secuencia artificial

<220>
<223> Reordenación alrededor de los residuos 137-139 de la cepa H37Rv de M. tuberculosis , incluyendo la supresión de Cys138
secuencia de polipéptido de proteína Rv3616c modificada basada en la separación y reordenación alrededor de los residuos 137-139 a partir de la cepa H37Rv de *M. tuberculosis*, incluyendo la supresión de Cys138
<400> 179

```
Met Ala Gly Ala Met Ala Val Val Gly Gly Ala Leu Ala Tyr Leu Val
1               5                  10                  15

Val Lys Thr Leu Ile Asn Ala Thr Gln Leu Leu Lys Leu Leu Ala Lys
            20                  25                  30
```

Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp Ile Ile Ser Asp Val
35 40 45

Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu Val Trp Glu Phe Ile Thr
50 55 60

Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp Asp Lys Leu Thr Gly Trp
65 70 75 80

Val Thr Gly Leu Phe Ser Arg Gly Trp Ser Asn Leu Glu Ser Phe Phe
85 90 95

Ala Gly Val Pro Gly Leu Thr Gly Ala Thr Ser Gly Leu Ser Gln Val
100 105 110

Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser Ser Gly Leu Ala
115 120 125

His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu Pro Ala Leu Ala Gly
130 135 140

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

His Ala Ala Ser Thr Arg Gln Ala Leu Arg Pro Arg Ala Asp Gly Pro
165 170 175

Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln Ser Gln Leu Val Ser
180 185 190

Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val Gly Met Gly Gly Met
195 200 205

His Pro Ser Ser Gly Ala Ser Lys Gly Thr Thr Thr Lys Lys Tyr Ser
210 215 220

Glu Gly Ala Ala Ala Gly Thr Glu Asp Ala Glu Arg Ala Pro Val Glu
225 230 235 240

Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val Arg Asn Val Val Ser
245 250 255

Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly Leu Tyr
260 265 270

Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu Tyr Ser
275 280 285

Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala Ala Phe

290 295 300

Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala Gly Lys
305 310 315 320

Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu Asp Arg
325 330 335

Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln Thr Thr
340 345 350

Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val Arg Pro
355 360 365

Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala Leu Ser
370 375 380

Ala Ala Phe Gln Ala Pro Phe
385 390

<210> 180
<211> 392
<212> PRT
<213> Secuencia artificial

<220>
<223> Reordenación alrededor de los residuos 152-153 de la cepa H37Rv de M. tuberculosis

<400> 180

```
Met Val Val Lys Thr Leu Ile Asn Ala Thr Gln Leu Leu Lys Leu Leu
1               5                  10                  15

Ala Lys Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp Ile Ile Ser
            20                  25                  30

Asp Val Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu Val Trp Glu Phe
        35                  40                  45

Ile Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp Asp Lys Leu Thr
    50                  55                  60

Gly Trp Val Thr Gly Leu Phe Ser Arg Gly Trp Ser Asn Leu Glu Ser
65                  70                  75                  80

Phe Phe Ala Gly Val Pro Gly Leu Thr Gly Ala Thr Ser Gly Leu Ser
                85                  90                  95

Gln Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser Ser Gly
            100                 105                 110
```

```
Leu Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu Pro Ala Leu
        115                 120                 125

Ala Gly Ile Gly Gly Gly Ser Gly Phe Gly Gly Leu Pro Ser Leu Ala
    130                 135                 140

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

Gly Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln Ser Gln Leu
                165                 170                 175

Val Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val Gly Met Gly
            180                 185                 190

Gly Met His Pro Ser Ser Gly Ala Ser Lys Gly Thr Thr Thr Lys Lys
        195                 200                 205

Tyr Ser Glu Gly Ala Ala Ala Gly Thr Glu Asp Ala Glu Arg Ala Pro
    210                 215                 220

Val Glu Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val Arg Asn Val
225                 230                 235                 240

Val Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
                245                 250                 255

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
            260                 265                 270

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
        275                 280                 285

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
    290                 295                 300

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
305                 310                 315                 320

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
                325                 330                 335

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
            340                 345                 350

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
        355                 360                 365
```

```
Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
    370                 375                 380

Val Gly Gly Ala Leu Ala Tyr Leu
385                 390
```

REIVINDICACIONES

1. Una proteína Rv3616c modificada, en la que se ha alterado la hidrofobicidad de los residuos de aminoácidos correspondientes a los residuos 134-183 de la secuencia de H37Rv, comprendiendo dicha proteína Rv3616c modificada un primer polipéptido y un segundo polipéptido, estándo el primer polipéptido localizado hacia el extremo N con respecto al segundo polipéptido, y en la que:

(i) el primer polipéptido es una secuencia que tiene al menos una identidad del 98 % con los residuos 1-134 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos una identidad del 98 % con los residuos 155-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están directamente unidos o indirectamente unidos por medio de un tercer polipéptido, teniendo dicho tercer polipéptido al menos una identidad del 80 % con una secuencia que corresponde a los residuos 135-154 de la SEQ ID NO: 1, en la que se ha suprimido una porción contigua de al menos 3 aminoácidos.

2. La proteína Rv3616c modificada de acuerdo con la reivindicación 1, en la que el enlace peptídico es indirecto.

3. La proteína Rv3616c modificada de acuerdo con la reivindicación 2, en la que el tercer polipéptido es una secuencia que tiene al menos una identidad del 95 % con una secuencia correspondiente a los residuos 135-154 de la SEQ ID NO: 1, en la que se ha suprimido una porción contigua de al menos 3 aminoácidos.

4. La proteína Rv3616c modificada de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en la que la porción contigua suprimida de los residuos correspondientes a 135-154 en la SEQ ID NO: 1 es de al menos 4 aminoácidos.

5. La proteína Rv3616c modificada de acuerdo con la reivindicación 4, en la que la porción contigua suprimida de los residuos correspondientes a 135-154 en la SEQ ID NO: 1 es de al menos 5 aminoácidos.

6. La proteína Rv3616c modificada de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en la que el tercer polipéptido tiene 20 aminoácidos o menos de longitud.

7. La proteína Rv3616c modificada de acuerdo con una cualquiera de las reivindicaciones 1 a 6, teniendo dicha proteína menos de 500 residuos de aminoácidos de longitud.

8. La proteína Rv3616c modificada de acuerdo con la reivindicación 1, que comprende la secuencia de aminoácidos de SEQ ID NO: 163.

9. La proteína Rv3616c modificada de acuerdo con la reivindicación 1, que comprende la secuencia de aminoácidos de SEQ ID NO: 165.

10. La proteína Rv3616c modificada de acuerdo con la reivindicación 1, que comprende la secuencia de aminoácidos de SEQ ID NO: 166.

11. La proteína Rv3616c modificada de acuerdo con la reivindicación 1, que comprende la secuencia de aminoácidos de SEQ ID NO: 167.

12. La proteína Rv3616c modificada de acuerdo con la reivindicación 1, que comprende la secuencia de aminoácidos de SEQ ID NO: 168.

13. La proteína Rv3616c modificada de acuerdo con la reivindicación 1, que comprende la secuencia de aminoácidos de SEQ ID NO: 169.

14. La proteína Rv3616c modificada de acuerdo con una cualquiera de las reivindicaciones 1 a 13, en la que dicha proteína consiste en un primer polipéptido y un segundo polipéptido, estando el primer polipéptido localizado hacia el extremo N con respecto al segundo polipéptido, y en la que:

(i) el primer polipéptido es una secuencia que tiene al menos una identidad del 98 % con los residuos 1-134 de la SEQ ID NO: 1; y
(ii) el segundo polipéptido es una secuencia que tiene al menos una identidad del 98 % con los residuos 155-392 de la SEQ ID NO: 1;

en la que los primero y segundo polipéptidos están directamente unidos o indirectamente unidos por medio de un tercer polipéptido, teniendo dicho tercer polipéptido al menos una identidad del 80 % con una secuencia que corresponde a los residuos 135-154 de la SEQ ID NO: 1, en la que se ha suprimido una porción contigua de al menos 3 aminoácidos.

15. Un polinucleótido que comprende una secuencia de ácido nucleico que codifica una proteína Rv3616c modificada de acuerdo con una cualquiera de las reivindicaciones 1 a 14.

16. El polinucleótido de acuerdo con la reivindicación 15, que comprende una secuencia que codifica un polipéptido de acuerdo con una cualquiera de las SEQ ID NO: 163 o 165 a 169.

17. La proteína Rv3616c modificada de acuerdo con una cualquiera de las reivindicaciones 1 a 14, o el polinucleótido de acuerdo con la reivindicación 15 o 16 para su uso como medicamento.

18. Uso de una proteína Rv3616c modificada de acuerdo con una cualquiera de las reivindicaciones 1 a 14 o el polinucleótido de acuerdo con la reivindicación 15 o 16 en la elaboración de un medicamento para el tratamiento, la mejora o la prevención de TB.

19. Una composición inmunogénica, que comprende:

(a) una proteína Rv3616c modificada de acuerdo con una cualquiera de las reivindicaciones 1 a 14 o un polinucleótido de acuerdo con la reivindicación 15 o 16;
y
(b) un potenciador no específico de la respuesta inmunitaria.

