

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 647 321**

51 Int. Cl.:

A61K 39/04 (2006.01)

C07K 14/35 (2006.01)

G01N 33/569 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.07.2009 PCT/EP2009/059586**

87 Fecha y número de publicación internacional: **28.01.2010 WO10010180**

96 Fecha de presentación y número de la solicitud europea: **24.07.2009 E 09781058 (4)**

97 Fecha y número de publicación de la concesión europea: **23.08.2017 EP 2315597**

54 Título: **Composiciones y procedimientos novedosos**

30 Prioridad:

25.07.2008 US 83692 P

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

20.12.2017

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:

**BROWN, JAMES;
METTENS, PASCAL y
MURPHY, DENNIS**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 647 321 T3

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

DESCRIPCIÓN

Composiciones y procedimientos novedosos

Campo de la invención

5 La presente invención también se refiere a polipéptidos y polinucleótidos para su uso en el tratamiento o la prevención de la tuberculosis, en particular para su uso en el tratamiento o prevención de la tuberculosis latente y en la prevención o retraso de la reactivación de la tuberculosis (y también a los procedimientos relacionados). La presente invención se refiere además a composiciones farmacéuticas e inmunogénicas que comprenden dichos polipéptidos y polinucleótidos.

Antecedentes de la invención

10 La tuberculosis (TB) es una enfermedad infecciosa crónica causada por la infección con *Mycobacterium tuberculosis* y otras especies de *Mycobacterium*. Es una enfermedad importante en los países en desarrollo, así como un problema creciente en áreas desarrolladas del mundo. Se cree que hay más de 2 mil millones de personas infectadas con el bacilo de la tuberculosis, con aproximadamente 9,2 millones de nuevos casos de TB y 1,7 millones de muertes cada año. El 10 % de los infectados por el bacilo de la tuberculosis desarrollará tuberculosis activa, cada
15 persona con TB activa infectará a un promedio de 10 a 15 otras personas al año. Mientras que las tasas de incidencia anuales han alcanzado su máximo nivel mundial, el número de muertes y casos sigue aumentando debido al crecimiento de la población (Organización Mundial de la Salud, Datos sobre tuberculosis 2008).

20 *Mycobacterium tuberculosis* infecta a las personas a través de las vías respiratorias. Los macrófagos alveolares engullen la bacteria, pero esta es capaz de sobrevivir y proliferar inhibiendo la fusión del fagosoma con los lisosomas ácidos. Se produce una respuesta inmunitaria compleja que implica células T CD4+ y CD8+, que, en última instancia, da como resultado la formación de un granuloma. Fundamental para el éxito de *Mycobacterium tuberculosis* como patógeno es el hecho de que la bacteria aislada, pero no erradicada, puede persistir durante largos períodos de tiempo, dejando a un individuo vulnerable al desarrollo posterior de la TB activa.

25 Menos del 5 % de los individuos infectados desarrolla la TB activa en los primeros años después de la infección. El granuloma puede persistir durante décadas y se cree que contiene *Mycobacterium tuberculosis* viva en un estado de latencia, privado de oxígeno y nutrientes. No obstante, recientemente se ha sugerido que la mayoría de las bacterias en estado de latencia se encuentra en tipos de células no macrófagos repartidas por todo el cuerpo (Locht y col., Expert Opin. Biol. Ther. 2007, 7(11):1665-1677). El desarrollo de TB activa se produce cuando el equilibrio entre la inmunidad natural del huésped y el patógeno cambia por ejemplo como resultado de un evento inmunosupresor
30 (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 TB latente y la TB activa (Cardana P-J Inflammation y Allergy - Drug Targets 2006 6:27-39; Cardana P-J Infection 2009 37(2):80-86).

35 Aunque una infección puede ser asintomática durante un período considerable de tiempo, la enfermedad activa se manifiesta más habitualmente como inflamación aguda de los pulmones, lo que da lugar a cansancio, pérdida de peso, fiebre y una tos persistente. Si no se trata, normalmente se producen complicaciones graves y muerte.

40 La tuberculosis en general se puede controlar mediante tratamiento antibiótico prolongado, aunque dicho tratamiento no es suficiente para prevenir la propagación de la enfermedad. Los individuos infectados pueden ser asintomáticos, pero contagiosos, durante un tiempo. Además, aunque el cumplimiento del régimen de tratamiento es crucial, el comportamiento del paciente es difícil de controlar. Algunos pacientes no completan el curso del tratamiento, lo que puede conducir a un tratamiento ineficaz y al desarrollo de resistencia a los medicamentos.

45 La TB resistente a múltiples fármacos (MDR-TB) es una forma que no responde a los medicamentos de primera línea. El 5 % de todos los casos de TB son MDR-TB, con una estimación de 490.000 nuevos casos de MDR-TB anuales. Se produce TB extensamente resistente a fármacos cuando se desarrolla resistencia a los medicamentos de segunda línea además de la MDR-TB. Se estima que anualmente se producen 40.000 nuevos casos de TB-XDR prácticamente intratable (Organización Mundial de la Salud, Datos sobre tuberculosis 2008).

Incluso si se ha completado un ciclo completo de tratamiento con antibióticos, la infección con M. tuberculosis puede no quedar erradicada de la persona infectada y puede permanecer como infección latente que puede reactivarse.

Con el fin de controlar la propagación de la tuberculosis, la vacunación eficaz y un diagnóstico precoz preciso de la enfermedad son de suma importancia.

50 El diagnóstico de la infección latente de TB normalmente se establece mediante la prueba de la tuberculina, que implica la exposición intradérmica al derivado proteico purificado de tuberculina (PPD). Las respuestas de células T específicas de antígeno dan como resultado una induración mensurable en el sitio de la inyección durante 48-72 horas después de la inyección, lo que indica la exposición a antígenos micobacterianos. Sin embargo, la sensibilidad y la especificidad han sido un problema con esta prueba y no siempre se puede distinguir fácilmente a los individuos

vacunados con BCG de los individuos infectados (esto es particularmente importante a la luz del hecho de que la BCG no protege contra la infección latente). En general, los individuos que han recibido BCG pero no están infectados por *M. tuberculosis* muestran una reacción de PPD por debajo de 10 mm de diámetro, mientras que las personas que tienen una reacción de PPD por encima de 10 mm de diámetro se considera que están infectadas por *M. tuberculosis*. No obstante, esta regla no es aplicable a individuos con inmunosupresión debido a la infección por VIH, que puede dar como resultado una reacción de PPD por debajo de 10 mm de diámetro; o en los países endémicos donde las personas infectadas por micobacterias no tuberculosis pueden mostrar una reacción de PPD por encima de 10 mm de diámetro.

El progreso en los últimos años ha visto el desarrollo de ensayos in vitro basados en células T, sobre la base de la liberación de interferón-gamma y el uso de antígenos que son más específicos de *M. tuberculosis* que el PPD, a saber, ESAT-6 y CFP-10. Estas pruebas de alta especificidad parecen ser al menos tan sensibles como la prueba cutánea de la tuberculina y también demuestran menos reactividad cruzada debido a la vacunación con BCG. Véase en Pai M y col. Expert Rev. Mol. Diagn. 2006 6(3):413-422 una reciente revisión del diagnóstico de la TB latente. Sin embargo, dado que ESAT-6/CFP-10 son antígenos de fases iniciales, los ensayos basados en ESAT-6/CFP-10 solo pueden funcionar de manera óptima en personas con infección reciente. Por consiguiente, la identificación de nuevos antígenos asociados específicamente con la tuberculosis latente puede ayudar al desarrollo de ensayos más sensibles que podrían detectar infecciones latentes a más largo plazo.

Sigue habiendo la necesidad de estrategias eficaces para el tratamiento y prevención de la tuberculosis, en particular, para el tratamiento y la prevención de la TB latente y la prevención de la reactivación de la TB.

El documento US2003/236393 divulga procedimientos para identificar, aislar y mutageneizar genes de virulencia de las micobacterias.

Breve resumen de la invención

La presente invención se refiere en general a la identificación de Rv1753c como antígeno de la TB (en particular, un antígeno asociado con la TB latente) y a procedimientos y usos relacionados en la prevención y el tratamiento de la TB, en especial, la prevención y el tratamiento de la TB latente y la prevención o retraso de la reactivación de la TB.

La presente invención proporciona un polipéptido aislado, que comprende:

- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
- (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;

para su uso en el tratamiento o prevención de la tuberculosis.

El uso de un polipéptido que comprende:

- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
- (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;

en la fabricación de un medicamento para el tratamiento o la prevención de la TB, representa otro aspecto de la invención.

La presente invención proporciona un polinucleótido aislado que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende:

- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
- (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;

para su uso en el tratamiento o prevención de la tuberculosis.

El uso de un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende:

- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
- (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;

en la fabricación de un medicamento para el tratamiento o la prevención de la TB, representa otro aspecto de la

invención.

Adicionalmente, se proporciona una composición que comprende un componente Rv1753c y un componente M72, en la que dicho componente Rv1753c es un polipéptido que comprende:

- 5 (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
 (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;

y dicho componente M72 es:

- 10 (i) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos el 90 % con la SEQ ID No: 25; o
 (ii) un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos el 90 % con la SEQ ID No: 25;

15 y en el que dicho componente M72 se proporciona en forma de un polipéptido, el componente Rv1753c y un componente M72 se pueden proporcionar como dos componentes polipeptídicos individuales o como una proteína de fusión que comprende ambos componentes polipeptídicos.

Además, se proporciona una composición que comprende un componente Rv1753c y un componente M72, en la que dicho componente Rv1753c es un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende:

- 20 (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
 (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;

y dicho componente M72 es:

- 25 (i) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos el 90 % con la SEQ ID No: 25; o
 (ii) un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos el 90 % con la SEQ ID No: 25;

30 y en el que dicho componente M72 se proporciona en forma de un polipéptido, el componente Rv1753c y un componente M72 se pueden proporcionar como dos componentes polipeptídicos individuales o como una proteína de fusión que comprende ambos componentes polipeptídicos.

Por otro lado, se proporciona una composición que comprende:

(a) un polipéptido que comprende:

- 35 (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
 (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;

y

(b) monofosforil lípido A 3-de-O-acilado y QS21 en una formulación de liposoma.

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

45 Un sujeto que recibe un polipéptido, polinucleótido o composición de la invención puede haber sido vacunado anteriormente para la tuberculosis (por ejemplo, vacunado contra la infección por *M. tuberculosis*), tal como que ha sido vacunado con el bacilo de *Bacillus Calmette-Guerin* (BCG). Como alternativa, un sujeto que recibe un polipéptido, polinucleótido o composición de la invención puede no haber sido vacunado anteriormente para la tuberculosis (por ejemplo, no vacunado contra la infección por *M. tuberculosis*), tal como que no ha sido vacunado con el bacilo de *Bacillus Calmette-Guerin* (BCG).

Descripción de las figuras

50 Figura 1: Porcentaje de células CD4 y CD8 de ratones CB6F1 inmunizados que expresan las citocinas IFN-gamma y/o IL-2 y/o TNF-alfa el día 21 (es decir, 7 días después de la segunda inmunización).

- Figura 2: Perfil de citocinas el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD4 específica de antígeno en ratones CB6F1 inmunizados.
- Figura 3: Perfil de citocinas el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD8 específica de antígeno en ratones CB6F1 inmunizados.
- 5 Figura 4: Porcentaje de células CD4 y CD8 de ratones CB6F1 inmunizados que expresan las citocinas IFN-gamma y/o IL-2 y/o TNF-alfa el día 35 (es decir, 7 días después de la tercera inmunización).
- Figura 5: Perfil de citocinas el día 35 (es decir, 7 días después de la tercera inmunización) de la respuesta de CD4 específica de antígeno en ratones CB6F1 inmunizados.
- Figura 6: Perfil de citocinas el día 35 (es decir, 7 días después de la tercera inmunización) de la respuesta de CD8 específica de antígeno en ratones CB6F1 inmunizados.
- 10 Figura 7: Porcentaje de células CD4 y CD8 de ratones C57BL/6 inmunizados que expresan las citocinas IFN-gamma y/o IL-2 y/o TNF-alfa el día 21 (es decir, 7 días después de la segunda inmunización).
- Figura 8: Perfil de citocinas el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD4 específica de antígeno en ratones C57BL/6 inmunizados.
- 15 Figura 9: Perfil de citocinas el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD8 específica de antígeno en ratones C57BL/6 inmunizados.
- Figura 10: Respuestas de células T CD4 específicas de antígeno en seres humanos sin tratamiento previo y con infección latente.