Figura 1

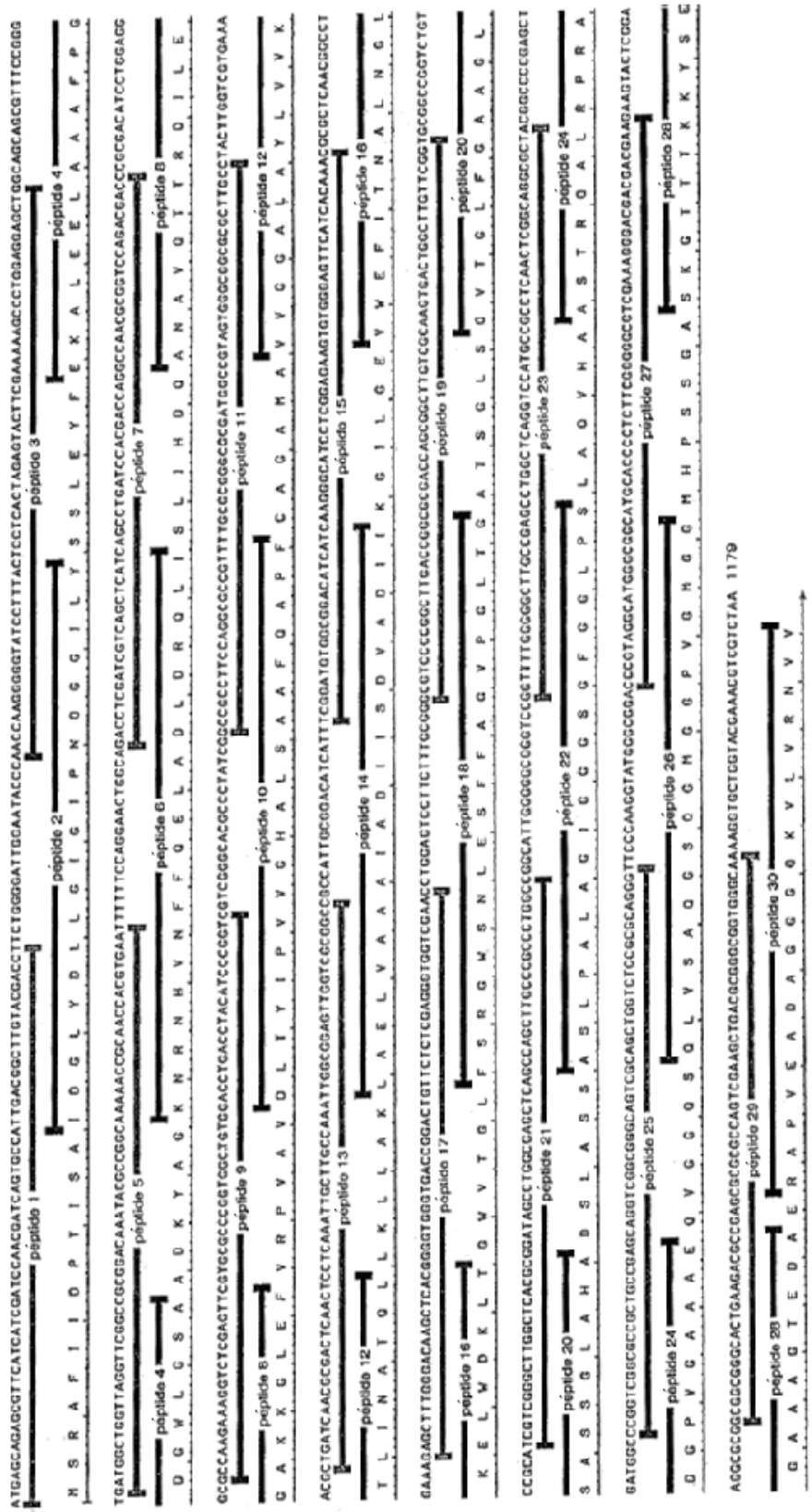


Figura 2

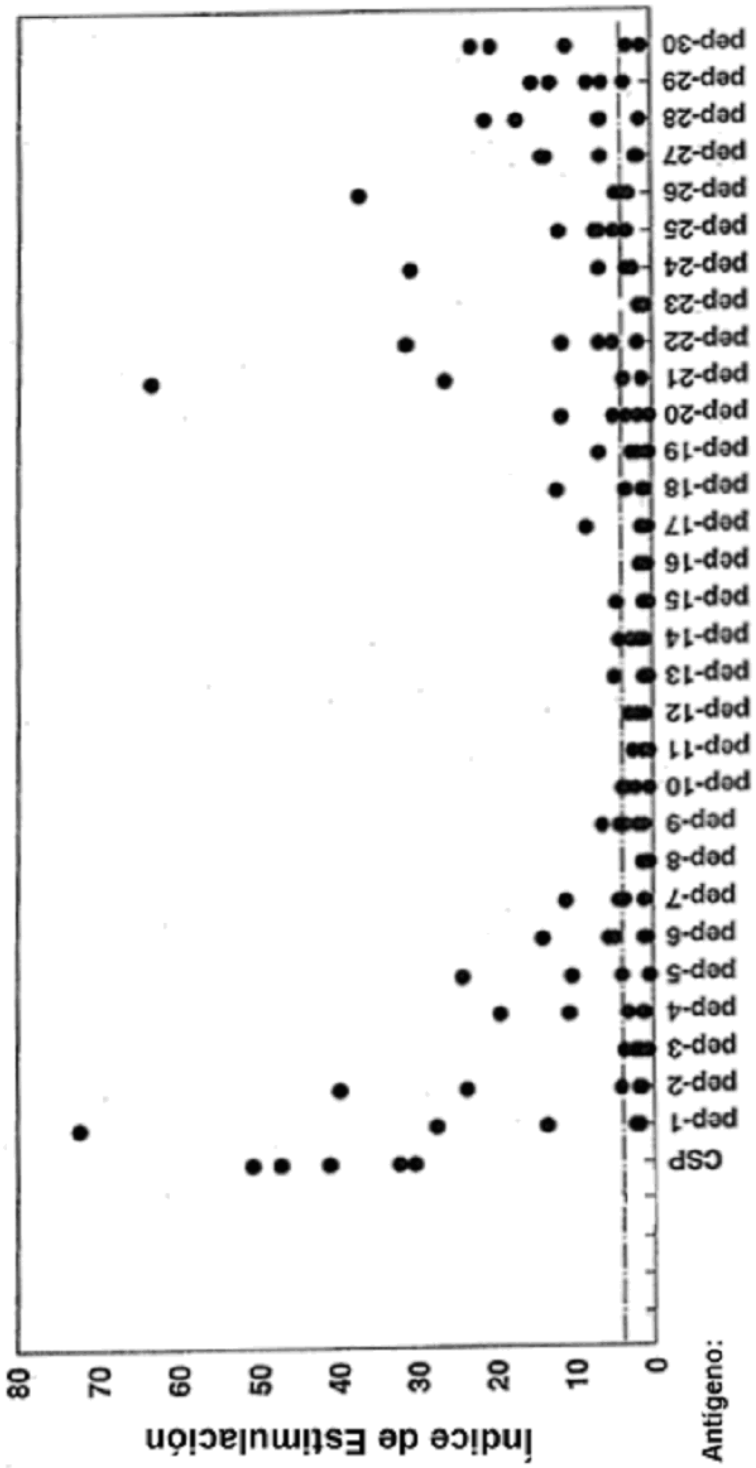

Antigeno:
CSP
pep-1
pep-2
pep-3
pep-4
pep-5
pep-6
pep-7
pep-8
pep-9
pep-10
pep-11
pep-12
pep-13
pep-14
pep-15
pep-16
pep-17
pep-18
pep-19
pep-20
pep-21
pep-22
pep-23
pep-24
pep-25
pep-26
pep-27
pep-28
pep-29
pep-30
Indice de Estimulación
0
10
20
30
40
50
60
70
80

Figura 3

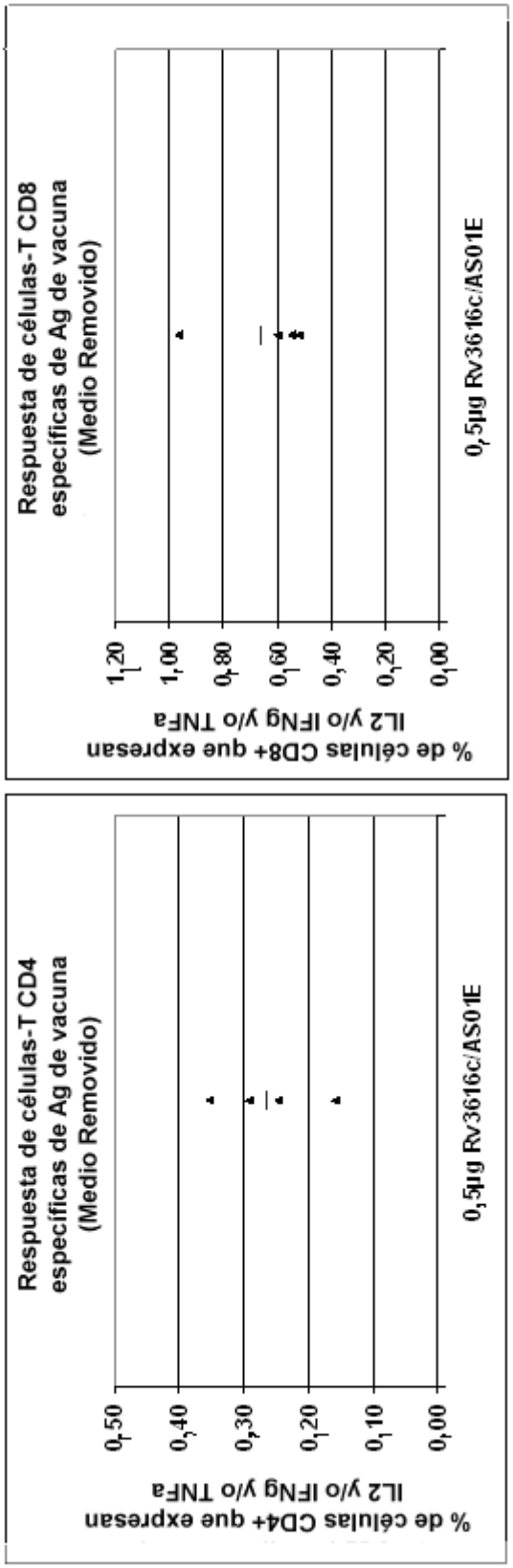

Respuesta de células-T CD4 específicas de Ag de vacuna (Medio Removido)
% de células CD4+ que expresan IL2 y/o IFNg y/o TNFa
0,00
0,10
0,20
0,30
0,40
0,50
0,5µg Rv3616c/AS01E
Respuesta de células-T CD8 específicas de Ag de vacuna (Medio Removido)
% de células CD8+ que expresan IL2 y/o IFNg y/o TNFa
0,00
0,20
0,40
0,60
0,80
1,00
1,20
0,5µg Rv3616c/AS01E

Figura 4

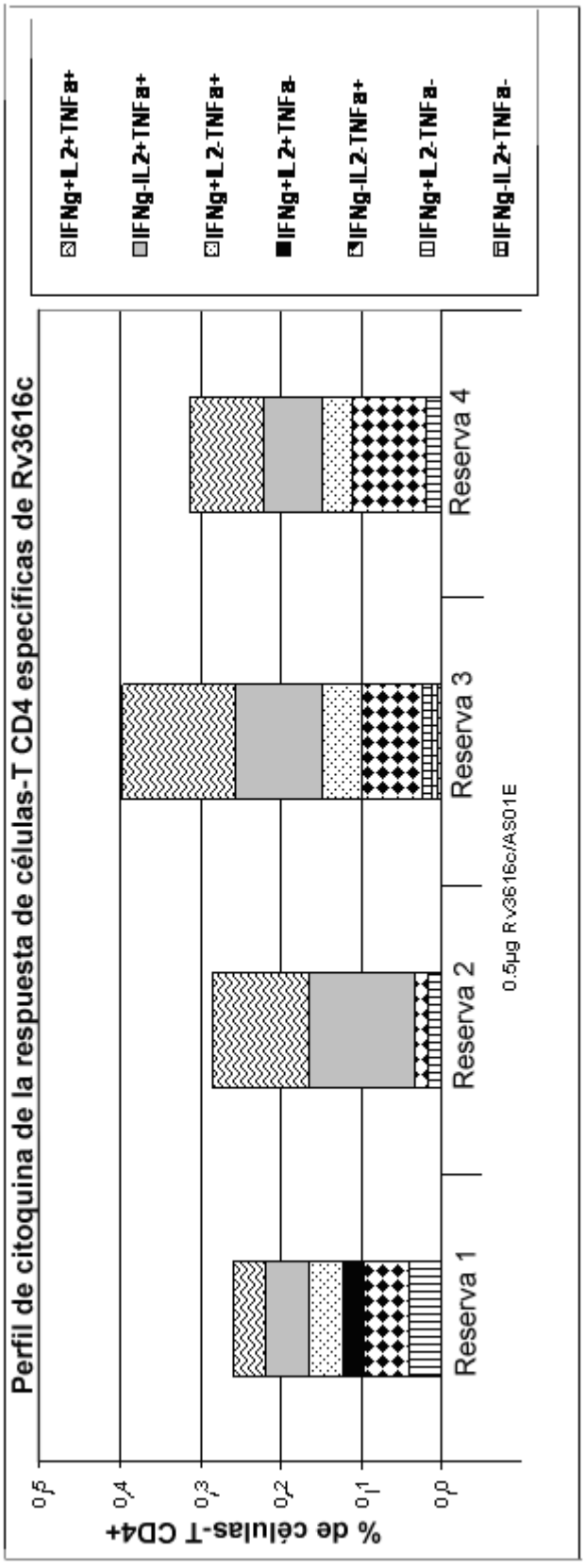
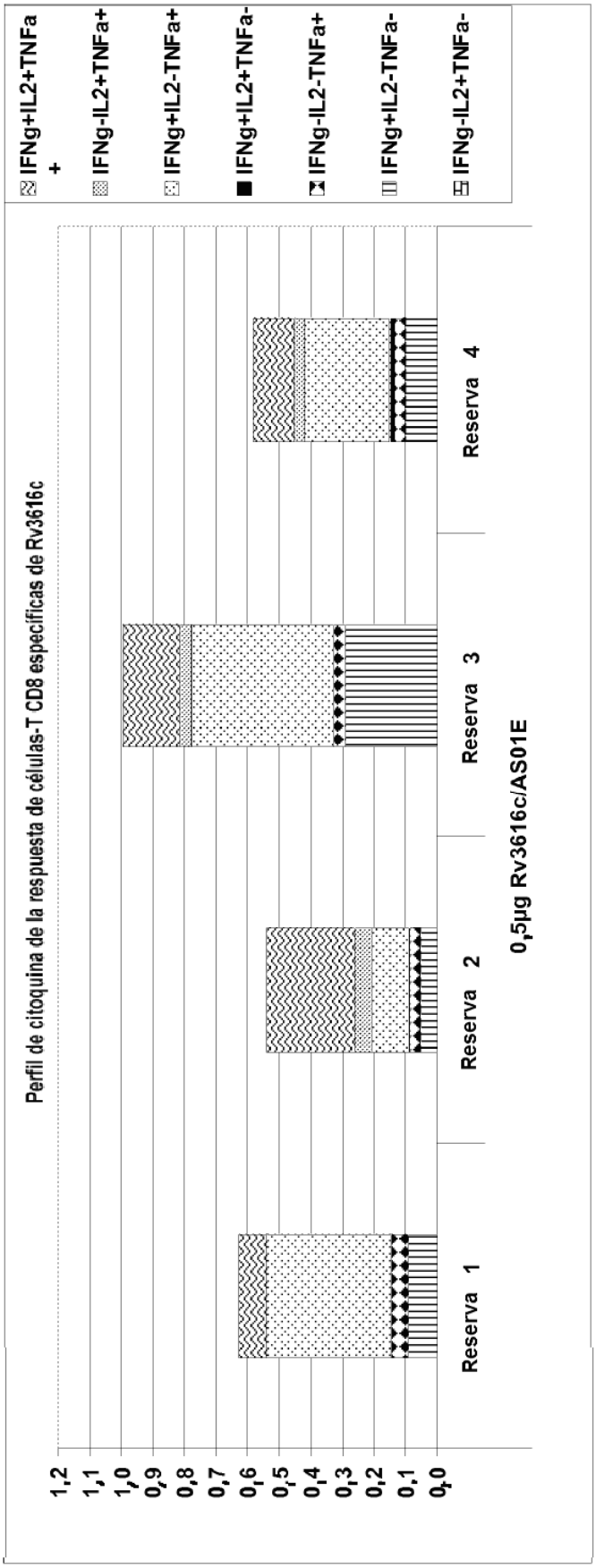
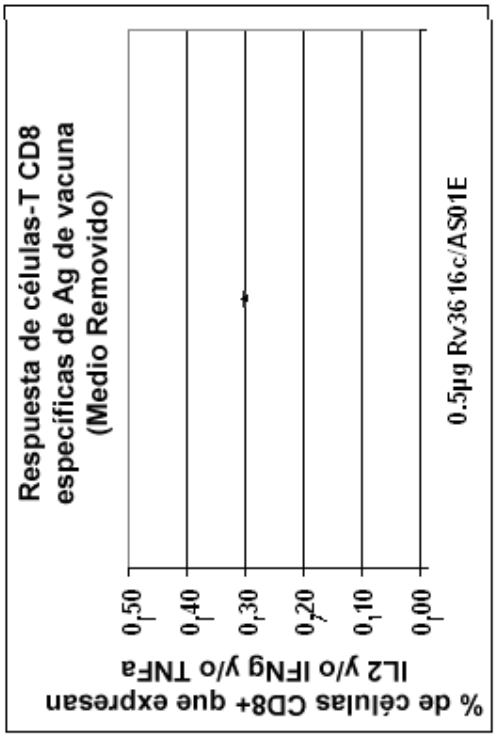

Perfil de citoquina de la respuesta de células-T CD4 especificas de Rv3616c
% de células-T CD4+
0,0
0,1
0,2
0,3
0,4
0,5
IFNg+IL2+TNFa+
IFNg-IL2+TNFa+
IFNg+IL2-TNFa+
IFNg+IL2+TNFa-
IFNg-IL2-TNFa+
IFNg+IL2-TNFa-
IFNg-IL2+TNFa-
Reserva 1
Reserva 2
Reserva 3
Reserva 4
0.5µg Rv3616c/AS01E

Figura 5


Perfil de citoquina de la respuesta de células-T CD8 específicas de Rv3616c
IFNg+IL2+TNFa+
IFNg-IL2+TNFa+
IFNg+IL2-TNFa+
IFNg+IL2+TNFa-
IFNg-IL2-TNFa+
IFNg+IL2-TNFa-
IFNg-IL2+TNFa-
1,2
1,1
1,0
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0,0
Reserva 1
Reserva 2
Reserva 3
Reserva 4
0,5µg Rv3616c/AS01E


Respuesta de células-T CD8 específicas de Ag de vacuna (Medio Removido)
% de células CD8+ que expresan IL2 y/o IFNg y/o TNFa
0,00
0,10
0,20
0,30
0,40
0,50
0.5µg Rv3616c/AS01E

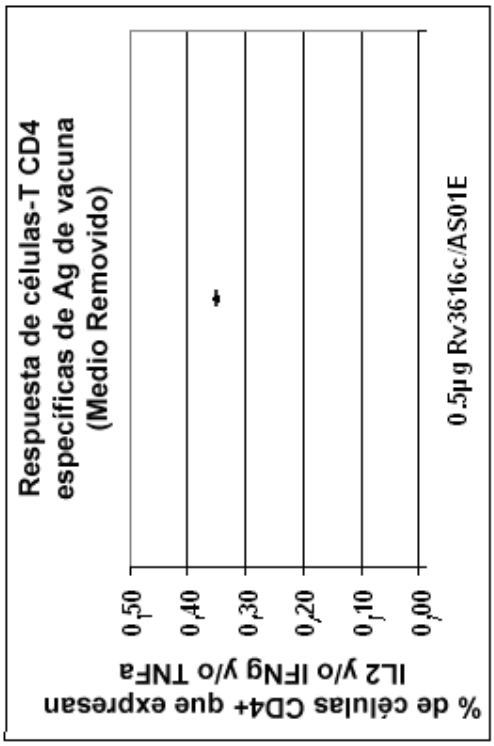

Respuesta de células-T CD4 específicas de Ag de vacuna (Medio Removido)
% de células CD4+ que expresan IL2 y/o IFNg y/o TNFa
0,00
0,10
0,20
0,30
0,40
0,50
0.5µg Rv3616c/AS01E

Figura 6

Figura 7

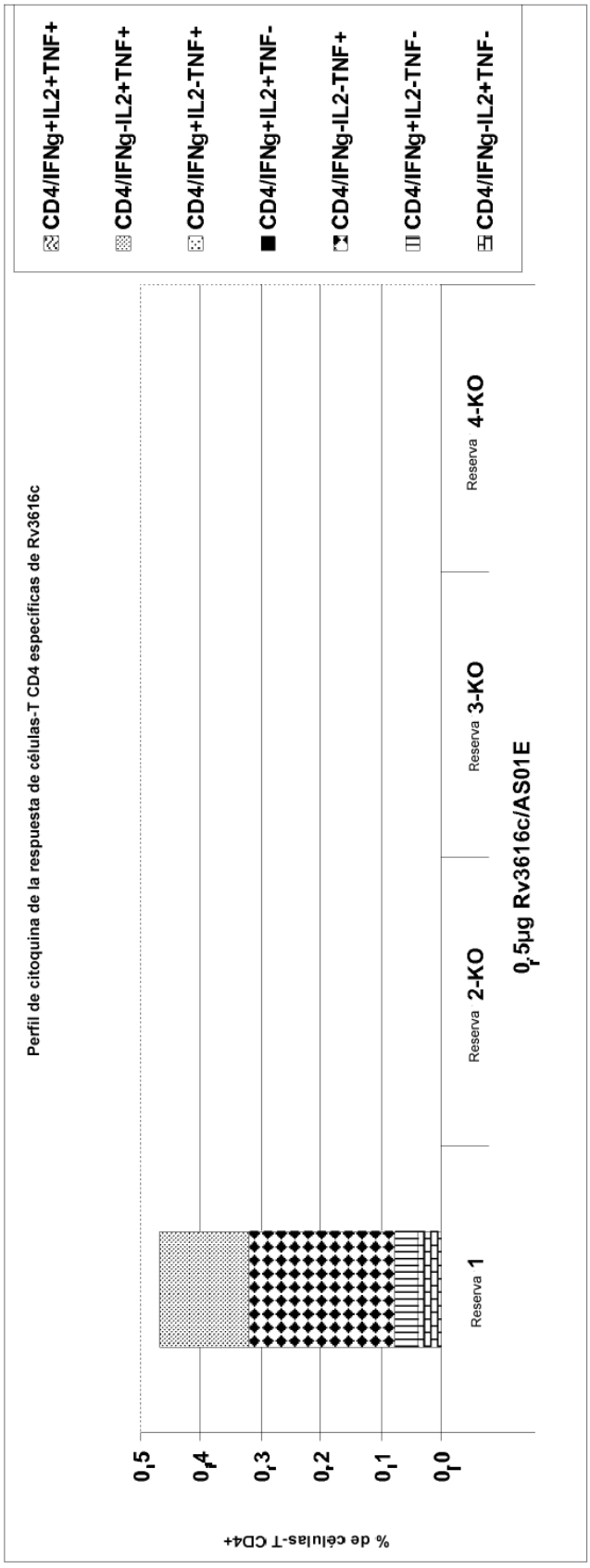

Perfil de citoquina de la respuesta de células-T CD4 específicas de Rv3616c
CD4/IFNg+IL2+TNF+
CD4/IFNg-IL2+TNF+
CD4/IFNg+IL2-TNF+
CD4/IFNg+IL2+TNF-
CD4/IFNg-IL2-TNF+
CD4/IFNg+IL2-TNF-
CD4/IFNg-IL2+TNF-
% de células-T CD4+
0,5
0,4
0,3
0,2
0,1
0,0
Reserva 1
Reserva 2-KO
Reserva 3-KO
Reserva 4-KO
0,5µg Rv3616c/AS01E

Figura 8

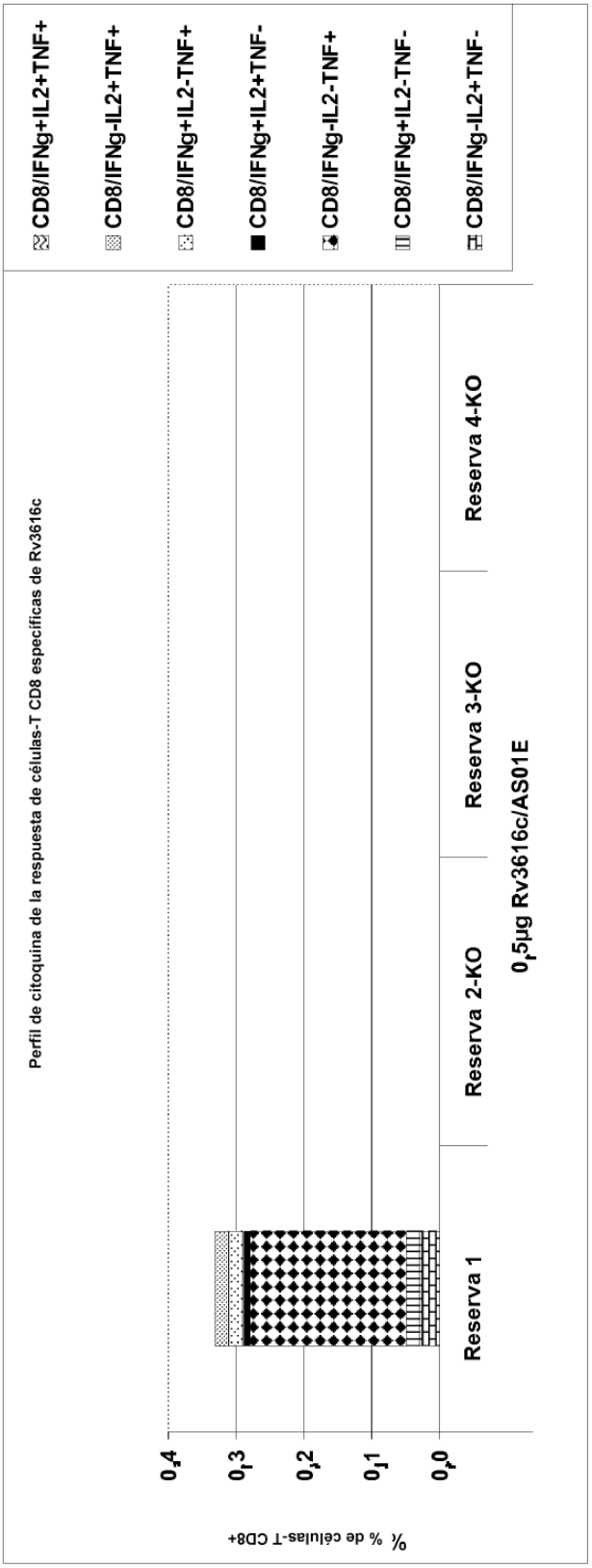
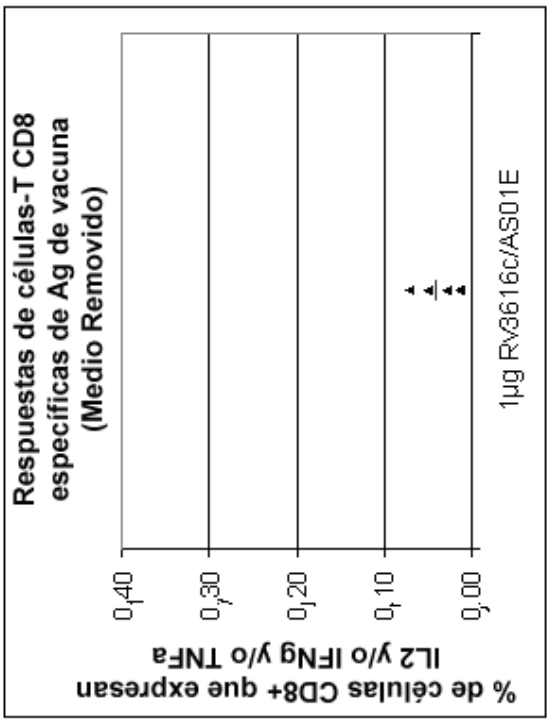