DESCRIPCIÓN DE LAS SECUENCIAS DE LA LISTA

- 20 SEQ ID NO: 1: secuencia polipeptídica de Rv1753c de la cepa H37Rv de *M. tuberculosis*.
 SEQ ID NO: 2: secuencia polipeptídica de Rv1753c de la cepa H37Rv de *M. tuberculosis*.
 SEQ ID NO: 3: secuencia polipeptídica de Rv1753c de la cepa CDC1551 de *M. tuberculosis*.
 SEQ ID NO: 4: secuencia polipeptídica de Rv1753c de la cepa F11 de *M. tuberculosis*.
 SEQ ID NO: 5: secuencia polipeptídica de Rv1753c de la cepa Haarlem de *M. tuberculosis*.
 25 SEQ ID NO: 6: secuencia polipeptídica de Rv1753c de la cepa C de *M. tuberculosis*.
 SEQ ID NO: 7: secuencia polipeptídica de Rv1753c de BCG.
 SEQ ID NO: 8: secuencia polipeptídica de Mtb8.4.
 SEQ ID NO: 9: secuencia polipeptídica de Mtb9.8.
 SEQ ID NO: 10: secuencia polipeptídica de Mtb9.9.
 30 SEQ ID NO: 11: secuencia polipeptídica de Ra12.
 SEQ ID NO: 12: secuencia polipeptídica de Ra35.
 SEQ ID NO: 13: secuencia polipeptídica de TbH9.
 SEQ ID NO: 14: secuencia polipeptídica de Mtb40.
 SEQ ID NO: 15: secuencia polipeptídica de Mtb41.
 35 SEQ ID NO: 16: secuencia polipeptídica de ESAT-6.
 SEQ ID NO: 17: secuencia polipeptídica de Ag85A.
 SEQ ID NO: 18: secuencia polipeptídica de Ag85B.
 SEQ ID NO: 19: secuencia polipeptídica de alfa-cristalina.
 SEQ ID NO: 20: secuencia polipeptídica de MPT64.
 40 SEQ ID NO: 21: secuencia polipeptídica de Mtb32A.
 SEQ ID NO: 22: secuencia polipeptídica de Mtb32A madura mutada en Ser/Ala.
 SEQ ID NO: 23: secuencia polipeptídica de TB10.4.
 SEQ ID NO: 24: secuencia polipeptídica de Mtb72f.
 SEQ ID NO: 25: secuencia polipeptídica de M72.
 45 SEQ ID NO: 26: secuencia polipeptídica de Mtb71f.
 SEQ ID NO: 27: secuencia polipeptídica de la fusión M92.
 SEQ ID NO: 28: secuencia polipeptídica de la fusión M103.
 SEQ ID NO: 29: secuencia polipeptídica de la fusión M114.
 SEQ ID NO: 30: epítipo 1 de células CD4 humanas putativas.
 50 SEQ ID NO: 31: epítipo 2 de células CD4 humanas putativas.
 SEQ ID NO: 32: epítipo 3 de células CD4 humanas putativas.
 SEQ ID NO: 33: epítipo 4 de células CD4 humanas putativas.
 SEQ ID NO: 34: epítipo 5 de células CD4 humanas putativas.
 SEQ ID NO: 35: epítipo 6 de células CD4 humanas putativas.
 55 SEQ ID NO: 36: epítipo 7 de células CD4 humanas putativas.
 SEQ ID NO: 37: epítipo 8 de células CD4 humanas putativas.
 SEQ ID NO: 38: epítipo 9 de células CD4 humanas putativas.
 SEQ ID NO: 39: epítipo 10 de células CD4 humanas putativas.
 SEQ ID NO: 40: epítipo 11 de células CD4 humanas putativas.
 60 SEQ ID NO: 41: epítipo 12 de células CD4 humanas putativas.
 SEQ ID NO: 42: epítipo 13 de células CD4 humanas putativas.
 SEQ ID NO: 43: epítipo 14 de células CD4 humanas putativas.
 SEQ ID NO: 44: epítipo 15 de células CD4 humanas putativas.
 SEQ ID NO: 45: epítipo 16 de células CD4 humanas putativas.

	SEQ ID NO: 244:	epítoto 185 de células CD8 humanas putativas.
	SEQ ID NO: 245:	epítoto 186 de células CD8 humanas putativas.
	SEQ ID NO: 246:	epítoto 187 de células CD8 humanas putativas.
	SEQ ID NO: 247:	epítoto 188 de células CD8 humanas putativas.
5	SEQ ID NO: 248:	secuencia polipeptídica de Rv2386c de la cepa H37Rv de <i>M. tuberculosis</i> .
	SEQ ID NO: 249:	secuencia polipeptídica de Rv2707c de la cepa H37Rv de <i>M. tuberculosis</i> .

Descripción detallada

Actualmente, la vacunación con bacterias vivas es el procedimiento más eficaz para inducir inmunidad protectora. El *Mycobacterium* más habitual empleado para este fin es el *bacilo de Calmette-Guerin* (BCG), una cepa no virulenta de *M. bovis* que se desarrolló hace más de 60 años. Sin embargo, la seguridad y la eficacia de BCG es una fuente de controversia, mientras que la protección contra la manifestación de la enfermedad grave en niños, el BCG no impide el establecimiento de TB latente o reactivación de la enfermedad pulmonar en la vida adulta. Adicionalmente, algunos países, tal como Estados Unidos, no vacunan al público en general con este agente.

Casi todas las vacunas de TB de nueva generación que actualmente se encuentran en desarrollo clínico han sido diseñadas como vacunas preexposición. Estas incluyen vacunas de subunidades, que han sido particularmente eficaces en aumentar la inmunidad inducida por la vacunación previa con BCG y vacunas de micobacterias vivas avanzadas que tienen por objeto sustituir al BCG con cepas más eficientes y/o seguras. Aunque estas vacunas tienen como objetivo mejorar la resistencia a la infección, que es probable que sean menos eficaces como postexposición o vacunas terapéuticas en los casos de TB latentes (Lin MY y col. *Endocrine, Metabolic y Immune Disorders - Drug Targets* 2008 8:15-29).

Se ha demostrado que varias de las proteínas que se expresan fuertemente durante las primeras etapas de la infección por *Mycobacterium* proporcionan un fuerte eficacia protectora en modelos de vacunación animal. Sin embargo, la vacunación con antígenos que son altamente expresado durante las primeras etapas de la infección puede no proporcionar una respuesta inmunitaria óptima para hacer frente a las etapas posteriores de la infección. El control adecuado durante las etapas posteriores de la infección puede requerir células T que son específicas de antígenos particulares que se expresan en ese momento.

Las vacunas postexposición dirigidas directamente a las bacterias persistentes latentes pueden ayudar en la protección contra la reactivación de la TB, mejorando así el control de la TB o incluso permitiendo el aclaramiento de la infección. Una vacuna dirigida a la TB latente podría, por tanto, de manera significativa y económica reducir las tasas de infección de tuberculosis a nivel mundial.

Las vacunas de subunidades basadas en antígenos de la última etapa también se podrían utilizar en combinación con antígenos en fase inicial para proporcionar una vacuna de múltiples fases. Como alternativa, los antígenos en fase tardía podrían usarse para complementar y mejorar la vacunación con BCG (ya sea reforzando la respuesta de BCG o mediante el desarrollo de cepas de BCG recombinantes avanzadas).

Recientemente, se ha propuesto una serie de candidatos a vacunas de *M. tuberculosis* según un análisis de la bioinformática, de la totalidad del genoma de *M. tuberculosis* (Zvi y col. *BMC Medical Genetics* 2008 1: 18) y en el ensayo de proteínas expresadas diferencialmente en individuos con infección activa y latente (Schuck SD y col. *PLoS ONE* 2009 4 (5):. e5590 China).

Aunque se ha demostrado que los macrófagos actúan como los principales efectores de la inmunidad de *Mycobacterium*, las células T son los inductores predominantes de dicha inmunidad. El papel esencial de las células T en la protección contra la tuberculosis se ilustra mediante los mayores índices de reactivación de la TB en individuos infectados por el virus de la inmunodeficiencia humana, debido la depleción asociada de células T CD4 +. Por otro lado, se ha demostrado que la transferencia adoptiva de células T CD4 + tomadas a la altura de la respuesta inmunitaria primaria a *M. tuberculosis* confiere protección contra *M. tuberculosis* en ratones deficientes en células T (Orme y col. *J. Exp Med.* 1983, 158:74-83).

Las células T CD4 + reactivas a *Mycobacterium* han demostrado ser potentes productores de γ -interferón (IFN- γ), que, a su vez, se ha demostrado que desencadenan los efectos anti-micobacterianos de los macrófagos en ratones (Flynn y col., *J. Exp. Med.* 1993, 178:2249-2254). Mientras que el papel de IFN- γ en los seres humanos está menos claro, los estudios han demostrado que la 1, 25-dihidroxi-vitamina D3, ya sea sola o en combinación con IFN- γ o factor de necrosis tumoral alfa, activa los macrófagos humanos para inhibir la infección por *M. tuberculosis*. Por otro lado, se sabe que el IFN- γ estimula los macrófagos humanos para hacer 1, 25-dihidroxi-vitamina D3. De manera similar, se ha demostrado que la interleucina-12 (IL-12) desempeña un papel en la estimulación 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 ed., 1994), *Tuberculosis* (2ª ed., Rom y Garay, eds., 2003) y *Harrison's Principles of Internal Medicine*, Capítulo 150, pág. 953-966 (16ª ed., Braunwald, y col., eds., 2005).

La presente invención se refiere en general a la identificación de Rv1753c como antígeno de la TB (en particular, un antígeno asociado con la TB latente) y a usos relacionados en la prevención y el tratamiento de la TB, en especial, la

prevención y el tratamiento de la TB latente y la prevención o retraso de la reactivación de la TB.

5 Por consiguiente, la invención proporciona una proteína Rv1753c, variante de la misma o un fragmento inmunogénico de la misma, o un polinucleótido que codifica dicha proteína, variante o fragmento, para su uso en el tratamiento o la prevención de la TB. De manera adecuada, el uso puede ser específicamente en la prevención y tratamiento de la TB latente (especialmente el tratamiento de la TB latente). Como alternativa, el uso puede ser en la prevención o retraso de la reactivación de la TB (especialmente el retraso de la reactivación de la TB, por ejemplo, por un período de meses, años o incluso indefinidamente).

10 La expresión "especie de *Mycobacterium* del complejo de tuberculosis" incluye las especies que tradicionalmente se consideran causantes de la enfermedad tuberculosis, así como especies ambientales y oportunistas de *Mycobacterium* que causan la tuberculosis y enfermedades pulmonares en pacientes inmunocomprometidos, tal como 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, pág. 953-966 (16ª ed., Braunwald, y col., eds., 2005). La presente invención se refiere particularmente a la infección por *M. tuberculosis*.

15 La expresión "infección activa" se refiere a una infección (por ejemplo, infección por *M. tuberculosis*) con síntomas y/o lesiones manifestados enfermedad (adecuadamente con síntomas de la enfermedad manifestados).

Las expresiones "infección inactiva", "infección durmiente" o "infección latente" se refieren a una infección (por ejemplo, infección por *M. tuberculosis*) sin síntomas y/o lesiones manifestados de la enfermedad (adecuadamente sin síntomas manifestados de la enfermedad).

20 La expresión "tuberculosis primaria" se refiere a la enfermedad clínica (manifestación de síntomas de la enfermedad) directamente después de la infección (por ejemplo, infección por *M. tuberculosis*). Véase, Harrison's Principles of Internal Medicine, Capítulo 150, pág. 953-966 (16ª ed., Braunwald, y col., eds., 2005).

25 Los términos "tuberculosis secundaria" o "tuberculosis posprimaria" se refieren a la reactivación de una infección durmiente, inactiva o latente (por ejemplo, infección por *M. tuberculosis*). Véase, Harrison's Principles of Internal Medicine, Capítulo 150, pág. 953-966 (16ª ed., Braunwald, y col., eds., 2005).

30 La expresión "reactivación de la tuberculosis" se refiere a la manifestación posterior de síntomas de la enfermedad en un individuo que ha dado un resultado positivo en la prueba para infección (por ejemplo, en una prueba de la tuberculina, adecuadamente en un ensayo *in vitro* basado en células T), pero que no tiene síntomas de la enfermedad aparentes. La prueba de diagnóstico positivo indica que el individuo está infectado, sin embargo, el individuo puede o no puede haber manifestado previamente síntomas de enfermedad activa que habían sido tratados suficiente y han llevado a la tuberculosis a un estado inactivo o latente. Se reconocerá que los procedimientos para la prevención, retraso o tratamiento de la reactivación de la tuberculosis se pueden iniciar en un individuo que manifiesta síntomas activos de la enfermedad.

35 La expresión tuberculosis "resistente a fármacos" se refiere a una infección (por ejemplo, infección por *M. tuberculosis*), en la que la cepa infectante no se mantiene estática o muerta (es decir, es resistente a) uno o más de los llamados agentes quimioterapéuticos "de primera línea" eficaces en el tratamiento de la tuberculosis (por ejemplo, isoniazida, rifampicina, etambutol, estreptomycin y pirazinamida).

40 La expresión tuberculosis "resistente a múltiples fármacos" se refiere a una infección (por ejemplo, infección por *M. tuberculosis*) en el que la cepa infectante es resistente a dos o más de los agentes quimioterapéuticos de "primera línea" eficaces en el tratamiento de la tuberculosis.

45 Un "agente quimioterapéutico" se refiere a un agente farmacológico conocido y utilizado en la técnica para tratar la tuberculosis (por ejemplo, infección por *M. tuberculosis*). Los agentes farmacológicos de ejemplo usados para tratar la tuberculosis incluyen, pero no se limitan a los mismos, amikacina, ácido aminosalicílico, capreomicina, cicloserina, etambutol, etionamida, isoniazida, kanamicina, pirazinamida, rifamicinas (es decir, rifampicina, rifapentina y rifabutin), estreptomycin, ofloxacin, ciprofloxacino, claritromicina, azitromicina y fluoroquinolonas. Entre los agentes quimioterapéuticos de "primera línea" usados para tratar la tuberculosis que no es resistente a fármacos se incluyen isoniazida, rifampicina, etambutol, estreptomycin y pirazinamida. Entre los agentes quimioterapéuticos de "segunda línea" usados para tratar la tuberculosis que ha mostrado resistencia farmacológica a uno o más fármacos de "primera línea" se incluyen ofloxacin, ciprofloxacino, etionamida, ácido aminosalicílico, cicloserina, amikacina, kanamicina y capreomicina. Dichos agentes farmacológicos se revisan en el Capítulo 48 de *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, Hardman y Limbird eds., 2001.

55 Los términos "polipéptido", "péptido" y "proteína" se usan de forma intercambiable en el presente documento para hacer referencia a un polímero de restos de aminoácidos. Los términos también se aplican a polímeros de aminoácidos en los que uno o más restos de aminoácidos es un mimético químico artificial de un aminoácido natural correspondiente, así como a polímeros de aminoácidos de origen natural y polímeros de aminoácidos de origen natural. Adecuadamente, un polipéptido de acuerdo con la presente invención consistirá solamente en restos de aminoácidos de origen natural, especialmente los aminoácidos codificados por el código genético.

El término "aminoácido" se refiere a aminoácidos naturales y sintéticos, así como análogos de aminoácidos y miméticos de aminoácidos que funcionan de un modo similar a los aminoácidos naturales. Los aminoácidos naturales son los codificados por el código genético, así como los aminoácidos que se modifican después, por ejemplo hidroxiprolina, γ -carboxiglutamato y O-fosfoserina. Análogos de aminoácidos se refiere a compuestos que tienen la misma estructura química básica que un aminoácido natural, es decir un carbono que está unido a un hidrógeno, un grupo carboxilo, un grupo amino y un grupo R, tal como homoserina, norleucina, metionina sulfóxido, metilmetionina sulfonio. Dichos análogos tienen grupos R modificados (p. ej., norleucina) o estructuras peptídicas modificadas, peor conservan la misma estructura química básica que un aminoácido natural. "Miméticos de aminoácidos" hace referencia a compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido pero que funciona de un modo similar a un aminoácido natural. Adecuadamente, un aminoácido es un aminoácido de origen natural o un análogo de aminoácido, especialmente un aminoácido de origen natural y, en particular, los aminoácidos codificados por el código genético.

"Ácido nucleico" se refiere a desoxirribonucleótidos o ribonucleótidos y polímeros de los mismos en forma monocatenaria o bicatenaria. El término abarca ácidos nucleicos que contienen análogos nucleotídicos conocidos o restos del esqueleto modificados o enlaces, que son sintéticos, de origen natural y de origen no natural, que tienen propiedades de unión similares al ácido nucleico de referencia y que se metabolizan de una manera similar a la de los nucleótidos de referencia. Ejemplos de tales análogos incluyen, sin limitación, fosforotioatos, fosforamidatos, fosfonatos de metilo, fosfonatos de metilo quiral, ribonucleótidos de 2-O-metilo, ácidos nucleicos peptídicos (PNAS). Adecuadamente, la expresión "ácido nucleico" se refiere a que ocurre naturalmente desoxirribonucleótidos o ribonucleótidos de origen natural y polímeros de los mismos.

A menos que se indique lo contrario, una secuencia de ácido nucleico concreta también abarca implícitamente variantes de la misma modificadas de forma conservadora (p. ej., sustituciones de codones degeneradas) y secuencias complementarias, así como la secuencia que se indica explícitamente (adecuadamente se refiere a la secuencia indicada explícitamente). Específicamente, se pueden conseguir sustituciones de codones degeneradas generando secuencias en las que la tercera posición de uno o más (o todos) codones seleccionados está sustituida con restos de base mixta y/o desoxinosina (Batzer y col., *Nucleic Acids Res.* 19:5081 (1991); Ohtsuka y col., *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini y col., *Mol. Cell. Probes* 8:91-98 (1994)). La expresión ácido nucleico se usa de forma intercambiable con gen, ADNc, ARNm, oligonucleótido y polinucleótido.

En el presente documento se hace referencia a los aminoácidos bien por su símbolos de tres letras conocido o por el símbolo de una letra recomendados por la IUPAC-IUB Biochemical Nomenclature Commission. Asimismo, se puede hacer referencia a los nucleótidos por sus códigos de una sola letra comúnmente aceptados.

Con la expresión "secuencia de la proteína Rv1753c", tal como se utiliza en el presente documento, se entiende la secuencia polipeptídica proporcionada en la SEQ ID No: 1 o un homólogo del mismo 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 es ambiental u oportunista y que causa infecciones oportunistas, tales como infecciones pulmonares en huéspedes inmunodeprimidos (por ejemplo, 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, pág. 953-966, 16ª ed., Braunwald, y col., eds., 2005).

Para asegurar una alta tasa de eficacia entre los huéspedes vacunados, los componentes de una vacuna deben estar bien conservados entre las cepas de importancia clínica. De manera adecuada, la proteína Rv1753c deriva de H37Rv de *M. tuberculosis* (es decir, la secuencia polipeptídica proporcionada en la SEQ ID No: 1)) o un homólogo de la misma 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 la resistencia a fármacos son una base particularmente valiosa para la secuencia de la proteína Rv1753c. Las cepas de interés incluyen:

CDC1551 - cepa transmisible y virulenta

Familia Haarlem (tal como, Haarlem A) - cepas resistentes a fármacos que se encuentran en poblaciones humanas hacinadas. 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 se descubrió en Haarlem, Países Bajos.

KZN4207 - aislada sensible a fármacos de pacientes en KwaZulu-Natal, Sudáfrica

KZN1435 - aislada resistente a múltiples fármacos (MDR) de pacientes en KwaZulu-Natal, Sudáfrica

KZN605 - aislada extensamente resistente a fármacos (MDR) de pacientes en KwaZulu-Natal, Sudáfrica

C - Altamente transmitida en la ciudad de Nueva York. En un estudio se encontró que esta cepa era más habitual entre los usuarios de fármacos inyectables y resistente a los productos intermedios de nitrógeno reactivos (Friedman y col., *J. Infect. Dis.* 1997 176(2):478-84)

94-M4241A - Aislado en San Francisco en 1994 de un paciente nacido en China. Esta cepa se analizó

anteriormente mediante análisis de delección genómica (Gagneux y col., PNAS 2006 103(8):2869-2873).

02-1987 - Aislado en San Francisco en 2002 de un paciente nacido en Corea del Sur. Esta cepa se analizó anteriormente mediante análisis de delección genómica (Gagneux y col., PNAS 2006 103(8):2869-2873).

5 T92 - Aislado en San Francisco en 1999 de un paciente nacido en Filipinas. Esta cepa se publicó en Hirsh y col., PNAS 2004 101:4871-4876).

T85 A - aislada en San Francisco en 1998 de un paciente nacido en China. Esta cepa se publicó en Hirsh y col., PNAS 2004 101:4871-4876).

EAS054 - aislada en San Francisco en 1993 de un paciente nacido en India. Esta cepa se analizó anteriormente mediante análisis de delección genómica (Gagneux y col., PNAS 2006 103(8):2869-2873).

10 Gagneux y col., PNAS 2006 103(8):2869-2873 y Herbert y col., Infect. Immun. 2007 75(12):5798-5805 proporcionan valiosos antecedentes en la serie de cepas de *M. tuberculosis* que se sabe que existen.

Lo más adecuado, la proteína Rv1753c se selecciona de las secuencias polipeptídicas proporcionadas en las SEQ ID No: 1 y 3-7, en particular en las SEQ ID No: 1 y 3-6, tal como la SEQ ID No: 1.

15 Los polinucleótidos de interés particular son los que comprenden (tal como que consisten en) una secuencia que codifica:

- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
- (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7.

20 Los polinucleótidos comprenderán adecuadamente (tal como que consisten en) una variante de la SEQ ID NO: 2.

COMBINACIONES

25 Los polipéptidos relacionados con Rv1753c de la presente invención pueden comprender además otros componentes diseñados para mejorar su inmunogenicidad o para mejorar estos antígenos en otros aspectos. Por ejemplo, la mejora de aislamiento de los antígenos polipeptídicos puede facilitarse mediante la adición de un tramo de restos de histidina (habitualmente conocido como cola de his) hacia un extremo del antígeno.

30 El término "cola de his" se refiere a una cadena de restos de histidina, típicamente seis restos, que se insertan dentro de la secuencia de referencia. Para minimizar la interrupción de la actividad asociada con la secuencia de referencia, un marcador his se inserta típicamente en el extremo N, por lo general inmediatamente después del resto de metionina de iniciación o bien en el extremo C-terminal. Por lo general son heterólogos a la secuencia nativa, pero están incorporados, ya que facilitan el aislamiento mediante la mejora de la proteína de unión a resinas de cromatografía de afinidad de metal inmovilizado (IMAC). En términos generales, la presencia o ausencia de una cola de His no es de importancia desde el punto de vista de provocar una respuesta inmune deseable contra la proteína de referencia. Sin embargo, para evitar el riesgo de una reacción adversa frente a la propia cola de his, se considera mejor minimizar la longitud de la cola de his, por ejemplo, a cuatro o menos restos, en particular a dos restos (o excluir el uso de una cola de his completamente).

35 Para mejorar la magnitud y/o amplitud de la respuesta inmunitaria provocada, las composiciones, polipéptidos y ácidos nucleicos de la invención pueden comprender múltiples copias del antígeno de la invención y/o polipéptidos heterólogos adicionales (o polinucleótidos que los codifican) de especies de *Mycobacterium* (en particular, *M. tuberculosis*).

40 Un experto en la materia reconocerá que cuando se utiliza un número de componentes combinados, la presentación precisa se puede modificar. Por ejemplo, un componente Rv1753c y una copia adicional del antígeno de la invención o un componente antigénico heterólogo adicional se podrían presentar:

- (1) dos componentes polipeptídicos individuales;
- (2) como una proteína de fusión que comprende ambos componentes polipeptídicos;
- 45 (3) como un componente polipeptídico y uno polinucleotídico;
- (4) dos componentes polinucleotídicos individuales;
- (5) como un único polinucleótido que codifica dos componentes polipeptídicos individuales; o
- (6) como un único polinucleótido que codifica una proteína de fusión que comprende ambos componentes polipeptídicos.

50 Esta flexibilidad se aplica igualmente a las situaciones en que se usan tres o más componentes en combinación. Sin embargo, por comodidad, a menudo es deseable que cuando están presentes un número de componentes, estén contenidos dentro de una única proteína de fusión o un polinucleótido que codifica una proteína de fusión única. En una realización de la invención, todos los componentes antigénicos se proporcionan como polipéptidos (por ejemplo,

dentro de una única proteína de fusión). En una realización alternativa de la invención, todos los componentes antigénicos se proporcionan como polinucleótidos (por ejemplo, un polinucleótido único, tal como uno que codifica una proteína de fusión individual).

5 El término "heterólogo", cuando se usa 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 entre sí en la naturaleza. Por ejemplo, el ácido nucleico típicamente se produce de forma recombinante, teniendo dos o más secuencias de genes no relacionados dispuestos para hacer un nuevo ácido nucleico funcional, por ejemplo, un promotor de una fuente y una región codificante de otra fuente. De manera similar, una proteína heteróloga indica que la proteína comprende dos o más subsecuencias que no se encuentran en la misma relación entre sí en la naturaleza (por ejemplo, una
10 proteína de fusión).

"Polipéptido de fusión" o "proteína de fusión" hace referencia a una proteína que tiene al menos dos polipéptidos heterólogos (p. ej., al menos dos polipéptidos de *Mycobacterium* sp) unidos covalentemente, ya sea directamente o a través de un enlazador de aminoácidos. Los polipéptidos que forman la proteína de fusión están normalmente unidos al extremo C-terminal o al extremo N-terminal, aunque también pueden unirse de extremo C-terminal a extremo C-terminal, de extremo N-terminal a extremo N-terminal o de extremo N-terminal a extremo C-terminal. Los polipéptidos de la proteína de fusión pueden encontrarse en cualquier orden. Esta expresión también se refiere a variantes modificadas de forma conservadora, variantes polimórficas, alelos, mutantes, fragmentos inmunogénicos y homólogos entre especies de los antígenos que forman la proteína de fusión. Los antígenos de la tuberculosis de *Mycobacterium* se describen en Cole y col., Nature 393:537 (1998), que divulga la totalidad del genoma de *Mycobacterium tuberculosis*. Los antígenos de otras especies de *Mycobacterium* que corresponden a antígenos de *M. tuberculosis* se pueden identificar, por ejemplo, usando algoritmos de comparación de secuencias, como se describe en el presente documento u otros procedimientos conocidos por los expertos en la materia, por ejemplo, ensayos de hibridación y ensayos de unión a anticuerpo.
15
20

El término "fusionado" se refiere a la unión covalente entre dos polipéptidos en una proteína de fusión. Las polipéptidos están normalmente unidos covalentemente a través de un enlace peptídico, ya sea directamente o a través de un enlazador de aminoácidos. Opcionalmente, los péptidos se pueden unir a través de enlaces covalentes no peptídicos conocidos por los expertos en la técnica.
25

Antígenos de *M. tuberculosis* de ejemplo que se pueden combinar con Rv1753c incluyen uno o más de (por ejemplo, de 1 a 5, tal como de 1 a 3, en particular 1) los siguiente (tal como uno o más de (i) a (xii)):

30 (i) Mtb8.4 (también conocido como DPV y Rv1174c), cuya secuencia polipeptídica se describe en SEQ ID No: 102 del documento WO97/09428 (ADNc en la SEQ ID No: 101) y en Coler y col. Journal of Immunology 1998 161:2356-2364. De particular interés es la secuencia de Mtb8.4 madura que carece del péptido señal líder (es decir, los restos de aminoácidos 15-96 de la SEQ ID No: 102 del documento WO97/09428). La secuencia polipeptídica de longitud completa de Mtb8.4 se muestra en la SEQ ID No: 8;

35 (ii) Mtb9.8 (también conocido como MSL y Rv0287), cuya secuencia polipeptídica se describe en la SEQ ID No: 109 del documento WO98/53075 (fragmentos de MSL se divulgan en las SEQ ID No: 110-124 del documento WO98/53075, siendo las SEQ ID No: 119 y 120 de particular interés) 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 polipeptídica de longitud completa de Mtb9.8 se muestra en la SEQ ID No: 9;