Perfil de citoquina de la respuesta de células-T CD8 específicas de Rv3616c
CD8/IFNg+IL2+TNF+
CD8/IFNg-IL2+TNF+
CD8/IFNg+IL2-TNF+
CD8/IFNg+IL2+TNF-
CD8/IFNg-IL2-TNF+
CD8/IFNg+IL2-TNF-
CD8/IFNg-IL2+TNF-
% de células-T CD8+
0,4
0,3
0,2
0,1
0,0
Reserva 1
Reserva 2-KO
Reserva 3-KO
Reserva 4-KO
0,5µg Rv3616c/AS01E

Figura 9


Respuestas de células-T CD8 específicas de Ag de vacuna (Medio Removido)
% de células CD8+ que expresan IL2 y/o IFNg y/o TNFa
0,40
0,30
0,20
0,10
0,00
1µg Rv3616c/AS01E

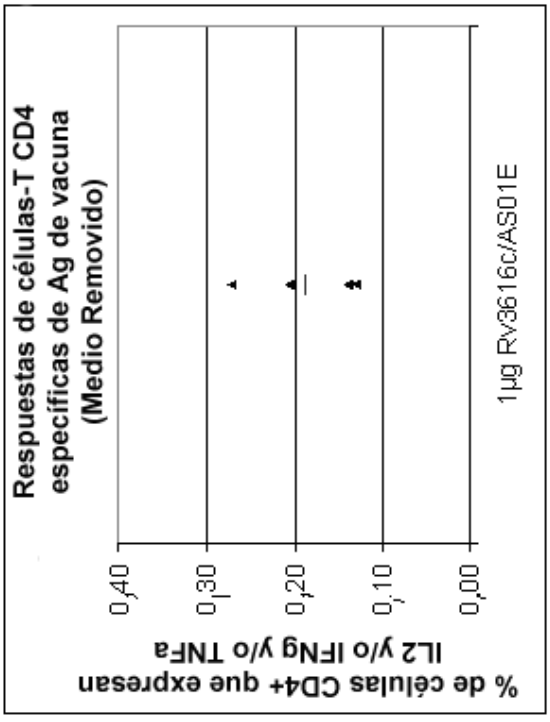

Respuestas de células-T CD4 específicas de Ag de vacuna (Medio Removido)
% de células CD4+ que expresan IL2 y/o IFNg y/o TNFa
0,40
0,30
0,20
0,10
0,00
1µg Rv3616c/AS01E

Figura 10

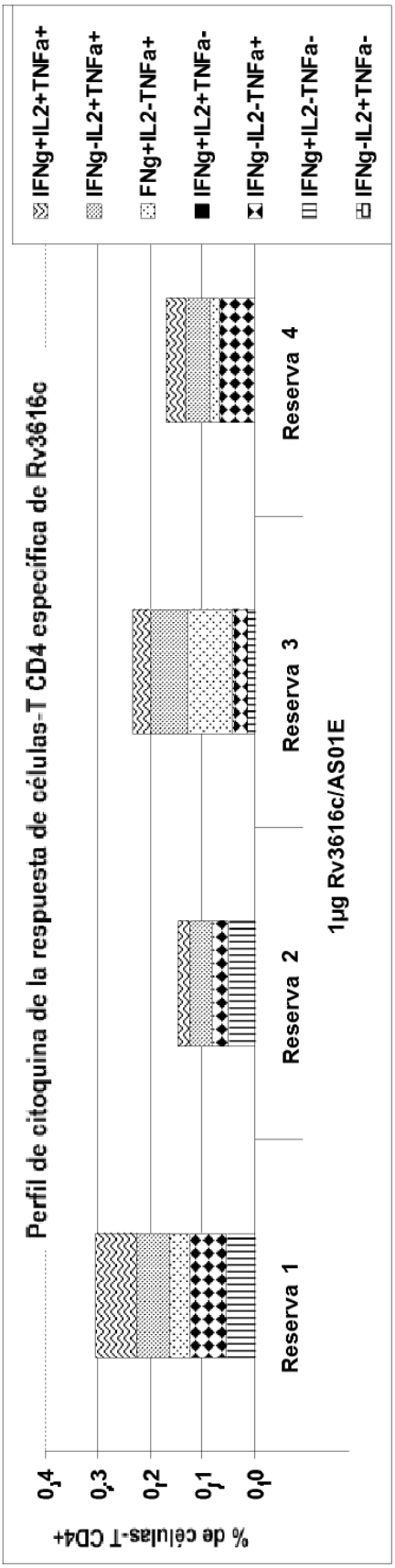
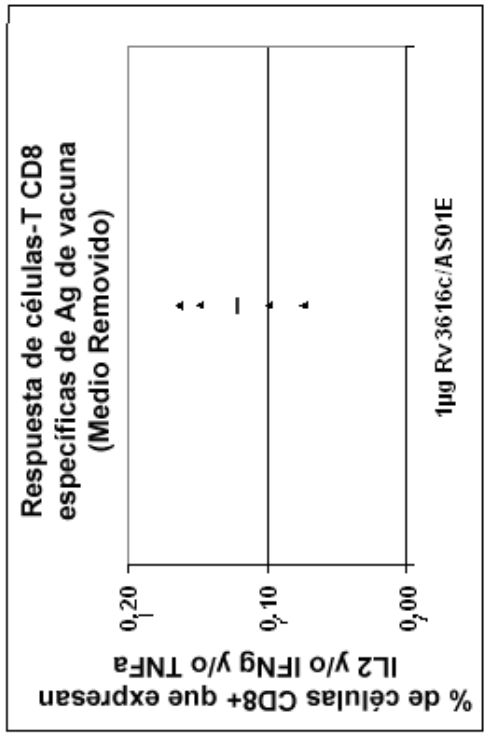

Perfil de citoquina de la respuesta de células-T CD4 especifica de Rv3616c
% de células-T CD4+
0,4
0,3
0,2
0,1
0,0
Reserva 1
Reserva 2
Reserva 3
Reserva 4
1µg Rv3616c/AS01E
IFNg+IL2+TNFa+
IFNg-IL2+TNFa+
FNg+IL2-TNFa+
IFNg+IL2+TNFa-
IFNg-IL2-TNFa+
IFNg+IL2-TNFa-
IFNg-IL2+TNFa-

Figura 11


Respuesta de células-T CD8
específicas de Ag de vacuna
(Medio Removido)
% de células CD8+ que expresan
IL2 y/o IFNg y/o TNFa
0,20
0,10
0,00
1µg Rv3616c/AS01E

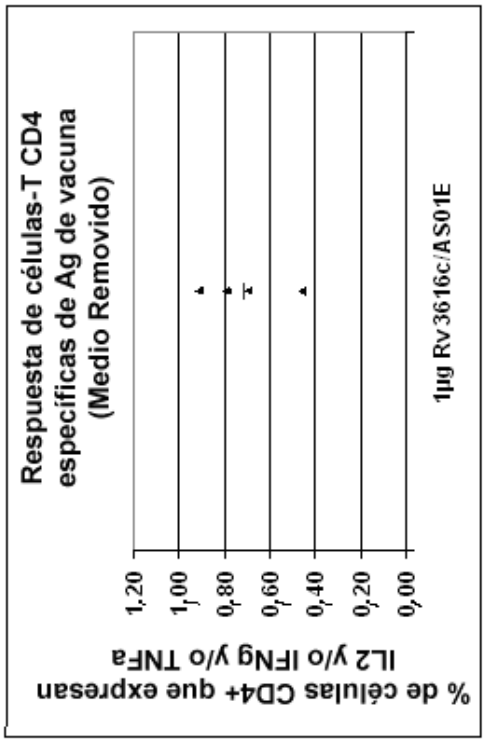

Respuesta de células-T CD4
específicas de Ag de vacuna
(Medio Removido)
% de células CD4+ que expresan
IL2 y/o IFNg y/o TNFa
1,20
1,00
0,80
0,60
0,40
0,20
0,00
1µg Rv3616c/AS01E

Figura 12

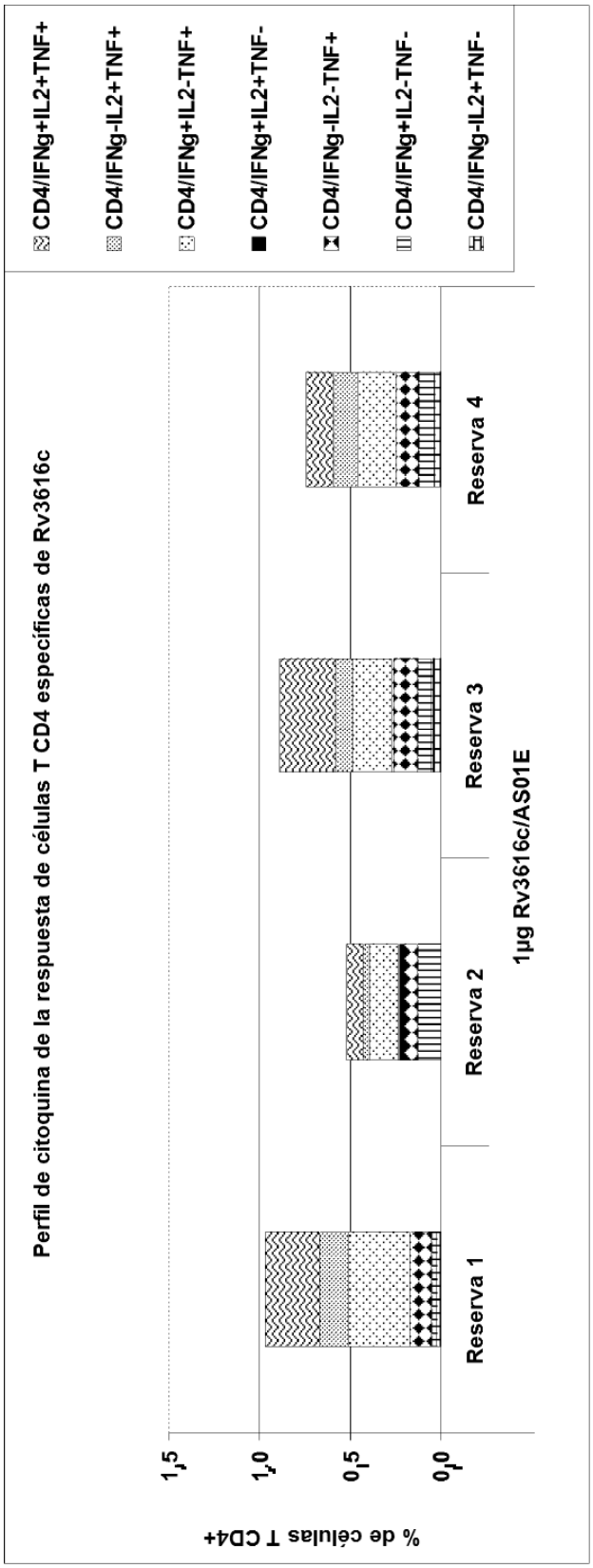

Perfil de citoquina de la respuesta de células T CD4 específicas de Rv3616c
CD4/IFNg+IL2+TNF+
CD4/IFNg-IL2+TNF+
CD4/IFNg+IL2-TNF+
CD4/IFNg+IL2+TNF-
CD4/IFNg-IL2-TNF+
CD4/IFNg+IL2-TNF-
CD4/IFNg-IL2+TNF-
% de células T CD4+
1,5
1,0
0,5
0,0
Reserva 1
Reserva 2
Reserva 3
Reserva 4
1µg Rv3616c/AS01E

Figura 13

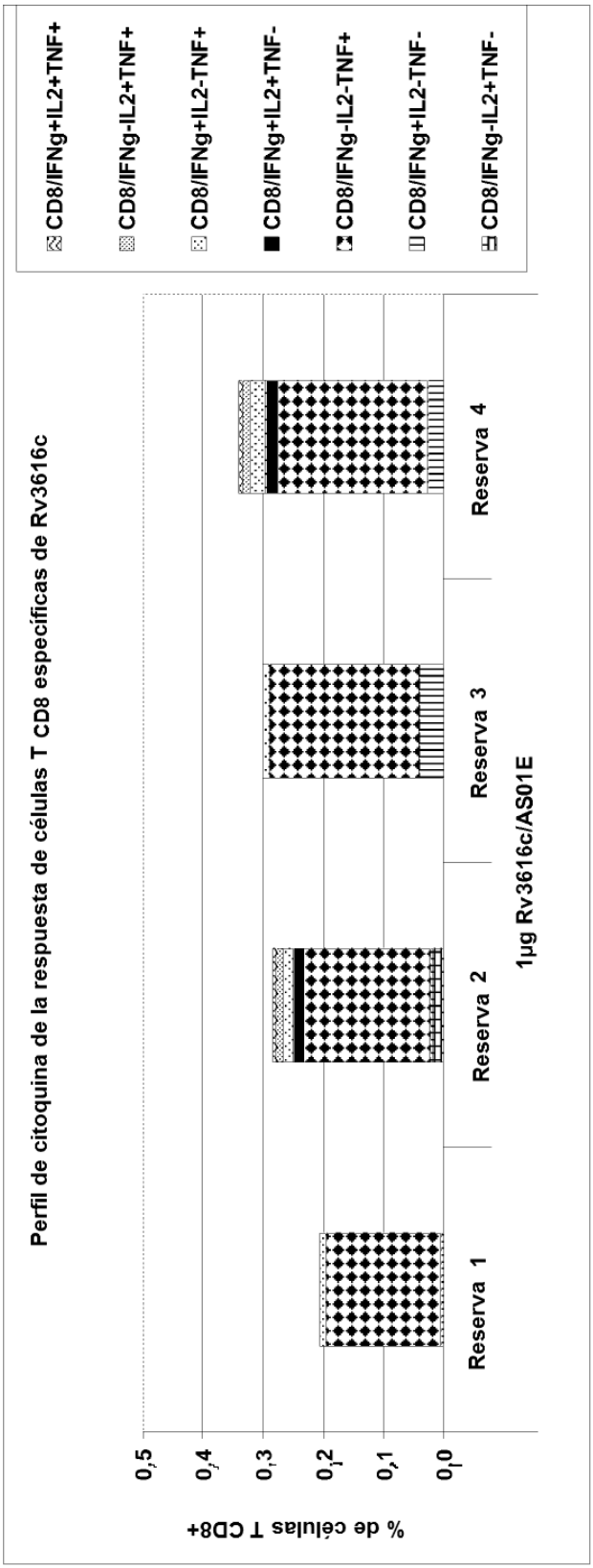

Perfil de citoquina de la respuesta de células T CD8 específicas de Rv3616c
CD8/IFNg+IL2+TNF+
CD8/IFNg-IL2+TNF+
CD8/IFNg+IL2-TNF+
CD8/IFNg+IL2+TNF-
CD8/IFNg-IL2-TNF+
CD8/IFNg+IL2-TNF-
CD8/IFNg-IL2+TNF-
% de células T CD8+
0,5
0,4
0,3
0,2
0,1
0,0
Reserva 1
Reserva 2
Reserva 3
Reserva 4
1µg Rv3616c/AS01E

Figura 14

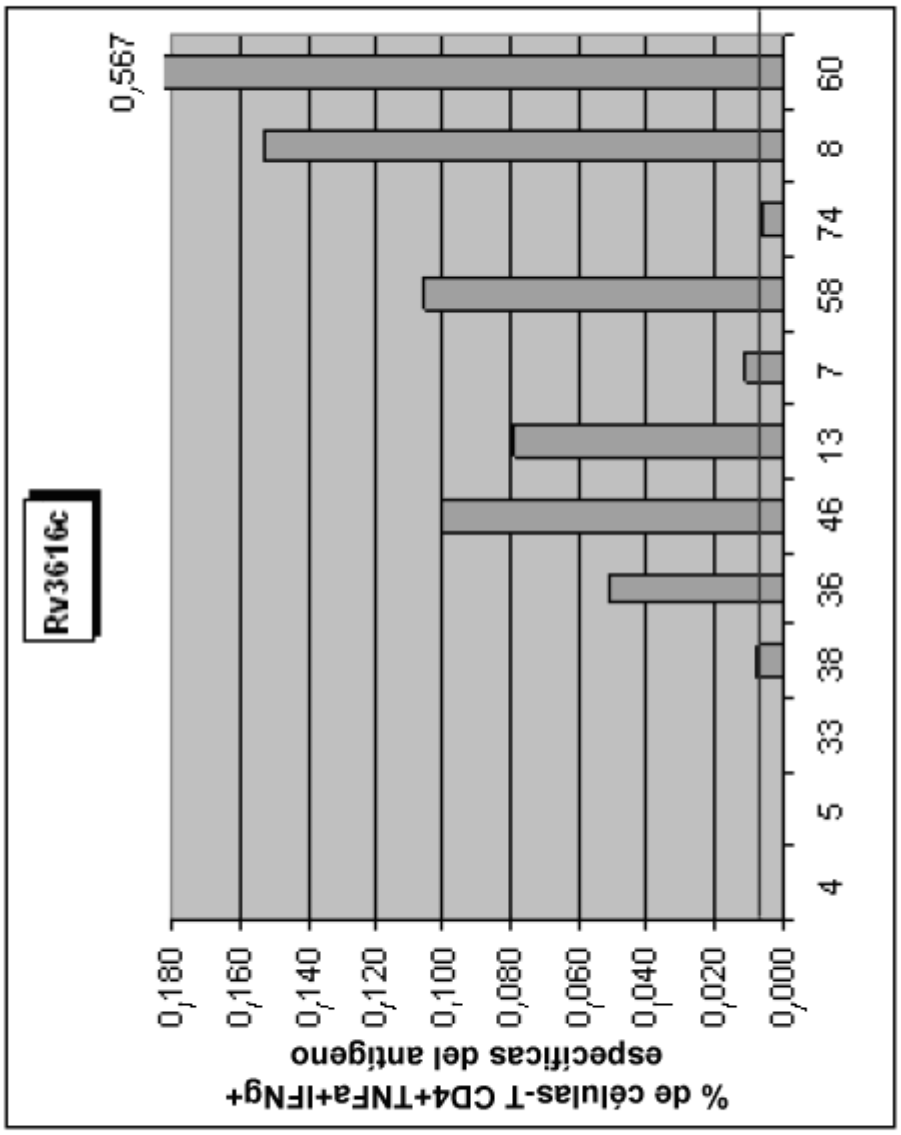


Figura 14

Figura 15

```
H37Rv       1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
CDC1551     1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
F11         1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
Haarlem     1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
Cepa C      1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
BCG         1 MSRVFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA

H37Rv      61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
CDC1551    61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
F11        61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
Haarlem    61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
Cepa C     61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
BCG        61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT

H37Rv     121 YIPVVGHALSAAFQAPFCAGAMAVVGGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
CDC1551   121 YIPVVGHALSAAFQAPFCAGAMAVVGGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
F11       121 YIPVVGHALSAAFQAPFCAGAMAVVGGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
Haarlem   121 YIPVVGHALSAAFQAPFCAGAMAVVGGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
Cepa C    121 YIPVVGHALSAAFQAPFCAGAMAVVGGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
BCG       121 YIPVVGHALSAAFQAPFCAGAMAVVGGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD

H37Rv     181 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
CDC1551   181 IISDVADIIKGILGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
F11       181 IISDVADIIKGILGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
Haarlem   181 IISDVADIIKGILGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
Cepa C    181 IISDVADIIKGILGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
BCG       181 IISDVADIIKGILGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT

H37Rv     241 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
CDC1551   241 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
F11       241 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
Haarlem   241 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
Cepa C    241 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
BCG       241 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA

H37Rv     301 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
CDC1551   301 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
F11       301 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
Haarlem   301 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
Cepa C    301 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
BCG       301 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY

H37Rv     361 SEGAAAGTEDAERAPVEADAGGGQKVLVRNVV
CDC1551   361 SEGAAAGTEDAERAPVEADAGGGQKVLVRNVV
F11       361 SEGAAAGTEDAERAPVEADAGGGQKVLVRNVV
Haarlem   361 SEGAAAGTEDAERAPVEADAGGGQKVLVRNVV
Cepa C    361 SEGAAAGTEDAERAPVEADAGGGQKVLVRNVV
BCG       361 SEGAAAGTEDAERAPVEADAGGGQKVLVRNVV
```

Figura 16A

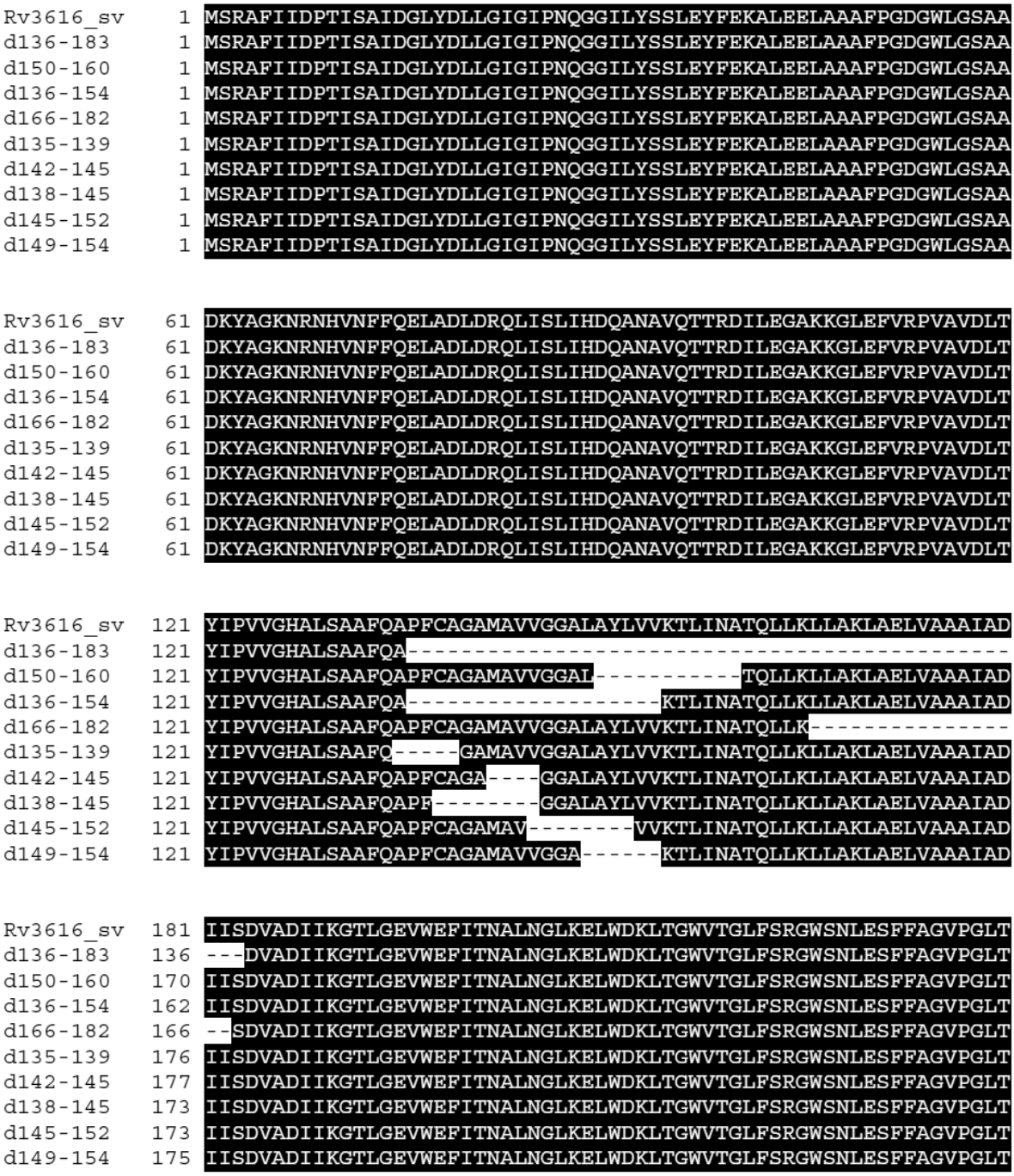
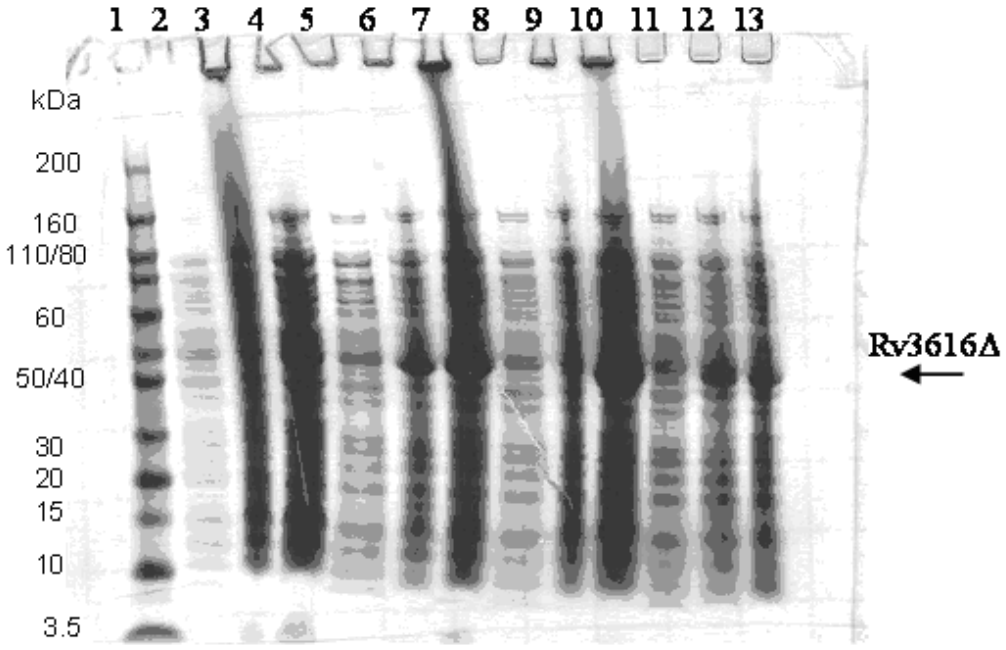