40 (iii) Mtb9.9 (también conocido como Mtb9.9A, MTI, MTI-A y Rv1793) cuya secuencia polipeptídica se describe en la SEQ ID No: 19 del documento WO98/53075 y e Alderson y col. Journal of Experimental Medicine 2000 7:551-559 (los fragmentos de MTI se divulgan en las SEQ ID No: 17 y 51-66 del documento WO98/53075, siendo las SEQ ID No: 17, 51, 52, 53, 56 y 62-65 de interés particular). Un número de variantes de polipéptidos de MTI se describen en las SEQ ID No: 21, 23, 25, 27, 29 y 31 del documento WO98/53075 y en Alderson y col. Journal of Experimental Medicine 2000 7:551-559. La secuencia polipeptídica de longitud completa de Mtb9.9 se muestra en la SEQ ID No: 10;
45

(iv) Ra12 (también conocido como antígeno en C terminal de Mtb32A), cuya secuencia polipeptídica se describe en la SEQ ID No: 10 del documento WO01/98460 en Skeiky y col. Journal of Immunology 2004 172:7618-7682. La secuencia polipeptídica de longitud completa de Ra12 se muestra en la SEQ ID No: 11;

50 (v) Ra35 (también conocido como antígeno en N termina de Mtb32A), cuya secuencia polipeptídica se describe en la SEQ ID No: 8 del documento WO01/98460 en Skeiky y col. Journal of Immunology 2004 172:7618-7682. La secuencia polipeptídica de longitud completa de Ra35 se muestra en la SEQ ID No: 12;

(vi) TbH9 (también conocida como Mtb39, Mtb39A, TbH9FL y Rv1196) cuya secuencia polipeptídica se describe en la SEQ ID No: 107 del documento WO97/09428, y también en Dillon y col. Infection and Immunity 1999 67(6):2941-2950 y Skeiky y col. Journal of Immunology 2004 172:7618-7682. La secuencia polipeptídica de longitud completa de TbH9 se muestra en la SEQ ID No: 13;
55

(i) Mtb40 (también conocida como HTCC1 y Rv3616c) cuya secuencia polipeptídica se describe en la SEQ ID

No: 138 del documento WO98/53075 (ADNc en la SEQ ID No: 137). La secuencia polipeptídica de longitud completa de Mtb40 se muestra en la SEQ ID No: 14;

5 (viii) Mtb41 (también conocida como MTCC2 y Rv0915c) cuya secuencia polipeptídica se describe en SEQ ID No: 142 del documento WO98/53075 (ADNc en la SEQ ID No: 140) y en Coler y col. *Journal of Immunology* 2000 165:7140-7149. La secuencia polipeptídica de longitud completa de Mtb41 se muestra en la SEQ ID No: 15;

(ix) ESAT -6 (también conocida como esxA y Rv3875) cuya secuencia polipeptídica se describe en la SEQ ID No: 103 del documento WO97/09428 (ADNc en la SEQ ID No: 104) y en Sorensen y col. *Infection and Immunity* 1995 63(5):1710-1717. La secuencia polipeptídica de longitud completa de ESAT-6 se muestra en la SEQ ID No: 16;

10 (x) Antígenos del complejo Ag85 (por ejemplo, Ag85A, también conocidos como fbpA y Rv3804c; o Ag85B, también conocidos como bpB y Rv1886c) que se tratan, 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 polipeptídica de longitud completa de Ag85A se muestra en la SEQ ID No: 17 (la proteína madura de restos 43-338, es decir, que carece del péptido señal, es de particular interés). La secuencia polipeptídica de longitud completa de Ag85B se muestra en la SEQ ID No: 18 (la proteína madura de restos 41-325, es decir, que carece del péptido señal, es de particular interés);

15 (xi) Alfa-cristalina (también conocida como hspX y Rv2031c), que se describe en Verbon y col. *Journal of Bacteriology* 1992 174:1352-1359 y Friscia y col. *Clinical and Experimental Immunology* 1995 102:53-57 (son de particular interés los fragmentos correspondientes a los restos 71-91, 21-40, 91-110 y 111-130). La secuencia polipeptídica de longitud completa de alfa-cristalina se muestra en la SEQ ID No: 19;

20 (xii) Mpt64 (también conocida como Rv1980c) que se describe en Roche y col. *Scandinavian Journal of Immunology* 1996 43:662-670. La secuencia polipeptídica de longitud completa de MPT64 se muestra en la SEQ ID No: 20 (la proteína madura de restos 24-228, es decir, que carece del péptido señal, es de particular interés):

25 (xiii) Mtb32A, cuya secuencia polipeptídica se describe en la SEQ ID No: 2 (longitud completa) y los restos 8-330' de la SEQ ID No: 4 (madura) del documento WO01/98460, especialmente variantes que tienen al menos una de la tríada catalítica mutada (por ejemplo, el residuo de serina catalítico, que puede estar mutado, por ejemplo, a alanina). La secuencia polipeptídica de longitud completa de Mtb32A se muestra en la SEQ ID No: 21. La forma madura de Mtb32A que tiene una mutación de Ser/Ala se muestra en la SEQ ID No: 22;

(xiv) TB10.4, la secuencia polipeptídica de longitud completa de TB10.4 se muestra en la SEQ ID No: 23;

30 (xv) Rv2386c, la secuencia polipeptídica de longitud completa de Rv2386c de H37Rv de *Mycobacterium tuberculosis* se muestra en la SEQ ID No: 248; y/o

(xvi) Rv2707c, la secuencia polipeptídica de longitud completa de Rv2707c de H37Rv de *Mycobacterium tuberculosis* se muestra en la SEQ ID No: 249,

o combinaciones de los mismos, tales como (por ejemplo, combinaciones tales como (a) a (g)):

35 (a) una combinación de los componentes Ra12, TbH9 y Ra35, por ejemplo, en forma de una proteína de fusión, tal como Mtb72f. La secuencia polipeptídica de Mtb72f se describe en la SEQ ID No: 6 del documento WO2006/117240 (ADNc en la SEQ ID No: 5) y en Skeiky y col. *Journal of Immunology* 2004 172:7618-7682 (donde incorpora una cola de his opcional para ayudar a la purificación, cuando se utiliza en la presente invención, adecuadamente Mtb72f carece de los restos de His opcionales). La secuencia polipeptídica de Mtb72f se muestra en la SEQ ID No: 24;

40 (b) una combinación de los componentes Ra12, TbH9 y Ra35 con mutación de Ser/Ala (es decir, cuando el residuo de serina catalítico se ha sustituido con alanina), por ejemplo, en forma de una proteína de fusión, tal como M72. La secuencia polipeptídica de M72 se describe en la SEQ ID No: 4 del documento WO2006/117240 (ADNc en la SEQ ID No: 3) donde incorpora una doble histidina opcional para ayudar a la fabricación, cuando se utiliza en la presente invención, M72 también puede incorporar una doble histidina, aunque adecuadamente M72 carece de la doble histidina opcional (es decir, los restos 4-725 de la SEQ ID No: 4 del documento WO2006/117240 son de particular interés). La secuencia polipeptídica de M72 se muestra en la SEQ ID No: 25;

45 (c) una combinación de los componentes Mtb8.4, Mtb9.8, Mtb9.9 y Mtb41, por ejemplo, en forma de una proteína de fusión, tal como Mtb71f. La secuencia polipeptídica de Mtb71f se describe en la SEQ ID No: 16 del documento WO99/051748 (ADNc en la SEQ ID No: 15) donde incorpora una cola de his opcional para ayudar a la purificación, cuando se utiliza en la presente invención, adecuadamente Mtb71f corresponde a los restos de aminoácidos 9-710 de la SEQ ID No: 16 del documento WO99/051748. La secuencia polipeptídica de Mtb71f se muestra en la SEQ ID No: 26;

50 (d) una combinación de Mtb72f o M72 (adecuadamente sin los restos de histidina opcionales para ayudar a la expresión) con Mtb9.8 y Mtb9.9, por ejemplo en una proteína de fusión. La secuencia polipeptídica de una fusión

M72-Mtb9.9-Mtb9.8 se muestra en la SEQ ID No: 27 (fusión M92), cuando se usa en la presente invención, la fusión M72-Mtb9.9-Mtb9.8 puede incorporar, opcionalmente, una doble histidina tras el residuo de metionina inicial para ayudar a la fabricación;

5 (e) una combinación de Mtb72f o M72 (adecuadamente sin los restos de histidina opcionales para ayudar a la expresión) con Ag85B, por ejemplo en una proteína de fusión, tal como Mtb103f. La secuencia polipeptídica de Mtb103f se describe en la SEQ ID No: 18 del documento WO03/070187 (ADNc en la SEQ ID No: 10) donde incorpora una cola de his opcional para ayudar a la purificación, cuando se utiliza en la presente invención, adecuadamente Mtb103f corresponde a los restos de aminoácidos 8-1016 de la SEQ ID No: 18 del documento WO03/070187. También es de particular interés M103, es decir, Mtb103f que incorpora una mutación Ser/Ala en el componente Ra35, cuando se utiliza en la presente invención, adecuadamente M103 corresponde a los restos de aminoácidos 8-1016 de la SEQ ID No: 18 del documento WO03/070187, en el que el residuo de Ser en la posición 710 se ha reemplazado por Ala. La secuencia polipeptídica de M103 se muestra en la SEQ ID No: 28, cuando se utiliza en la presente invención, la fusión M72-Mtb9.9-Mtb9.8 puede incorporar, opcionalmente, una doble histidina tras el residuo de metionina inicial para ayudar a la fabricación;

15 (f) una combinación de Mtb72f o M72 (adecuadamente sin los restos de histidina opcionales para ayudar a la expresión) con Mtb41, por ejemplo en una proteína de fusión, tal como Mtb114f. La secuencia polipeptídica de Mtb114f se describe en la SEQ ID No: 16 del documento WO03/070187 (ADNc en la SEQ ID No: 9) donde incorpora una cola de his opcional para ayudar a la purificación, cuando se utiliza en la presente invención, adecuadamente Mtb114f corresponde a los restos de aminoácidos 8-1154 de la SEQ ID No: 16 del documento WO03/070187. También es de particular interés M114, es decir, Mtb114f que incorpora una mutación Ser/Ala en el componente Ra35, cuando se utiliza en la presente invención, adecuadamente M114 corresponde a los restos de aminoácidos 8-1154 de la SEQ ID No: 16 del documento WO03/070187, en el que el residuo de Ser en la posición 710 se ha reemplazado por Ala. La secuencia polipeptídica de M114 se muestra en la SEQ ID No: 29, cuando se utiliza en la presente invención, la fusión M72-Mtb9.9-Mtb9.8 puede incorporar, opcionalmente, una doble histidina tras el residuo de metionina inicial para ayudar a la fabricación;

(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.

30 Las combinaciones de un componente Rv1753c y un componente Mtb40 son de particular interés. Obviamente, dichas combinaciones podrían contener, opcionalmente, otros componentes antigénicos adicionales (por ejemplo, un componente M72).

Otra combinación de interés comprende un componente Rv1753c y un componente M72.

Una combinación adicional de interés comprende un componente Rv1753c y un componente Rv2386c.

35 Otras combinaciones de interés incluyen las que comprenden un componente Rv1753c y un componente Rv2707c.

Una combinación adicional de interés comprende un componente Rv1753c y un componente alfa-cristalina.

El experto reconocerá que las combinaciones no tienen que basarse en las secuencias específicas descritas en (i) - (xvi) y (a) - (h) anteriores, y que las variantes modificadas de manera conservadora (por ejemplo, que tiene una de identidad de al menos el 70 %, tal como una identidad de al menos el 80 %, en particular una identidad de al menos el 90 % y especialmente una identidad de al menos el 95 %) o fragmentos inmunogénicos (por ejemplo, al menos un 20 % del antígeno de longitud completa, tal como al menos el 50 % del antígeno, en particular al menos el 70 % y especialmente al menos el 80 %) de las secuencias descritas se pueden usar para alcanzar el mismo efecto práctico.

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

50 Muchos de los antígenos anteriores también se divulgan en las solicitudes de patente números 08/523.435, 08/523.436, 08/658.000, 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 PCT PCT/US98/10407, PCT/US98/10514, PCT/US99/03265, PCT/US99/03268, PCT/US99/07717, los documentos WO97/09428 y WO97/09429, el documento WO98/16645, el documento WO98/16646, cada una de las cuales se incorpora al presente documento a modo de referencia.

55 Las composiciones, polipéptidos y ácidos nucleicos de la invención también pueden comprender polipéptidos adicionales de otras fuentes. Por ejemplo, las composiciones y proteínas de fusión de la invención pueden incluir

5 polipéptidos o ácidos nucleicos que codifican polipéptidos, en las que el polipéptido potencia la expresión del antígeno, por ejemplo, NS1, una proteína del virus de la gripe (véanse, por ejemplo, los documentos WO99/40188 y WO93/04175). Los ácidos nucleicos de la invención pueden modificarse por ingeniería genética basada en la preferencia de codones en una especie de elección, por ejemplo, seres humanos (en el caso de la expresión *in vivo*) o una bacteria en particular (en el caso de la producción de polipéptido).