Rv3616_sv 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d136-183 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d150-160 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d136-154 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d166-182 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d135-139 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d142-145 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d138-145 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d145-152 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
d149-154 1 MSRAFIIDPTISAIDGLYDLLGIGIPNQGGILYSSLEYFEKALEELAAAFPGDGWLGSAA
Rv3616_sv 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d136-183 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d150-160 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d136-154 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d166-182 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d135-139 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d142-145 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d138-145 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d145-152 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
d149-154 61 DKYAGKNRNHVNFFQELADLDRQLISLIHDQANAVQTTRDILEGAKKGLEFVRPVAVDLT
Rv3616_sv 121 YIPVVGHALSAAFQAPFCAGAMAVVGGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
d136-183 121 YIPVVGHALSAAFQA----------
d150-160 121 YIPVVGHALSAAFQAPFCAGAMAVVGGAI----------TQLLKLLAKLAELVAAAIAD
d136-154 121 YIPVVGHALSAAFQA----------KTLINATQLLKLLAKLAELVAAAIAD
d166-182 121 YIPVVGHALSAAFQAPFCAGAMAVVGGALAYLVVKTLINATQLLK----------
d135-139 121 YIPVVGHALSAAFQ-----GAMAVVGGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
d142-145 121 YIPVVGHALSAAFQAPFCAGA----GGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
d138-145 121 YIPVVGHALSAAFQAPF--------GGALAYLVVKTLINATQLLKLLAKLAELVAAAIAD
d145-152 121 YIPVVGHALSAAFQAPFCAGAMAV--------VVKTLINATQLLKLLAKLAELVAAAIAD
d149-154 121 YIPVVGHALSAAFQAPFCAGAMAVVGGA------KTLINATQLLKLLAKLAELVAAAIAD
Rv3616_sv 181 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d136-183 136 ---DVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d150-160 170 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d136-154 162 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d166-182 166 --SDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d135-139 176 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d142-145 177 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d138-145 173 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d145-152 173 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT
d149-154 175 IISDVADIIKGTLGEVWEFITNALNGLKELWDKLTGWVTGLFSRGWSNLESFFAGVPGLT

Figura 16B

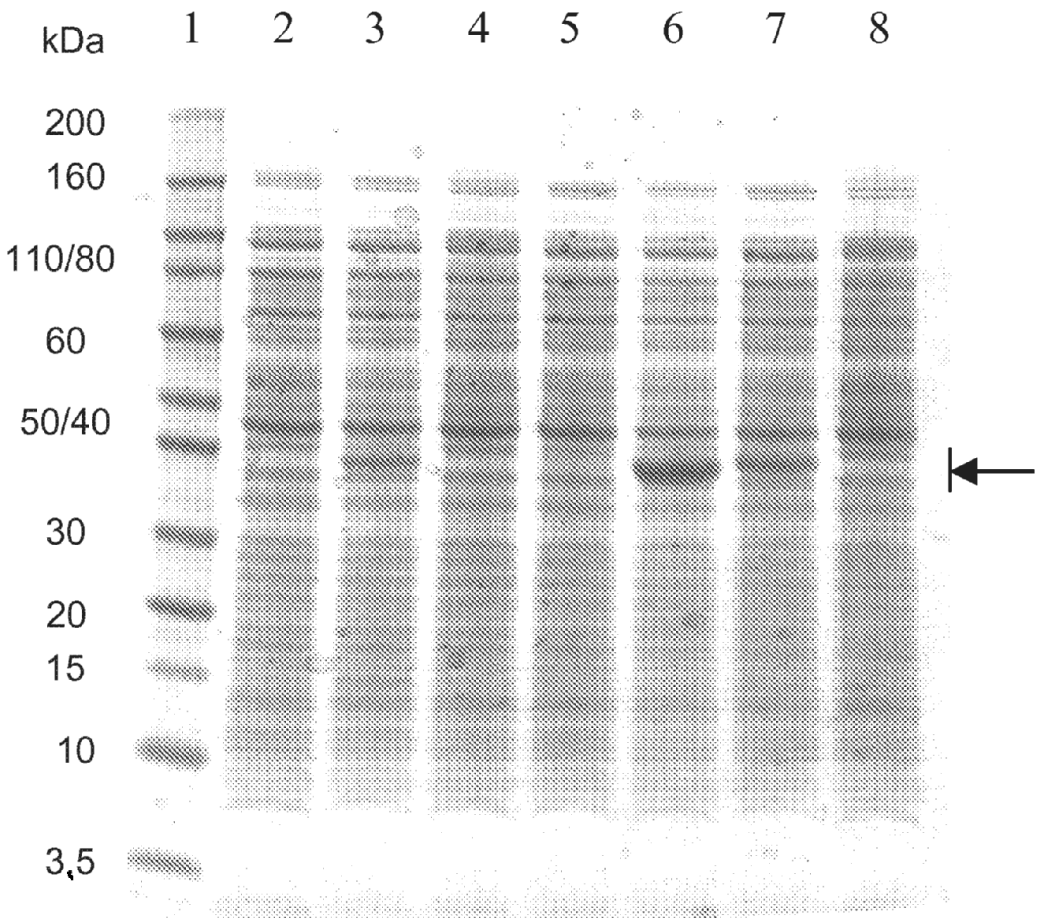
```
Rv3616_sv   241 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d136-183    193 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d150-160    230 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d136-154    222 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d166-182    224 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d135-139    236 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d142-145    237 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d138-145    233 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d145-152    233 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA
d149-154    235 GATSGLSQVTGLFGAAGLSASSGLAHADSLASSASLPALAGIGGGSGFGGLPSLAQVHAA

Rv3616_sv   301 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d136-183    253 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d150-160    290 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d136-154    282 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d166-182    284 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d135-139    296 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d142-145    297 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d138-145    293 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d145-152    293 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
d149-154    295 STRQALRPRADGPVGAAAEQVGGQSQLVSAQGSQGMGGPVGMGGMHPSSGASKGTTTKKY
```

Figura 17



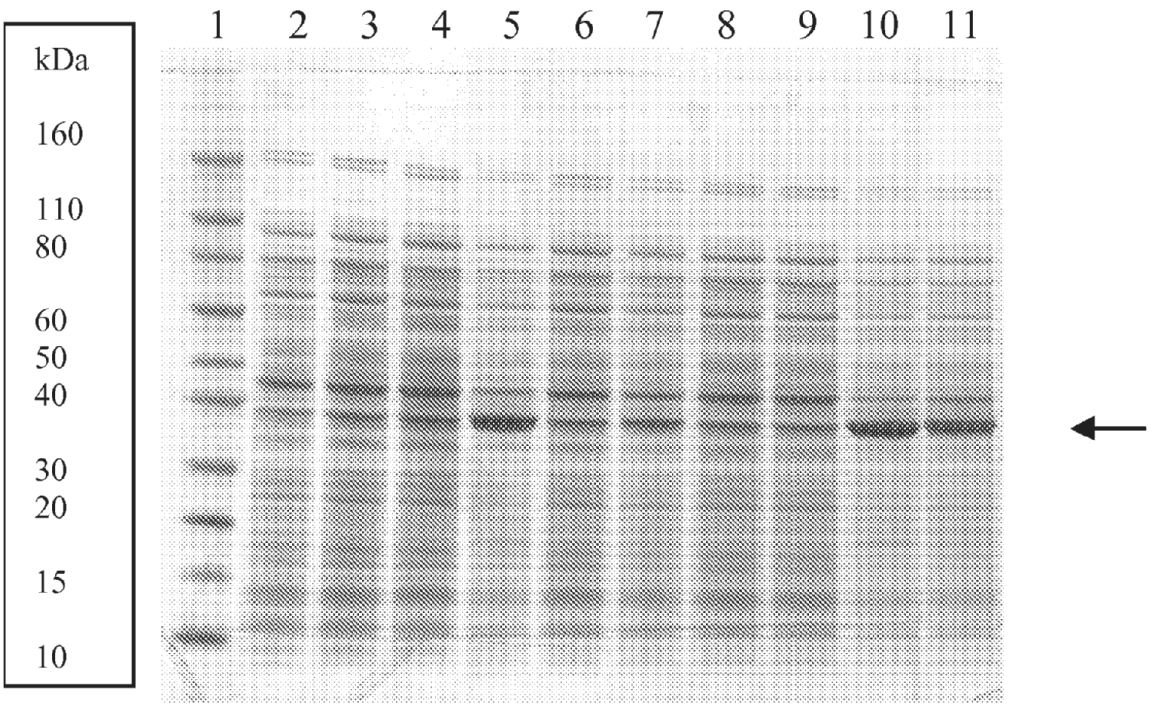
1- Estándar de proteína Novex
2- Rv3616Δ136-183 Antes de inducción (BI)
3- Rv3616Δ136-183 Después de inducción (AI) 37°C
4- Rv3616Δ136-183 Después de inducción (AI) 16°C
5- Rv3616Δ150-160 BI
6- Rv3616Δ150-160 AI 37°C
7- Rv3616Δ150-160 AI 16°C
8- Rv3616Δ136-154 BI
9- Rv3616Δ136-154 AI 37°C
10- Rv3616Δ136-154 AI 16°C
11- Rv3616Δ166-182 BI
12- Rv3616Δ166-182 AI 37°C
13- Rv3616Δ166-182 AI 16°C

Figura 18



1- **Estándar de proteína Novex**
2- **Rv3616 Antes de inducción (BI)**
3- **Rv3616Δ135-139 Después de inducción (AI) 37°C**
4- **Rv3616Δ142-145 AI 37°C**
5- **Rv3616Δ145-152 AI 37°C**
6- **Rv3616Δ138-145 AI 37°C**
7- **Rv3616Δ149-154 AI 37°C**
8- **Rv3616 AI 37°C**

Figura 19



1- Estándar de proteína Novex
2- Rv3616 Antes de inducción (BI)
3- Rv3616 Después de inducción (AI)
4- Rv3616Δ150-160 AI
5- Rv3616Δ136-154 AI
6- Rv3616Δ166-182 AI
7- Rv3616Δ135-139 AI
8- Rv3616Δ142-145 AI
9- Rv3616Δ145-152 AI
10- Rv3616Δ138-145 AI
11- Rv3616Δ149-154 AI

Figura 20

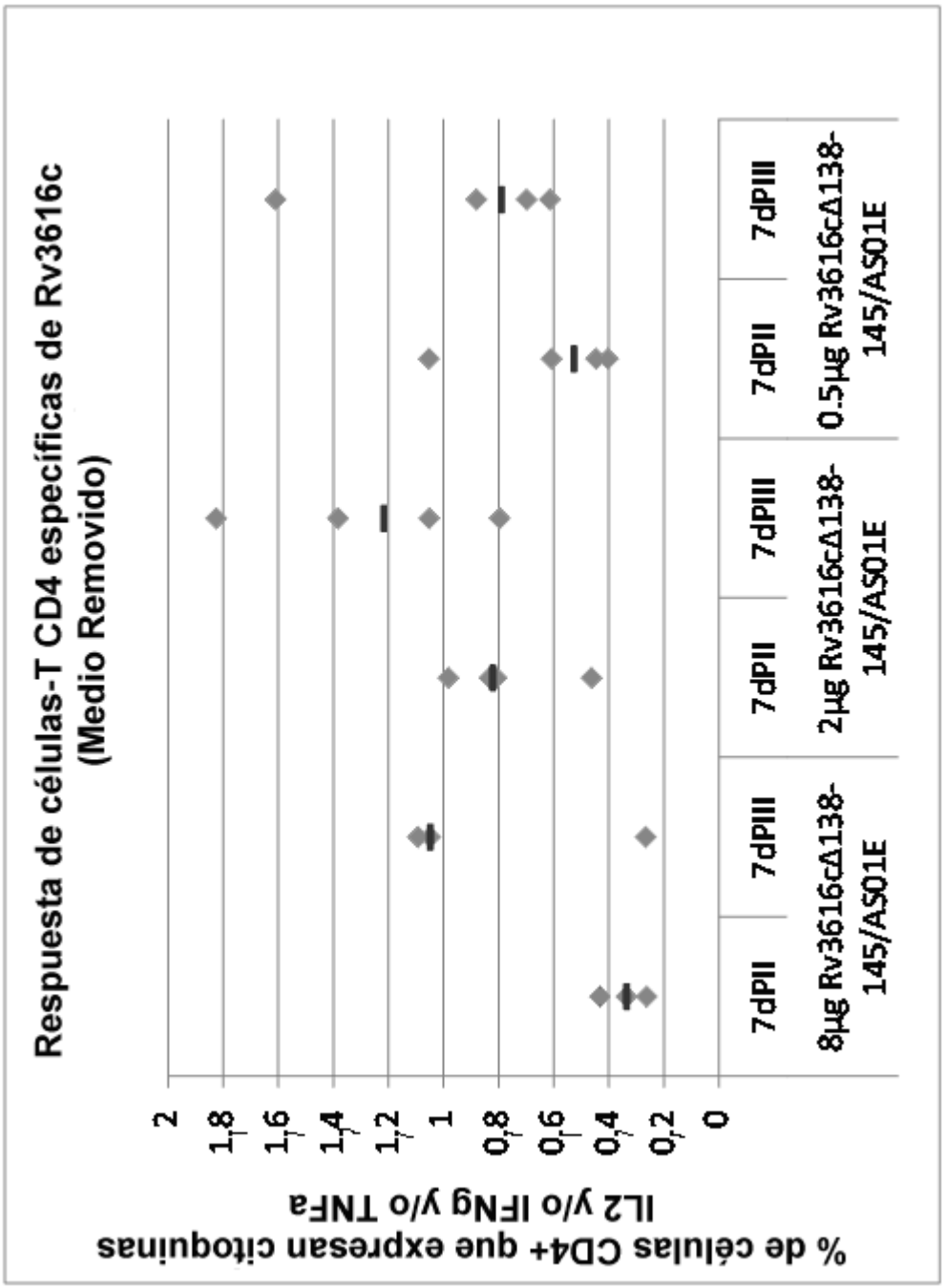

Respuesta de células-T CD4 específicas de Rv3616c
(Medio Removido)
% de células CD4+ que expresan citoquinas
IL2 y/o IFNg y/o TNFa
2
1,8
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0
7dPII
7dPIII
8µg Rv3616cΔ138-145/AS01E
7dPII
7dPIII
2µg Rv3616cΔ138-145/AS01E
7dPII
7dPIII
0.5µg Rv3616cΔ138-145/AS01E

Figura 21

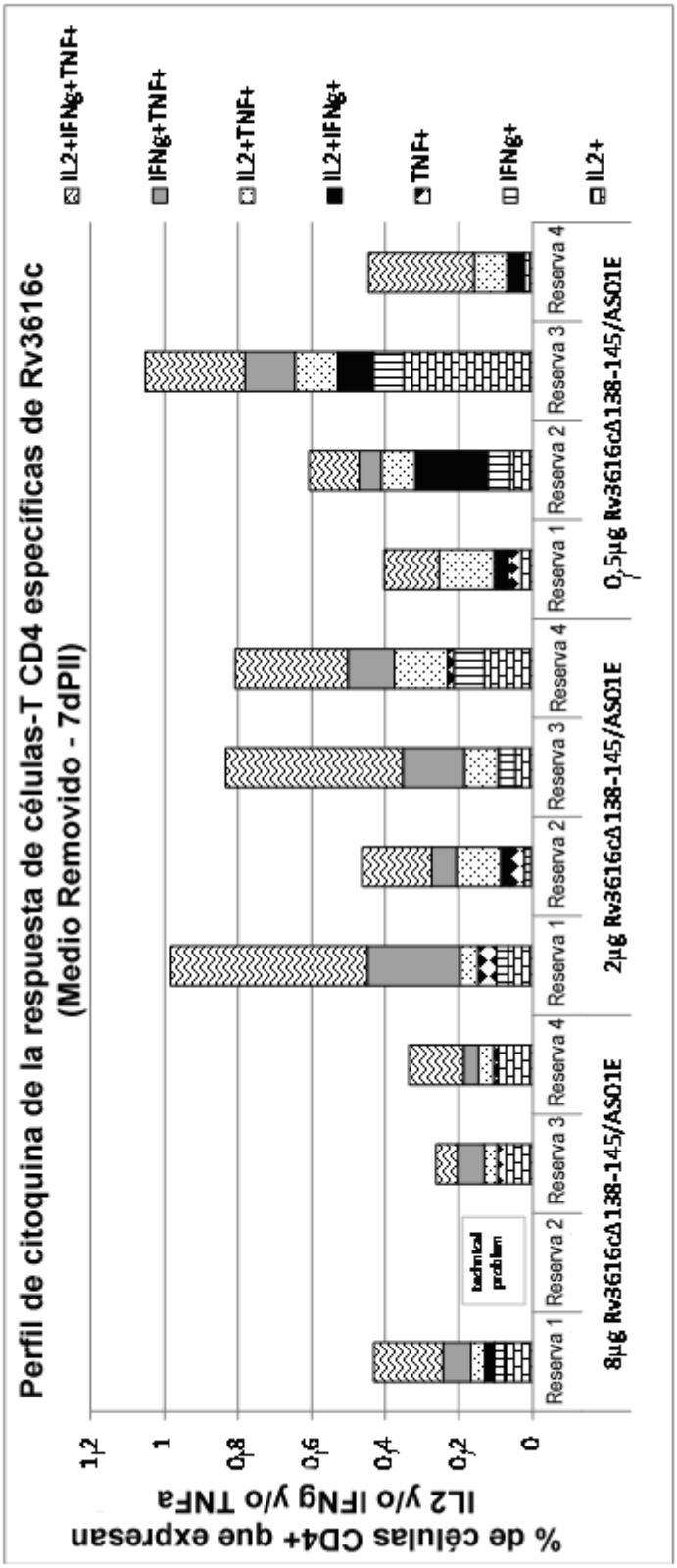

Perfil de citoquina de la respuesta de células-T CD4 específicas de Rv3616c
(Medio Removido - 7dPII)
% de células CD4+ que expresan IL2 y/o IFNg y/o TNFa
1,2
1
0,8
0,6
0,4
0,2
0
IL2+IFNg+TNF+
IFNg+TNF+
IL2+TNF+
IL2+IFNg+
TNF+
IFNg+
IL2+
Reserva 1
Reserva 2
Reserva 3
Reserva 4
technical problem
8µg Rv3616cΔ138-145/AS01E
2µg Rv3616cΔ138-145/AS01E
0,5µg Rv3616cΔ138-145/AS01E

Figura 22

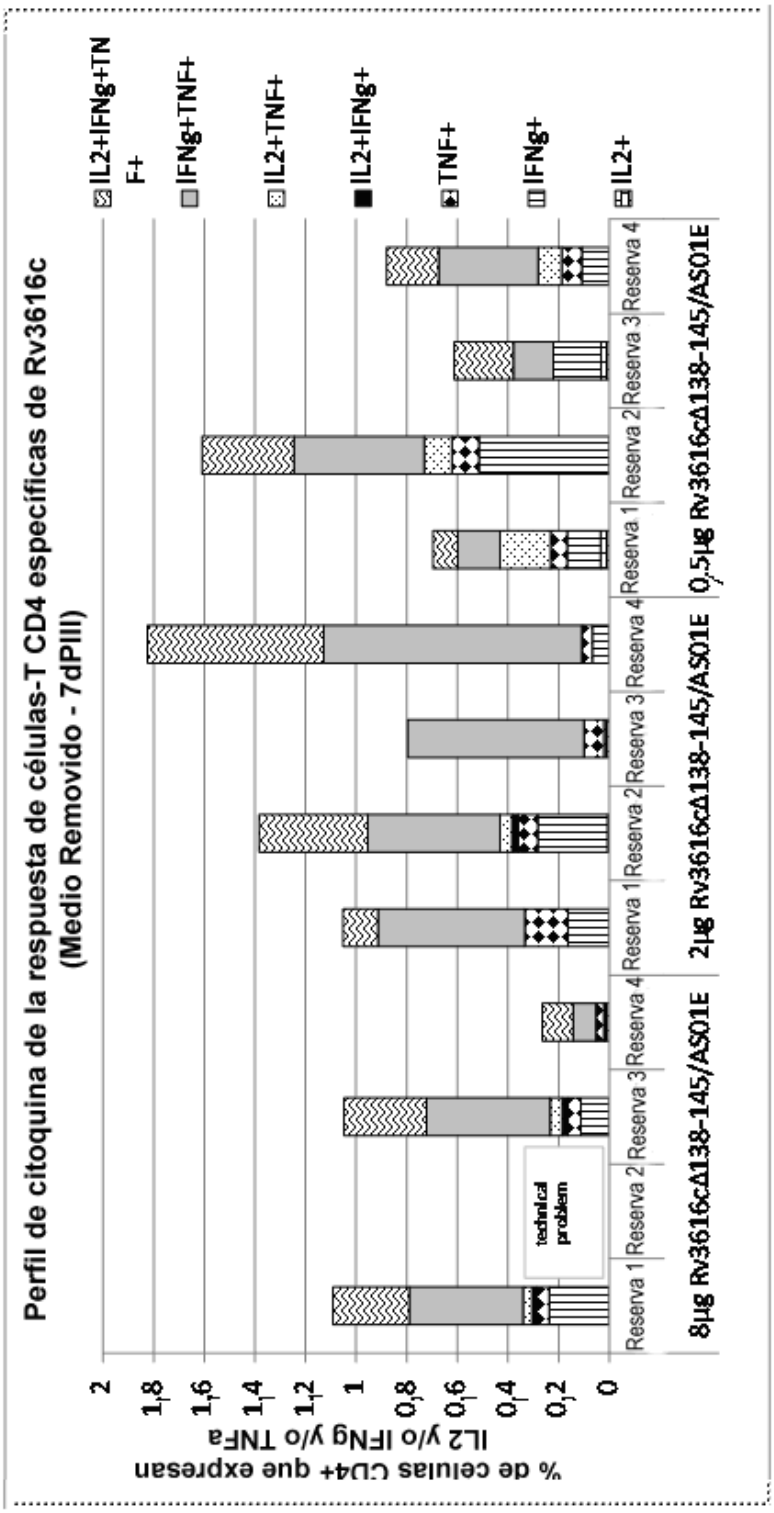

Perfil de citoquina de la respuesta de células-T CD4 específicas de Rv3616c
(Medio Removido - 7dPIII)
% de celulas CD4+ que expresan IL2 y/o IFNg y/o TNFa
IL2+IFNg+TNF+
IFNg+TNF+
IL2+TNF+
IL2+IFNg+
TNF+
IFNg+
IL2+
technical problem
Reserva 1
Reserva 2
Reserva 3
Reserva 4
8µg Rv3616cΔ138-145/AS01E
2µg Rv3616cΔ138-145/AS01E
0,5µg Rv3616cΔ138-145/AS01E