10 El componente Rv1753c también puede administrarse con uno o más agentes quimioterapéuticos eficaces contra la tuberculosis (por ejemplo, infección por *M. tuberculosis*). Entre los ejemplos de tales agentes quimioterapéuticos se incluyen, aunque sin limitación, amikacina, ácido aminosalicílico, capreomicina, cicloserina, etambutol, etionamida, isoniazida, kanamicina, pirazinamida, rifamicinas (es decir, rifampicina, rifapentina y rifabutina), estreptomina, ofloxacino, ciprofloxacino, claritromicina, azitromicina y fluoroquinolonas. Tal quimioterapia está determinada por el juicio del médico responsable del tratamiento usando combinaciones de fármacos preferentes. Los agentes quimioterapéuticos de "primera línea" usados para tratar la tuberculosis (por ejemplo, infección por *M. tuberculosis*) que no es resistente a fármacos incluyen isoniazida, rifampicina, etambutol, estreptomina y pirazinamida. Entre los agentes quimioterapéuticos de "segunda línea" usados para tratar la tuberculosis (por ejemplo, infección por *M. tuberculosis*) que ha mostrado resistencia farmacológica a uno o más fármacos de "primera línea" se incluyen ofloxacino, ciprofloxacino, etionamida, ácido aminosalicílico, cicloserina, amikacina, kanamicina y capreomicina.

15 Los agentes quimioterapéuticos convencionales se administran generalmente durante un período relativamente largo (aproximadamente 9 meses). La combinación de agentes quimioterapéuticos convencionales con la administración de un componente Rv1753c de acuerdo con la presente invención puede permitir reducir el período de tratamiento quimioterapéutico que se reduce (por ejemplo, a 8 meses, 7 meses, 6 meses, 5 meses, 4 meses, 3 meses o menos) sin una disminución de la eficacia.

20 De particular interés es el uso de un componente Rv1753c junto con el bacilo de Calmette-Guerin (BCG). Por ejemplo, en la forma de una BCG modificado que expresa de forma recombinante Rv1753c (o una variante o fragmento del mismo como se describe en el presente documento). Como alternativa, el componente Rv1753c se puede utilizar para mejorar la respuesta de un sujeto a la vacunación con BCG, ya sea mediante la coadministración o mediante refuerzo de una vacunación previa con el BCG. Cuando se usa para potenciar la respuesta de un sujeto a la vacunación con BCG, el componente Rv1753c, obviamente, se puede proporcionar en forma de un polipéptido o un polinucleótido (opcionalmente en combinación con componentes antigénicos adicionales, como se ha descrito anteriormente).

25 El experto en la materia reconocerá que no es necesario administrar las combinaciones de componentes juntas y se pueden aplicar: por separado o en combinación; al mismo tiempo, secuencialmente o dentro de un corto período de tiempo; a través de la misma vía o de vías diferentes. Sin embargo, por comodidad es generalmente deseable (cuando los regímenes de administración son compatibles) administrar una combinación de componentes como una sola composición.

30 Los polipéptidos, polinucleótidos y composiciones de la presente invención por lo general se administran a seres humanos, pero son eficaces en otros mamíferos, incluyendo mamíferos domésticos (por ejemplo, perros, gatos, conejos, ratas, ratones, cobayas, hámsters, chinchillas) y mamíferos agrícolas (por ejemplo, vacas, cerdos, ovejas, cabras, caballos).

FRAGMENTOS INMUNOGÉNICOS

35 Los epítomos de linfocitos T son tiras de aminoácidos contiguas y cortas que son reconocidas por los linfocitos T (por ejemplo, linfocitos T CD4+ o CD8+). La identificación de los epítomos de células T puede conseguirse mediante experimentos de mapeo de epítomos que son bien conocidos por el experto en la materia (véase, por ejemplo, Paul, *Fundamental Immunology*, 3ª ed., 243-247 (1993); Beißbarth y col. *Bioinformatics* 2005 21(Suppl. 1):i29-i37).

Como alternativa, los epítomos pueden predecirse utilizando los enfoques analizados en los Ejemplos.

40 Como resultado de la participación crucial de la respuesta de linfocitos T en la tuberculosis, es fácilmente evidente que los fragmentos del polipéptido Rv1753c de longitud completa que contienen al menos un epítomo de linfocitos T serán inmunogénicos y pueden contribuir a la inmunoprotección. Tales fragmentos se denominan en el presente documento como fragmentos inmunogénicos.

45 Normalmente, los fragmentos inmunogénicos comprenderán al menos 9 aminoácidos contiguos de la secuencia polipeptídica de longitud completa (por ejemplo, al menos 10), tales como al menos 12 aminoácidos contiguos (por ejemplo, al menos 15 o al menos 20 aminoácidos contiguos), en particular al menos 50 aminoácidos contiguos, tal como al menos 100 aminoácidos contiguos (por ejemplo, al menos 200 aminoácidos contiguos). Adecuadamente, los fragmentos inmunogénicos serán al menos un 20 %, tal como al menos un 50 %, al menos un 70 % o al menos un 80 % de la longitud de la secuencia polipeptídica de longitud completa.

55 Se entenderá que en una población exógama diversa, tal como los seres humanos, tipos diferentes de HLA significa que los epítomos específicos pueden no ser reconocidos por todos los miembros de la población. Por consiguiente, para maximizar el nivel de reconocimiento y escala de la respuesta inmunitaria a un polipéptido, generalmente es

deseable que un fragmento inmunogénico contenga una pluralidad de los epítomos de la secuencia de longitud completa (idóneamente todos los epítomos).

5 Entre los fragmentos particulares de la proteína Rv1753c que pueden ser útiles se incluyen los que contienen al menos un epítomo CD4+, adecuadamente al menos dos epítomos CD4+ y, especialmente, todos los epítomos CD4+ (tales como los epítomos descritos en los Ejemplos y en las SEQ ID Nos: 30-59, particularmente los asociados con una pluralidad de alelos de HLA, por ejemplo los asociados con 2, 3, 4, 5 o más alelos).

10 Entre otros fragmentos de la proteína Rv1753c que pueden ser útiles se incluyen los que contienen al menos un epítomo CD8, adecuadamente al menos dos epítomos CD8 y, especialmente, todos los epítomos CD8 (tales como los epítomos descritos en los Ejemplos y en las SEQ ID Nos: 60-247, particularmente los asociados con una pluralidad de alelos de HLA, por ejemplo los asociados con 2, 3, 4, 5 o más alelos).

15 Cuando se usa un fragmento individual del polipéptido de longitud completa, dicho fragmento se considera inmunogénico cuando provoca una respuesta que es al menos un 20 %, adecuadamente al menos un 50 % y, especialmente, al menos un 75 % (tal como al menos un 90 %) de la actividad de la secuencia de referencia en un ensayo de reestimulación *in vitro* de PBMC o sangre entera con antígenos específicos (por ejemplo, reestimulación durante un periodo de entre varias horas hasta dos semanas, tal como hasta un día, de 1 día a 1 semana o de 1 a 2 semanas) que mide la activación de las células mediante linfoproliferación, producción de citocinas en el sobrenadante de cultivo (medido mediante ELISA, CBA etc.) o caracterización de las respuestas de linfocitos T o B mediante tinción intracelular y extracelular (por ejemplo, usando anticuerpos específicos de marcadores inmunitarios, tales como CD3, CD4, CD8, IL2, TNFa, IFNg, CD40L, CD69 etc.), seguido de análisis con un citómetro de flujo. De manera adecuada, un fragmento se considera inmunogénico cuando provoca una respuesta que es al menos un 20 %, adecuadamente, al menos un 50 % y especialmente al menos un 75 % (tal como al menos un 90 %) de la actividad de la secuencia de referencia en un ensayo de proliferación de células T y/o de producción de IFN-gamma.

25 En algunas circunstancias se puede utilizar una pluralidad de fragmentos de polipéptido de longitud completa (que pueden o no solapar y puede o no puede cubrir la totalidad de la secuencia de longitud completa) para obtener una respuesta biológica equivalente a la propia secuencia de longitud completa. Por ejemplo, al menos dos fragmentos inmunogénicos (tal como tres, cuatro o cinco) como se ha descrito anteriormente, que en combinación proporcionan al menos un 50 %, adecuadamente, al menos un 75 % y especialmente al menos un 90 % de la actividad de la secuencia de referencia en un ensayo *in vitro* de reestimulación de PBMC o sangre entera (por ejemplo, un ensayo de proliferación de células T y/o de producción de IFN-gamma).

30 **VARIANTES**

35 "Variantes" o "variantes modificadas de forma conservadora" se aplica tanto a secuencias de aminoácidos como de ácidos nucleicos. Con respecto a secuencias de ácido nucleico concretas, "variantes modificadas de forma conservadora" se refiere a los ácidos nucleicos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o en las que el ácido nucleico no codifica una secuencia de aminoácidos o secuencias esencialmente idénticas.

40 Dada 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 ellos, el aminoácido alanina. Por lo tanto, en cada posición en la que una alanina está especificada por un codón, se puede alterar el codón y convertir en cualquiera de los correspondientes codones descritos sin que se altere el polipéptido codificado. Dichas variaciones en el ácido nucleico conducen a variantes "silentes" o "degeneradas", que son una especie de variantes modificadas de forma conservadora. En el presente documento, toda secuencia de ácido nucleico que codifica un polipéptido también describe toda posible variación silente del ácido nucleico. Un experto en la materia reconocerá que cada codón en un ácido nucleico (a excepción de AUG, que normalmente es el único codón para metionina, y TGG, que habitualmente es el único codón para triptófano) se puede modificar para dar una molécula funcionalmente idéntica. En consecuencia, cada variación silente de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita.

50 Un polipéptido de la invención puede contener una serie de variaciones silentes (por ejemplo, 1-50, tal como 1-25, en particular 1-5 y especialmente 1 codón, pueden estar alterados) cuando se comparan con la secuencia de referencia. Un polinucleótido de la invención puede contener una serie de variaciones conservadoras no silentes (por ejemplo, 1-50, tal como 1-25, en particular 1-5 y especialmente 1 codón, pueden estar alterados) cuando se comparan con la secuencia de referencia. Las variaciones no silentes son aquellas que dan lugar a un cambio en la secuencia de aminoácidos. codificada (bien a través de la sustitución, delección o adición de restos de aminoácidos). Los expertos en la materia reconocerán que una secuencia de polinucleótidos concreta puede contener variaciones conservadoras tanto silentes como no silentes.

55 En cuanto a las variantes de una secuencia de proteínas, el experto en la técnica reconocerá que sustituciones, delecciones o adiciones individuales en un polipéptido que alteren, añadan o deleccionen un solo aminoácido o un porcentaje pequeño de aminoácidos es una "variante modificada de forma conservadora" en la que la o las alteraciones tienen como resultado la sustitución de u aminoácido con una funcionalidad similar al aminoácido o la

sustitución/delección/adición de residuos que no afectan sustancialmente a la función biológica de la variante.

Las tablas de sustituciones conservadoras que proporcionan aminoácidos funcionalmente similares son bien conocidas en la materia. Tales variantes modificadas de manera conservadora se suman y no excluyen las variantes polimórficas, homólogos entre especies y alelos de la invención.

- 5 Un polipéptido de la invención puede contener una serie de sustituciones conservadoras (por ejemplo, 1-50, tal como 1-25, en particular 1-10 y especialmente 1 resto de aminoácido, pueden estar alterados) cuando se comparan con la secuencia de referencia. En general, dichas sustituciones conservadoras entrarán dentro del agrupamiento de aminoácidos especificado más adelante, aunque en algunas circunstancias pueden ser posibles otras sustituciones sin que ello afecte sustancialmente a las propiedades inmunogénicas del antígeno. Cada uno de los siguientes ocho grupos contiene aminoácidos que normalmente son sustituciones conservadoras uno de otro:

- 10
- 1) Alanina (A), Glicina (G);
 - 2) Ácido aspártico (D), Ácido glutámico (E);
 - 3) Asparagina (N), Glutamina (Q),
 - 4) Arginina (R), Lisina (K);
 - 15 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).

- 20 Idóneamente, dichas sustituciones no se producen en la región de un epítipo y, por tanto, no tienen un impacto significativo sobre las propiedades inmunogénicas del antígeno.

Las variantes polipeptídicas también incluyen aquéllas en las que se insertan aminoácidos adicionales en comparación con la secuencia de referencia, por ejemplo dichas inserciones pueden producirse en 1-10 localizaciones (tal como 1-5 localizaciones, idóneamente 1 o 2 localizaciones, en particular 1 localización) y pueden implicar, por ejemplo, la adición de 50 o menos aminoácidos en cada localización (tal como 20 o menos, en particular 10 o menos, especialmente 5 o menos). Idóneamente, dichas inserciones no se producen en la región de un epítipo y, por tanto, no tienen un impacto significativo sobre las propiedades inmunogénicas del antígeno. Un ejemplo de inserciones incluye una tira corta de restos de histidina (por ejemplo, 2-6 restos) para ayudar a la expresión y/o purificación del antígeno en cuestión.

- 30 Las variantes proteicas incluyen aquellas en las que los aminoácidos se han deletado en comparación con la secuencia de referencia, por ejemplo dichas delecciones pueden producirse en 1-10 localizaciones (tal como 1-5 localizaciones, idóneamente 1 o 2 localizaciones, en particular 1 localización) y pueden implicar, por ejemplo, la delección de 50 o menos aminoácidos en cada localización (tal como 20 o menos, en particular 10 o menos, especialmente 5 o menos). Idóneamente, dichas delecciones no se producen en la región de un epítipo y, por tanto, no tienen un impacto significativo sobre las propiedades inmunogénicas del antígeno.

El experto en la técnica reconocerá que una variante polipeptídica concreta puede comprender sustituciones, delecciones y adiciones (o cualquier combinación de los mismos).

Los procedimientos para determinar las regiones de epítipo de un antígeno se describen y se ilustran en los Ejemplos.

- 40 Las variantes exhiben, preferentemente, una identidad de al menos aproximadamente un 90 % (tal como al menos aproximadamente un 95 %, al menos aproximadamente un 98 % o al menos aproximadamente un 99 %) con la secuencia de referencia asociada.

Los términos "idéntico" o porcentaje de "identidad" en el contexto de dos o más secuencias de ácidos nucleicos o de polipéptidos se refiere a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de restos de aminoácidos o nucleótidos que son iguales (es decir, un 90 %, un 95 %, un 98 % o un 99 % de identidad sobre una región especificada), cuando se comparan y alinean para correspondencia máxima sobre una ventana de comparación, o una región designada medida usando uno de los siguientes algoritmos de comparación de secuencias o mediante alineamiento manual e inspección visual. Por tanto, se dice que dichas secuencias son "sustancialmente idénticas". Esta definición también se refiere a la complementaria de una secuencia problema.

50 Opcionalmente, la identidad existe sobre una región que tiene una longitud de al menos aproximadamente 25 a aproximadamente 50 aminoácidos o nucleótidos, u, opcionalmente, sobre una región que tiene una longitud de 75-100 aminoácidos o nucleótidos. De manera adecuada, la comparación se realiza sobre una ventana correspondiente con la longitud completa de la secuencia de referencia.

- 55 Para comparar secuencias, normalmente una secuencia actúa como secuencia de referencia con la que se comparan las secuencias problema. Cuando se usa un algoritmo de comparación de secuencias, las secuencias problema y de referencia se introducen en un ordenador, se designan las coordenadas de la subsecuencia, en caso

necesario, y se designan los parámetros del programa del algoritmo de secuencias. Se pueden usar los parámetros por defecto del programa o parámetros alternativos. A continuación, el algoritmo de comparación de secuencias calcula el porcentaje de identidades de secuencia para las secuencias problema con respecto a la secuencia de referencia, según los parámetros del programa.

- 5 Una “ventana de comparación”, como se usa en el presente documento, hace referencia a un segmento en el que una secuencia se puede comparar con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias están alineadas de forma óptima. Los procedimientos de alineación de secuencias para comparación son bien conocidos en la materia. La alineación óptima de secuencias para comparación se puede realizar, 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 la búsqueda para el procedimiento de similitud de Pearson y Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), mediante implementaciones computerizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el paquete de software 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., eds., suplemento de 1995)).

Un ejemplo de un algoritmo útil es PILEUP. PILEUP crea una alineación de múltiples secuencias de un grupo de secuencias relacionadas usando alineaciones pareadas progresivas para mostrar la relación y el porcentaje de identidad de secuencia. También representa un árbol o dendograma que muestra las relaciones de agrupamiento usadas para crear la alineación. PILEUP usa una simplificación del procedimiento de alineación progresiva de Feng y Doolittle, *J. Mol. Evol.* 35:351-360 (1987). El procedimiento usado es similar al procedimiento 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 alineación múltiple comienza con la alineación pareada de las dos secuencias más similares, produciendo un grupo de dos secuencias alineadas. Después, este grupo se alinea a la siguiente secuencia más relacionada, o grupo de secuencias alineadas. Dos grupos de secuencias se alinean mediante extensión simple de la alineación apareada de dos secuencias individuales. La alineación final se consigue mediante una serie de alineaciones pareadas progresivas. El programa se ejecuta diseñando secuencias específicas y sus coordenadas de aminoácidos o nucleótidos para regiones de comparación de secuencia y designando los parámetros del programa. Usando PILEUP, una secuencia de referencia se compara con otras secuencias problema para determinar la relación del porcentaje de identidad de secuencia usando los parámetros siguientes: peso del hueco por defecto (3,00), peso de la longitud del hueco por defecto (0,10) y huecos terminales pesados. PILEUP se puede obtener del paquete de software para análisis de secuencia GCG, por ejemplo, la versión 7.0 (Devareaux y col., *Nuc. Acids Res.* 12:387-395 (1984)).

Otro ejemplo de un algoritmo que es adecuado para determinar el porcentaje de identidad de secuencia y la similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que 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 realizar análisis BLAST está disponible para el público a través del National Center for Biotechnology Information (sitio web en www.ncbi.nlm.nih.gov/). Este algoritmo implica identificar primero pares de secuencia de alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia problema que coinciden o satisfacen alguna puntuación T umbral de valor positivo cuando se alinea con una palabra de la misma longitud en una secuencia de la base de datos. T es el umbral de la puntuación de la palabra vecina (Altschul y col., citado anteriormente). Estas coincidencias de la palabra vecina inicial actúan como semillas para iniciar las búsquedas para encontrar los HSP que las contienen. Las coincidencias con la palabra se extienden en ambas direcciones a lo largo de cada secuencia todo lo que la puntuación de la alineación acumulada se pueda incrementar. Las puntuaciones acumuladas se calculan usando, para secuencias nucleotídicas, los parámetros M (puntuación de recompensa por un par de residuos equivalentes; siempre > 0) y N (puntuación de penalización para residuos no equivalentes; siempre < 0). Para las secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulada. La extensión para las coincidencias de la palabra en cada dirección se detienen cuando: la puntuación de la alineación acumulada se salga de la cantidad X a partir de su valor máximo alcanzado; la puntuación acumulada llega a cero o menor, debido a la acumulación de una o más alineaciones de residuos que puntúan negativo; o se alcanza el final de cualquiera de las secuencias. Los parámetros del algoritmo BLAST W , T y X determinan la sensibilidad y la velocidad de la alineación. El programa BLASTN (para las secuencias de nucleótidos) usad como defecto una longitud de texto (W) de 11, una expectativa (E) de 10, $M=5$, $N=-4$ y una comparación de ambas hebras. Para las secuencias de aminoácidos, el programa BLAST usa como valores predeterminados una longitud de palabra de 3 y una expectativa (E) de 10 y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, *Proc. Nat'l. Acad. Sci. USA* 89:10915 (1989)) alineaciones (B) de 50, expectativa (E) de 10, $M= 5$, $N= -4$, y una comparación de ambas hebras.

El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). Una medida de la similitud proporcionada por el algoritmo BLAST es la menor probabilidad de la suma ($P(N)$), que proporciona una indicación de la probabilidad de que se produzca al azar una coincidencia entre dos secuencias de nucleótidos o de aminoácidos. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la menor probabilidad de la suma en una comparación del ácido nucleico problema con el ácido nucleico de referencia es inferior a aproximadamente 0,2, más preferentemente inferior a aproximadamente 0,01 y, de la forma más preferente inferior a aproximadamente 0,001.

La presente invención se extiende también a polinucleótidos que comprenden una primera secuencia de nucleótidos que se hibrida de forma selectiva en condiciones moderadamente rigurosas (tales como en condiciones muy rigurosas) con la complementaria de una segunda secuencia de nucleótidos que codifica un polipéptido que comprende:

- 5 (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
- (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7.

La frase "condiciones de hibridación altamente rigurosas" se refiere a condiciones bajo las cuales una sonda hibridará con su subsecuencia diana, típicamente en una mezcla compleja de ácido nucleico, pero no con otras secuencias. Las condiciones altamente rigurosas dependen de la secuencia y serán diferentes en circunstancias diferentes. Las secuencias más largas hibridan específicamente a temperaturas más altas. En Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Probes, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993) se puede encontrar una extensa guía de la hibridación de ácidos nucleicos. En general, las condiciones altamente estrictas se seleccionan de modo que sean aproximadamente 5-10 °C menos que el punto de fusión térmica (T_f) para la secuencia específica a un pH y fuerza iónica definidos. La T_f es la temperatura (a la fuerza iónica, el pH y la concentración de ácido nucleico definidas) a la cual el 50% de las sondas complementarias a la diana hibridan con la secuencia diana en el equilibrio (dado que las secuencias diana están presentes en exceso, a la T_f, el 50 % de las sondas están ocupadas en el equilibrio). Las condiciones altamente rigurosas serán aquellas en las que la concentración de sal es menor que aproximadamente 1,0 M de ion sodio, normalmente una concentración del ion sodio (u otras sales) de aproximadamente 0,01 a 1,0 M a un pH de 7,0 a 8,3 y la temperatura es de al menos aproximadamente 30 °C para sondas cortas (por ejemplo, de 10 a 50 nucleótidos) y al menos aproximadamente 60 °C para sondas largas (por ejemplo, de más de 50 nucleótidos). También se pueden alcanzar condiciones altamente rigurosas con la adición de agentes desestabilizantes, tales como formamida. Para una hibridación selectiva o específica, una señal positiva es de al menos dos veces la inicial, opcionalmente 10 veces la hibridación inicial.

Ejemplos de condiciones altamente rigurosas de hibridación pueden ser los siguientes: 50 % de formamida, 5x SSC y 1 % de SDS, incubando a 42 °C, o, 5x SSC, 1% de SDS, incubando a 65°C, con un lavado en 0,2x SSC y 0,1 % de SDS a 65 °C.

Los ácidos nucleicos que no hibridan entre sí bajo condiciones altamente rigurosas son todavía funcionalmente equivalentes si los polipéptidos que codifican son sustancialmente idénticos. Esto ocurre, por ejemplo, cuando se crea una copia de un ácido nucleico usando la máxima degeneración de codón permitida por el código genético. En tales casos, los ácidos nucleicos típicamente hibridan en condiciones de hibridación moderadamente rigurosas.

Entre los ejemplos de "condiciones de hibridación moderadamente rigurosas" se incluyen una hibridación en un tampón de formamida al 40 %, NaCl 1 M, 1 % de SDS a 37 °C, y un lavado en 1X SSC a 4 5°C. Una hibridación positiva es al menos dos veces la de fondo. Los expertos reconocerán fácilmente que se pueden utilizar condiciones alternativas de hibridación y lavado para proporcionar condiciones de rigurosidad similares.

La frase "hibrida de forma selectiva (o específicamente) con" se refiere a la unión, formación de dúplex o hibridación de una molécula solo con una secuencia nucleotídica concreta en condiciones de hibridación estrictas cuando dicha secuencia está presente en una mezcla compleja (incluidos, entre otros, ADN o ARN celular total o de biblioteca).

En cualquier caso, las variantes de una secuencia polipeptídica tendrán esencialmente la misma actividad que la secuencia de referencia (en el caso de polinucleótidos, las secuencias de polinucleótidos variantes codificarán un polipéptido que tiene esencialmente la misma actividad que la secuencia de referencia). Con esencialmente la misma actividad se quiere decir que al menos un 50 %, adecuadamente al menos un 75 % y, especialmente, al menos un 90 % de la actividad de la secuencia de referencia en un ensayo de reestimulación *in vitro* de PBMC o sangre entera con antígenos específicos (por ejemplo, reestimulación durante un periodo de entre varias horas hasta dos semanas, tal como hasta un día, de 1 día a 1 semana o de 1 a 2 semanas) que mide la activación de las células mediante linfoproliferación, producción de citocinas en el sobrenadante de cultivo (medido mediante ELISA, CBA etc.) o caracterización de las respuestas de linfocitos T o B mediante tinción intracelular y extracelular (por ejemplo, usando anticuerpos específicos de marcadores inmunitarios, tales como CD3, CD4, CD8, IL2, TNFa, IFNg, CD40L, CD69 etc.), seguido de análisis con un citómetro de flujo. De manera adecuada, por esencialmente la misma actividad se quiere decir al menos un 50 %, adecuadamente al menos un 75 % y especialmente al menos un 90 % de la actividad de la secuencia de referencia en un ensayo de proliferación de células T y/o de producción de IFN-gamma.

55 COMPOSICIONES DE POLINUCLEÓTIDOS

Tal como se usa en el presente documento, el término "polinucleótido" se refiere a una molécula que se ha aislado libre de ADN genómico total de una especie particular. Por lo tanto, un polinucleótido que codifica un polipéptido se refiere a un segmento de polinucleótido que contiene una o más secuencias de codificación aunque aislado

sustancialmente de, o purificado de, ADN genómico total de la especie de la que se obtiene el polinucleótido.

Como apreciarán los expertos en la materia, los polinucleótidos de la presente invención pueden incluir secuencias genómicas, secuencias extragenómicas y codificadas por plásmidos y segmentos de genes de ingeniería más pequeños que expresan, o pueden adaptarse para que expresen, proteínas, polipéptidos, péptidos y similares. Tales segmentos se pueden aislar de la naturaleza o se pueden modificar sintéticamente mediante la mano del hombre.

"Aislado", como se usa en el presente documento, significa que un polinucleótido está sustancialmente lejos de otras secuencias de codificación y que el polinucleótido no contiene porciones grandes de ADN codificante no relacionado, tales como fragmentos cromosómicos grandes u otros genes funcionales o regiones codificantes de polipéptidos. Un ácido nucleico aislado se separa de otros marcos de lectura abiertos que flanquean al gen y codifican una proteínas distintas a la del gen. Por supuesto, esto se refiere al segmento de ADN tal como se aisló originalmente, y no excluye genes o regiones codificantes añadidas más tarde al segmento por la mano del hombre.

Como apreciará el experto en la técnica, los polinucleótidos aislados pueden ser monocatenarios (de codificación o antisentido) o bicatenarios, y pueden ser moléculas de ADN (genómico, ADNc o sintético) o ARN. Las moléculas de ARN incluyen moléculas de ARNhn, que contienen intrones y corresponden a una molécula de ADN de una manera una a una y moléculas de ARNm, que no contienen intrones. Secuencias codificantes y no codificantes adicionales pueden estar presentes, aunque no necesariamente, dentro de un polinucleótido de la presente invención, y un polinucleótido puede estar unido a otras moléculas y/o materiales de soporte, aunque no necesariamente.