Figura 23

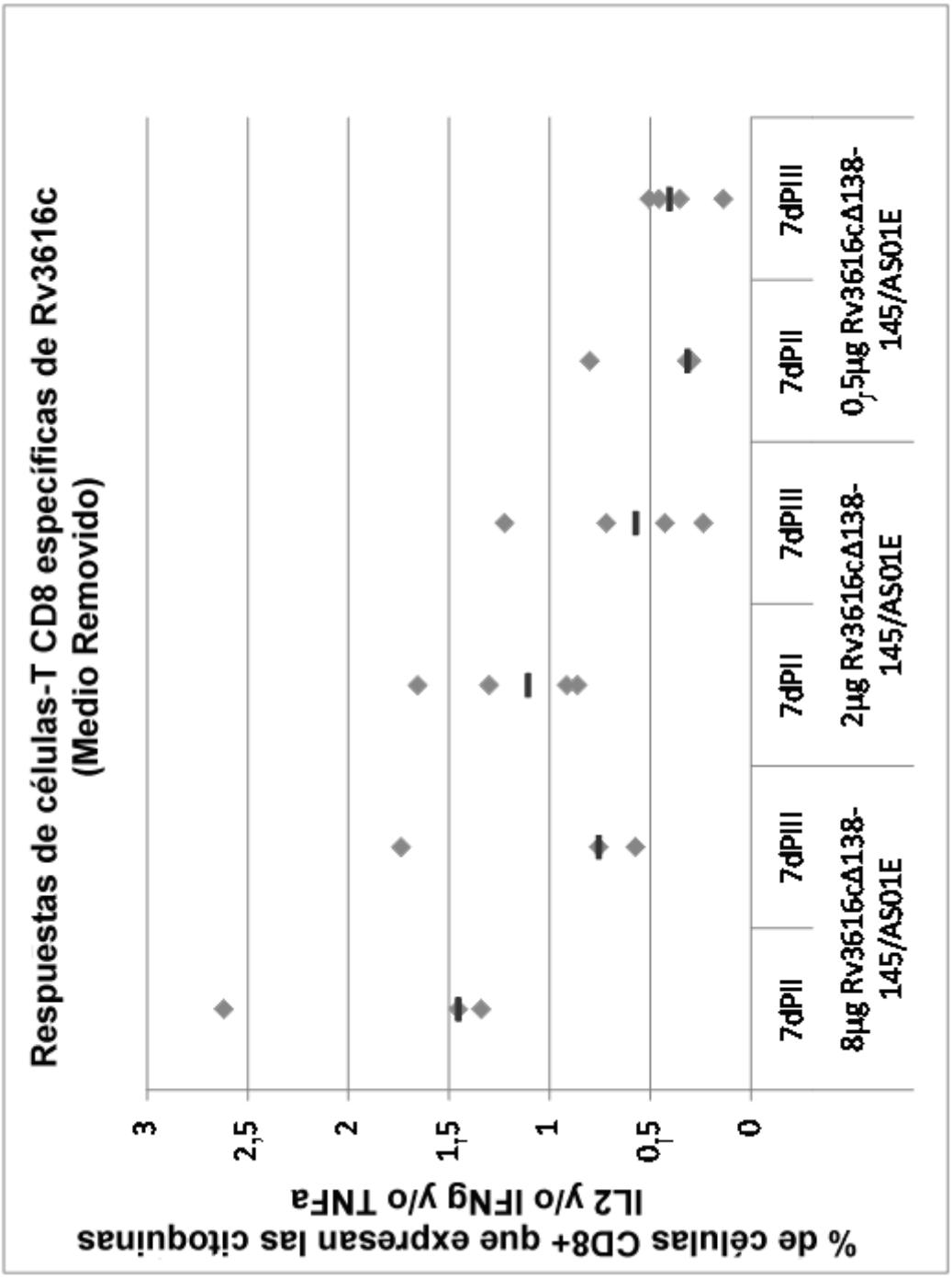

Respuestas de células-T CD8 específicas de Rv3616c
(Medio Removido)
% de células CD8+ que expresan las citoquinas
IL2 y/o IFNg y/o TNFa
3
2,5
2
1,5
1
0,5
0
7dPII
7dPIII
8µg Rv3616cΔ138-145/AS01E
7dPII
7dPIII
2µg Rv3616cΔ138-145/AS01E
7dPII
7dPIII
0,5µg Rv3616cΔ138-145/AS01E

Figura 24

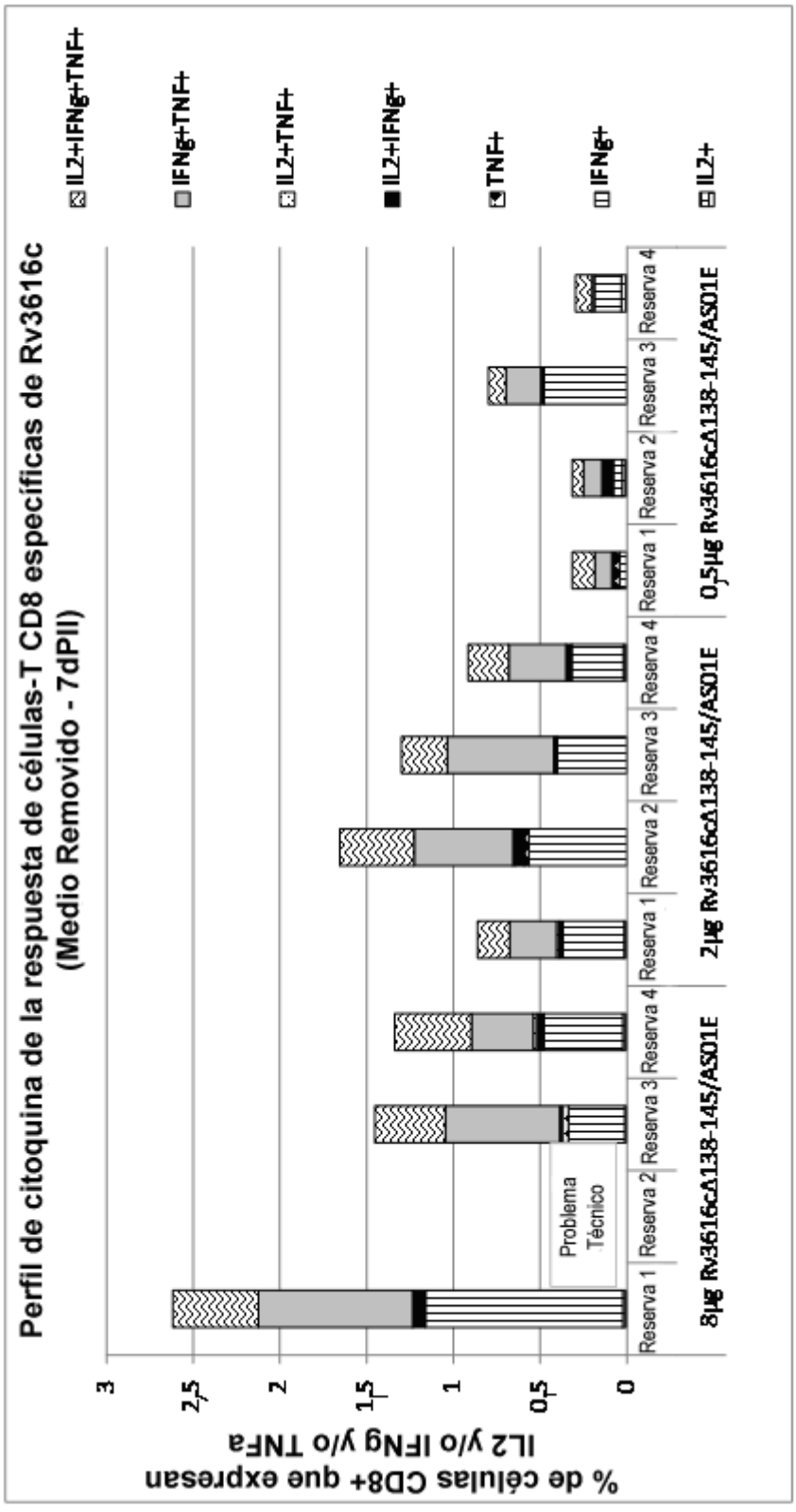
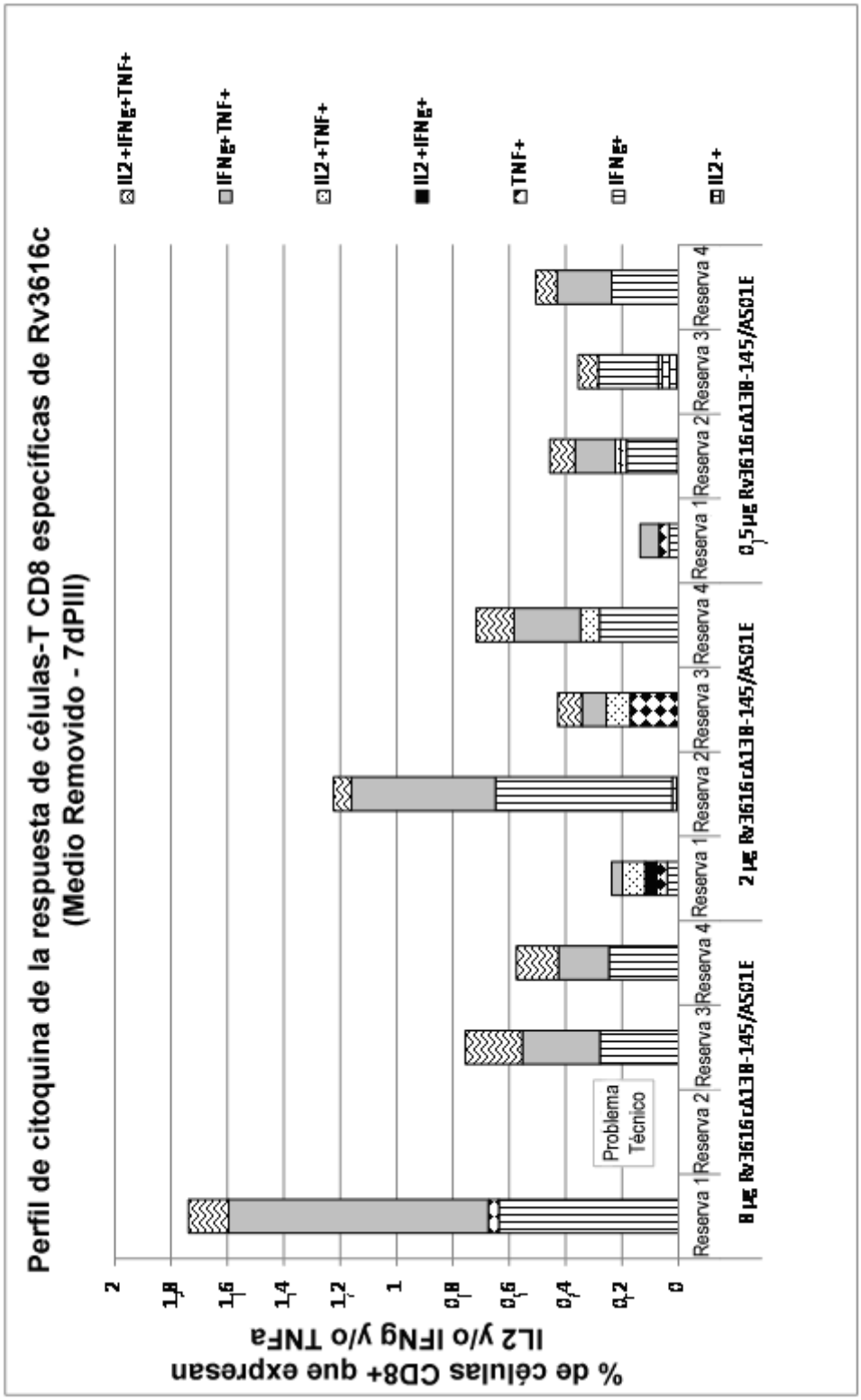

Perfil de citoquina de la respuesta de células-T CD8 específicas de Rv3616c
(Medio Removido - 7dPII)
% de células CD8+ que expresan
IL2 y/o IFNg y/o TNFa
3
2,5
2
1,5
1
0,5
0
IL2+IFNg+TNF+
IFNg+TNF+
IL2+TNF+
IL2+IFNg+
TNF+
IFNg+
IL2+
Problema Técnico
Reserva 1
Reserva 2
Reserva 3
Reserva 4
8µg Rv3616cΔ138-145/AS01E
2µg Rv3616cΔ138-145/AS01E
0,5µg Rv3616cΔ138-145/AS01E

Figura 25


Perfil de citoquina de la respuesta de células-T CD8 específicas de Rv3616c
(Medio Removido - 7dPIII)
% de células CD8+ que expresan IL2 y/o IFNg y/o TNFa
IL2+IFNg+TNF+
IFNg+TNF+
IL2+TNF+
IL2+IFNg+
TNF+
IFNg+
IL2+
2
1,8
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0
Problema Técnico
Reserva 1
Reserva 2
Reserva 3
Reserva 4
8 µg Rv3616cΔ138-145/AS01E
2 µg Rv3616cΔ138-145/AS01E
0,5µg Rv3616cΔ138-145/AS01E