Los polinucleótidos pueden comprender una secuencia nativa (es decir, una secuencia endógena que codifica un antígeno de *Mycobacterium* o una porción del mismo) o pueden comprender una variante o un equivalente biológico o funcional de dicha secuencia. Las variantes polinucleotídicas pueden contener una o más sustituciones, adiciones, deleciones y/o inserciones, como se describe adicionalmente más adelante, preferentemente, de forma tal que la inmunogenicidad del polipéptido codificado no disminuye, en relación con la proteína de referencia. El efecto sobre la inmunogenicidad del polipéptido codificado puede evaluarse generalmente como se describe en el presente documento.

En las realizaciones adicionales, la presente invención proporciona polinucleótidos y polipéptidos aislados que comprenden diversas longitudes de tramos contiguos de secuencia idéntica o complementaria a una o más de las secuencias divulgadas en el presente documento. Por ejemplo, la presente invención proporciona polinucleótidos que comprenden al menos aproximadamente 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 o 1000 o más nucleótidos contiguos de la secuencia de referencia divulgada en el presente documento, así como todas las longitudes intermedias entre los mismos. Se entenderá fácilmente que "longitudes intermedias", en este contexto, significa cualquier longitud entre los valores citados, tal como 30, 31, 32, etc.; 50, 51, 52, 53, etc.; 100, 101, 102, 103, etc.; 150, 151, 152, 153, etc.; incluyendo todos los números enteros de 200-500; 500-1,000, y similares.

Por otra parte, los expertos en la técnica apreciarán que, como resultado de la degeneración del código genético, hay muchas secuencias de nucleótidos que codifican un polipéptido como se describe en el presente documento. Algunos de estos polinucleótidos llevan una identidad relativamente baja con la secuencia de nucleótidos de cualquier gen nativo. Sin embargo, los polinucleótidos que varían debido a diferencias en el uso de codones se contemplan específicamente en la presente invención, por ejemplo los polinucleótidos que están optimizados para la selección de codones humanos y/o de primates. Además, los alelos de los genes que comprenden las secuencias de polinucleótidos proporcionadas en el presente documento están dentro del alcance de la presente invención. Los alelos son genes endógenos que están alterados como resultado de una o más mutaciones, tales como deleciones, adiciones y/o sustituciones de nucleótidos. El ARNm resultante y la proteína pueden, aunque no necesariamente, tener una estructura o función alteradas. Los alelos pueden identificarse utilizando técnicas estándar (tales como hibridación, amplificación y/o comparación de secuencia de base de datos).

IDENTIFICACIÓN Y CARACTERIZACIÓN DE POLINUCLEÓTIDOS

Los polinucleótidos se pueden identificar, preparar y/o manipular usando cualquiera de diversas técnicas bien conocidas. Por ejemplo, un polinucleótido puede identificarse, como se describe con más detalle a continuación, mediante el cribado de una micromatriz de ADNc. Tales cribados se pueden realizar, por ejemplo, utilizando una micromatriz Synteni (Palo Alto, CA) de acuerdo con las instrucciones del fabricante (y esencialmente como describen Schena y col., Proc. Natl. Acad. Sci. USA 93:10614-10619 (1996) y Heller y col., Proc. Natl. Acad. Sci. USA 94:2150-2155 (1997)). Como alternativa, los polinucleótidos se pueden amplificar a partir de ADNc preparado de células que expresan las proteínas descritas en el presente documento, tales como células de *M. tuberculosis*. Dichos polinucleótidos se pueden amplificar mediante una reacción en cadena de la polimerasa (PCR). Para este abordaje, se pueden diseñar cebadores específicos de secuencia según las secuencias proporcionadas en el presente documento, y se pueden adquirir o sintetizar.

Se puede usar una porción amplificada de un polinucleótido par aislar un gen de longitud completa de una biblioteca adecuada (por ejemplo, una biblioteca de ADNc de *M. tuberculosis*) usando técnicas bien conocidas. En dichas técnicas se realiza detección selectiva de una biblioteca (ADNc o genómico) usando una o más sondas o cebadores polinucleotídicos adecuados para amplificación. Preferentemente, una biblioteca se selecciona por tamaño de modo

que incluya moléculas más grandes. También se pueden preferir bibliotecas cebadas aleatorias para identificar las regiones de los genes 5' y cadena arriba. Se prefieren las bibliotecas genómicas para obtener intrones y extender las secuencias en 5'.

5 Para las técnicas de hibridación, se puede marcar una secuencia parcial (por ejemplo, mediante traducción por muesca o marcaje del extremo libre con ³²P) usando técnicas bien conocidas. Generalmente se realiza la detección selectiva de una biblioteca bacteriana o de bacteriófagos mediante hibridación de filtros que contienen colonias bacterianas desnaturalizadas (o láminas que contienen placas de fagos) con la sonda marcada (véase Sambrook y col., *Molecular Cloning: A Laboratory Manual* (2000)). Las colonias o placas que hibridan se seleccionan y expanden, y el ADN se aísla para su posterior análisis. Los clones de ADNc se pueden analizar para determinar la cantidad de secuencia adicional mediante, por ejemplo, por ejemplo, PCR usando un cebador a partir de la secuencia parcial y un cebador del vector. Se pueden generar mapas de restricción y secuencias parciales para identificar uno o más clones solapantes. Después, la secuencia completa se puede determinar usando técnicas estándar, que pueden implicar generar una serie de clones de delección. Después, las secuencias solapantes resultantes se pueden ensamblar en una única secuencia contigua. Se puede generar una molécula de ADNc de longitud completa uniendo fragmentos adecuados a través de técnicas bien conocidas.

20 Como alternativa, existen numerosas técnicas de amplificación para obtener una secuencia de codificación de longitud completa a partir de una secuencia parcial de ADNc. En dichas técnicas, la amplificación generalmente se realiza mediante PCR. Se puede usar cualquiera de varios kits disponibles comercialmente para realizar la etapa de amplificación. Los se pueden diseñar usando, por ejemplo, software bien conocido en la técnica. Los cebadores tienen, preferentemente, 22-30 nucleótidos de longitud, tienen un contenido de GC de al menos 50 % e hibridan con la secuencia diana a temperaturas de aproximadamente 68 °C a 72 °C. La región amplificada se puede secuenciar como se ha descrito anteriormente y las secuencias solapantes se ensamblan en una secuencia contigua.

25 Una de estas técnicas de amplificación es la PCR inversa (véase Triglia y col., *Nucl. Acids Res.* 16:8186 (1988)), que usa enzimas de restricción para generar un fragmento en la región conocida del gen. A continuación, el fragmento se circulariza mediante ligadura intramolecular y se usa como molde para la PCR con cebadores divergentes derivados de la región conocida. En un enfoque alternativo, las secuencias adyacentes a una secuencia parcial se pueden recuperar mediante amplificación con un cebador a una secuencia enlazadora y un cebador específico a una región conocida. Las secuencias amplificadas normalmente se someten a una segunda ronda de amplificación con el mismo cebador de unión y un segundo cebador específico de la región conocida. Una variación de este procedimiento, que usa dos cebadores que inician la extensión en direcciones opuestas desde la secuencia conocida, se describe en el documento WO 96/38591. Otra de estas técnicas se conoce como "amplificación rápida de los extremos del ADNc" o RACE. Esta técnica implica el uso de un cebador interno y un cebador externo, que hibridan con una región poliA o secuencia de vector, para identificar las secuencias que están en 5' y 3' de una secuencia conocida. Técnicas adicionales incluyen PCR de captura (Lagerstrom y col., *PCR Methods Applic.* 1:111-19 (1991)) y *walking PCR* (Parker y col., *Nucl. Acids. Res.* 19:3055-60(1991)). Otros procedimientos que usan amplificación también se pueden usar para obtener una secuencia de ADNc de longitud completa.

35 En ciertos casos, es posible obtener una secuencia de ADNc de longitud completa mediante análisis de las secuencias proporcionadas en una base de dato de marcadores de secuencia (EST), tales como las disponibles en GenBank. Las búsquedas de EST solapantes pueden realizarse, en general, usando programas bien conocidos (p. ej., búsquedas NCBI BLAST) y dichos EST se pueden usar para generar una secuencia contigua de longitud completa. Las secuencias de ADN de longitud completa también se pueden obtener mediante análisis de los fragmentos genómicos.

EXPRESIÓN DE POLINUCLEÓTIDOS EN LAS CÉLULAS HUÉSPED

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

50 Como apreciarán los expertos en la materia, en algunos casos puede ser ventajoso producir secuencias de nucleótidos que codifiquen polipéptidos que posean codones de origen no natural. Por ejemplo, se pueden seleccionar codones preferidos por un huésped procariontico o eucariótica concreto para incrementar el índice de expresión de proteínas o producir un transcrito de ARN recombinante que posee propiedades deseables tales como una semivida mayor que la de un transcrito generado a partir de la secuencia de origen natural.

55 Por otra parte, las secuencias polinucleotídicas se pueden someter a ingeniería usando procedimientos generalmente conocidos en la técnica para alterar las secuencias que codifican polipéptidos por diversas razones, incluidas, entre otras, alteraciones que modifican la clonación, procesamiento y/o expresión del producto génico. Por ejemplo, para realizar ingeniería de las secuencias nucleotídicas se puede usar arrastre de ADN mediante fragmentación aleatoria y reensamblaje por PCR de los fragmentos génicos y oligonucleótidos sintéticos. Además, se puede usar mutagénesis dirigida por sitio para insertar nuevos puntos de restricción, alterar patrones de

glucosilación, cambiar la preferencia de codones, producir variantes de corte y empalme o introducir mutaciones, etc.

Las secuencias de ácido nucleico naturales, modificadas o recombinantes se pueden unir a una secuencia heteróloga para codificar una proteína de fusión. Por ejemplo, la detección selectiva en bibliotecas peptídicas de inhibidores de la actividad polipeptídica puede ser útil codificar una proteína quimérica que pueda ser reconocida por un anticuerpo disponible comercialmente. Una proteína de fusión también se puede someter a ingeniería para contener un sitio de escisión localizado entre la secuencia que codifica el polipéptido y la secuencia de la proteína heteróloga, de modo que el polipéptido se puede escindir y purificar del resto heterólogo.

Las secuencias que codifican un polipéptido deseado se pueden sintetizar, completamente o en parte, usando procedimientos químicos bien conocidos en la técnica (véase, Caruthers, M. H. y col., Nucl. Acids Res. Symp. Ser. pág. 215-223 (1980), Horn y col., Nucl. Acids Res. Symp. Ser. pág. 225-232 (1980)). Como alternativa, la propia proteína se puede producir usando procedimientos químicos para sintetizar la secuencia de aminoácidos de un polipéptido, o una porción de la misma. Por ejemplo, la síntesis peptídica se puede realizar usando varias técnicas de fase sólida (Roberge y col., Science 269:202-204 (1995)) y la síntesis automática se puede realizar, por ejemplo, usando el sintetizador de péptidos ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

Un péptido recién sintetizado se puede purificar sustancialmente mediante cromatografía preparativa líquida de alto rendimiento (por ejemplo, Creighton, *Proteins, Structures and Molecular Principles* (1983)) u otras técnicas comparables disponibles en la materia. La composición de los péptidos sintéticos se puede confirmar mediante análisis o secuenciación de los aminoácidos (por ejemplo, el procedimiento de degradación de Edman). Adicionalmente, durante la síntesis directa, la secuencia de aminoácidos de un polipéptido, o cualquier parte del mismo, se puede alterar y/o combinar mediante procedimientos químicos con otras secuencias de otras proteínas, o cualquier parte de las mismas, para producir una variante del polipéptido.

Con el fin de expresar un polipéptido deseado, las secuencias de nucleótidos que codifican el polipéptido, o equivalentes funcionales, se pueden insertar en un vector de expresión adecuado, es decir un vector que contenga los elementos necesarios para la transcripción y traducción de la secuencia de codificación insertada. Se pueden usar procedimientos bien conocidos por los expertos en la técnica para construir vectores de expresión que contengan secuencias que codifican un polipéptido de interés y los elementos adecuados de control de la transcripción y la traducción. Estos procedimientos incluyen técnicas de ADN recombinante in vitro, técnicas sintéticas y recombinación genética *in vivo*. Dichas técnicas se describen en Sambrook y col., *Molecular Cloning, A Laboratory Manual* (2000) y Ausubel y col., *Current Protocols in Molecular Biology* (actualizada anualmente).

Se pueden usar varios sistemas de expresión en vector/huésped para contener y expresar secuencias de polinucleótidos. Estas incluyen, aunque sin limitación, microorganismos, tales como bacterias transformadas con bacteriófagos recombinantes, plásmidos, o vectores de expresión de ADN en cósmidos; levadura transformada con vectores de expresión de levadura; sistemas de células de insectos infectadas con vectores de expresión de virus (por ejemplo, baculovirus); sistemas de células vegetales transformadas con vectores de expresión de virus (por ejemplo, virus del mosaico de la coliflor, CaMV; virus del mosaico del tabaco, TMV) o con vector de expresión bacteriana (por ejemplo, plásmidos Ti, o pBR322); o sistemas de células animales.

Los "elementos control" o "secuencias reguladoras" presentes en un vector de expresión son las regiones no traducidas del vector, potenciadores, promotores y regiones no traducidas en 5' y 3', que interactúan con las proteínas de las células huésped para llevar a cabo la transcripción y la traducción. Tales elementos pueden variar de fuerza y especificidad. Dependiendo del sistema de vector y de la célula huésped utilizados, se puede usar cualquier número de elementos de transcripción y de traducción adecuados, incluidos los promotores constitutivos e inducibles. Se pueden usar. Por ejemplo, al clonar en sistemas bacterianos, se pueden usar promotores inducibles como el promotor híbrido LacZ del fagemido PBLUESCRIPT (Stratagene, LaJolla, Calif.) o el plásmido PSPORT1 (Gibco BRL, Gaithersburg, MD) y similares. En los sistemas de células de mamífero, generalmente se prefieren los promotores de genes de mamíferos o de virus de mamíferos. Si es necesario generar una línea celular que contenga múltiples copias de la secuencia de nucleótidos que codifica un polipéptido, los vectores basados en los virus SV40 o EBV se pueden usar de forma ventajosa con un marcador seleccionable adecuado.

En los sistemas bacterianos, se pueden seleccionar numerosos vectores de expresión en función del uso previsto para el polipéptido expresado. Por ejemplo, cuando se necesitan grandes cantidades para, por ejemplo, la inducción de anticuerpos, se pueden usar vectores que dirijan la expresión de altos niveles de proteínas de fusión que se purifican con facilidad. Dichos vectores incluyen, aunque sin limitación, los vectores de expresión y clonación de *E. coli*, tales como BLUESCRIPT (Stratagene), en los que la secuencia que codifica el polipéptido de interés se puede ligar en el vector en marco con secuencias para la Met en el extremo amino y los 7 restos posteriores de β -galactosidasa de modo que se produce una proteína híbrida; vectores pIN (Van Heeke y Schuster, J. Biol. Chem. 264:5503-5509(1989)); y similares. Vectores pGEX (Promega, Madison, también se pueden usar para expresar polipéptidos extraños como proteínas de fusión con glutatión-S-transferasa (GST). En general, tales proteínas de fusión son solubles y pueden purificarse con facilidad a partir de células lisadas mediante adsorción en perlas de glutatión-agarosa, seguida por elución en presencia de glutatión libre. Las proteínas hechas en tales sistemas se pueden diseñar para que incluyan heparina, trombina o sitios de escisión por la proteasa del factor XA de modo que el polipéptido clonado de interés se pueda liberar de la fracción GST a voluntad.

En levaduras, *Saccharomyces cerevisiae*, se puede usar una serie de vectores que contienen promotores constitutivos o inducibles, tales como el factor alfa, la alcohol oxidasa y PGH. Otros vectores que contienen promotores constitutivos e inducibles incluyen GAP, PGK, GAL y ADH. Para revisiones, véase Ausubel y col., (citado anteriormente) y Grant y col., *Methods Enzymol.* 153:516-544 (1987) y Romas y col., *Yeast* 8 423-88 (1992).

- 5 En los casos en los que se usan vectores de expresión en plantas, la expresión de las secuencias que codifican polipéptidos puede estar dirigida por cualquiera de una serie de promotores. Por ejemplo, los promotores vitales tales como los promotores 35S y 19S del CaMV pueden usarse solos o en combinación con la secuencia líder omega del TMV (Takamatsu, *EMBO J.* 6:307-311 (1987)). Como alternativa, se pueden usar promotores vegetales, como los promotores de la subunidad pequeña de RIBISCO o del shock térmico (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 células vegetales mediante transformación directa de ADN o mediante transfección mediada por patógenos. Dichas técnicas se describen en una serie de recopilaciones disponibles generalmente (véase, por ejemplo, Hobbs en *McGraw Hill Yearbook of Science and Technology* pág. 191-196 (1992)).
- 10
- 15 También se puede usar un sistema de insectos para expresar un polipéptido de interés. Por ejemplo, en uno de estos sistemas, se usa el virus de la polihedrosis nuclear de *Autographa californica* (AcNPV) como vector para expresar genes extraños en células de *Spodoptera frugiperda* o de *Trichoplusia larvae*. Las secuencias que codifican el polipéptido se pueden clonar en una región no esencial del virus, tal como el gen de polihedrina, y se pueden colocar bajo el control del promotor de la polihedrina. La inserción satisfactoria de la secuencia que codifica el polipéptido inactivará el gen de polihedrina y producirá virus recombinante que carece de la proteína de recubrimiento. Los virus recombinantes se pueden usar para infectar células de, por ejemplo, *S. frugiperda* o *Trichoplusia larvae*, en las que se puede expresar el polipéptido de interés (Engelhard y col., *Proc. Natl. Acad. Sci. U.S.A.* 91:3224-3227 (1994)).
- 20

25 En células huésped de mamífero, generalmente hay disponibles numerosos sistemas de expresión basados en virus. Por ejemplo, en los casos en los que como vector de expresión se usa un adenovirus, las secuencias que codifican un polipéptido de interés se pueden ligar en un complejo de transcripción/traducción del adenovirus que consista en el promotor tardío y la secuencia líder tripartita. La inserción en una región E1 o E3 no esencial del genoma viral puede usarse para obtener un virus viable que sea capaz de expresar el polipéptido en las células huésped infectadas (Logan y Shenk, *Proc. Natl. Acad. Sci. U.S.A.* 81:3655-3659(1984)). Además, se pueden usar potenciadores de la transcripción, tal como el potenciador del virus del sarcoma de Rous (RSV), para incrementar la expresión en células huésped de mamífero. procedimientos y protocolos para trabajar con vectores de adenovirus se revisan en Wold, 1998. Se pueden encontrar referencias adicionales con respecto al uso de vectores de adenovirus en *Adenovirus: A Medical Dictionary, Bibliography, and Annotated Research Guide to Internet References*, 2004,

30

35 También se pueden usar señales de iniciación específicas para conseguir una traducción más eficaz de secuencias que codifican un polipéptido de interés. Tales señales incluyen el codón de iniciación ATG y las secuencias adyacentes. En los casos en las que las secuencias codificadoras del polipéptido, su codón de iniciación y sus secuencias anteriores se insertan en el vector de expresión adecuado, pueden no ser necesarias las señales adicionales de control de la transcripción o de la traducción. Sin embargo, en los casos en los que solo se inserta la secuencia codificadora, o un fragmento de la misma, se deberán proporcionar señales exógenas de control traduccional, incluido el codón de iniciación ATG. Por otro lado, el codón de iniciación deberá estar en el correcto marco de lectura para garantizar la traducción de todo el inserto. Los elementos exógenos traduccionales y los codones de iniciación pueden tener varios orígenes, tanto naturales como sintéticos. La eficiencia de la expresión se puede potenciar mediante la inclusión de potenciadores que son adecuados para el sistema celular concreto que se use, tales como los descritos en la literatura (Scharf. y col., *Results Probl. Cell Differ.* 20:125-162(1994)).

40

45 Además, una cepa de célula huésped se puede escoger por su capacidad para modular la expresión de las secuencias insertadas o para procesar la proteína expresada del modo deseado. Tales modificaciones del polipéptido incluyen, entre otras, acetilación, carboxilación, glicosilación, fosforilación, lipidación y acilación. El procesamiento postraduccional que escinde una forma "prepro" de la proteína también se puede usar para facilitar la inserción, plegamiento y/o función correctas. Diferentes células huésped, tales como CHO, HeLa, MDCK, HEK293 y WI38, tienen una maquinaria celular específica y mecanismos característicos para dichas actividades postraduccionales se pueden escoger para garantizar una correcta modificación y procesamiento de la proteína extraña.

50

55 Para la producción a largo plazo y con un rendimiento alto de proteínas recombinantes generalmente se prefiere la expresión estable. Por ejemplo, las líneas celulares que expresan de forma estable un polinucleótido de interés se pueden transformar usando vectores de expresión que pueden contener varios orígenes de replicación virales y/p elementos de expresión endógena y un gen de un marcador seleccionable en el mismo vector o en otro diferente. Tras la introducción del vector, se puede dejar que las células crezcan durante 1-2 días en un medio enriquecido antes de cambiarlas a un medio selectivo. El FIN del marcador seleccionable es conferir resistencia a la selección y su presencia permite el crecimiento y recuperación de células que expresan satisfactoriamente las secuencias introducidas. La proliferación de clones de células transformadas de forma estable se puede realizar usando técnicas de cultivo tisular adecuadas al tipo de célula.

60

Para recuperar las líneas celulares transformadas se puede usar cualquier número de sistemas de selección. Entre estos se incluyen, aunque sin limitación, los genes de la timidina cinasa del virus del herpes simple (Wigler y col., Cell 11:223-32 (1977)) y de la adenina fosforibosiltransferasa (Lowy y col., Cell 22:817-23 (1990)), que se pueden emplear en células tk.sup.- o aprt.sup.-, respectivamente. Asimismo, se pueden usar la resistencia antimetabolitos, a antibióticos o a herbicidas como base de la selección; por ejemplo, dhfr que confiere resistencia al metotrexato (Wigler y col., Proc. Natl. Acad. Sci. U.S.A. 77:3567-70 (1980)); npt, que confiere resistencia a los aminoglucósidos, neomicina y G-418 (Colbere-Garapin y col., J. Mol. Biol. 150:1-14 (1981)); y als o pat, que confiere resistencia al clorosulfurón y a la fosfotricin acetiltransferasa, respectivamente (Murry, citado anteriormente). Se han descrito genes seleccionables adicionales, por ejemplo, trpB, que permite a las células utilizar indol en lugar de triptófano, o hisD, que permite a las células utilizar histinol en lugar de histidina (Hartman y Mulligan, Proc. Natl. Acad. Sci. U.S.A. 85:8047-51(1988)). Recientemente, el uso de marcadores visibles ha ganado popularidad con marcadores tales como las antocianinas, -glucuronidasa y su sustrato GUS, y luciferasa y su sustrato luciferina, siendo ampliamente usados no solo para identificar los transformantes sino también para cuantificar la cantidad de expresión de proteína transitoria o estable atribuible a un sistema de vector específico (Rhodes y col., Methods Mol. Biol. 55:121-131(1995)).

Aunque la presencia/ausencia de expresión de un gen marcador sugiere que también está presente el gen de interés, puede ser necesario confirmar su presencia y expresión. Por ejemplo, si la secuencia que codifica un polipéptido se inserta dentro de una secuencia del gen marcador, las células recombinantes que contienen secuencias se pueden identificar por la ausencia de la función del gen marcador. Como alternativa, se puede introducir un gen marcador en tándem con una secuencia que codifique un polipéptido bajo el control de un promotor sencillo. La expresión del gen marcador como respuesta a la inducción o selección normalmente indica la expresión del gen tándem también.

Como alternativa, las células huésped que contienen y expresan una secuencia polinucleotídica deseada se pueden identificar mediante diversos procedimientos conocidos por los expertos en la materia. Estos procedimientos incluyen, aunque sin limitación, hibridaciones ADN-ADN o ADN-ARN y bioensayo de proteínas o técnicas de inmunoensayo que incluyen membrana, solución o chip de tecnologías basadas para la detección y/o cuantificación de ácido nucleico o proteína.

En la materia se conocen diversos protocolos para detectar y medir la expresión de productos codificados en polinucleótidos usando anticuerpos policlonales o monoclonales específicos del producto. Entre los ejemplos se incluyen ensayo de inmunosorción ligado a enzimas (ELISA), radioinmunoensayo (RIA) y clasificación celular activada por fluorescencia (FACS). Para algunas aplicaciones se puede preferir un inmunoensayo de dos sitios basado en anticuerpos monoclonales usando anticuerpos monoclonales reactivos a dos epítopos no interferentes sobre un polipéptido dado, pero también se puede emplear un ensayo de unión competitiva. Estos y otros ensayos se describen, entre otros sitios, 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 técnica conocen una amplia variedad de marcadores y técnicas de conjugación, y se pueden usar en varios ensayos con ácido nucleico y aminoácidos. Entre los medios para producir sondas de hibridación o de PCR marcadas para detectar secuencias relacionadas con polinucleótidos se incluyen oligomarcaje, traducción in mella, marcaje del extremo libre o amplificación por PCR usando un nucleótido marcado. Como alternativa, se pueden clonar en un vector las secuencias o cualquier porción de las mismas secuencias para la producción de una sonda de ARNm. Dichos vectores se conocen en la técnica y están disponibles comercialmente y se pueden usar para sintetizar sondas de ARN in vitro mediante la adición una ARN polimerasa adecuada, tal como T7, T3 o SP6 y nucleótidos marcados. Estos procedimientos se pueden realizar usando diversos de kits disponibles comercialmente. Entre las moléculas indicadoras o marcadores adecuados se incluyen radionúclidos, enzimas, agentes fluorescentes, quimioluminiscentes o cromogénicos, así como sustratos, cofactores, inhibidores, partículas magnéticas y similares.

Las células huésped transformadas con una secuencia polinucleotídica de interés se pueden cultivar en condiciones adecuadas para la expresión y recuperación de la proteína a partir del cultivo celular. La proteína producida mediante una célula recombinante puede secretarse o estar contenida en el interior de la célula dependiendo de la secuencia y/o el vector usado. Como apreciarán los expertos en la materia, se pueden diseñar vectores de expresión que contienen polinucleótidos de modo que contengan secuencias señal que dirigen la secreción del polipéptido codificado a través de una membrana celular procariótica o eucariótica. Se pueden usar otras construcciones recombinantes para unir las secuencias que codifican un polipéptido de interés a una secuencia nucleotídica que codifica un dominio polipeptídico que facilitará la purificación de las proteínas solubles. Tales dominios que facilitan purificación incluyen, entre otros, péptidos quelantes de metales tales como módulos de histidina-triptófano que permiten la purificación en metales inmovilizados, dominios de proteína A que permiten la purificación en inmunoglobulina inmovilizada y el dominio usado en el sistema de purificación de extensión/afinidad FLAGS (Immunex Corp., Seattle, Wash.). La inclusión de secuencias enlazadoras que se pueden escindir tales como las específicas del factor XA o enteroquinas (San Diego, Calif.) entre el dominio de purificación y el polipéptido codificado se puede usar 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 restos de histidina que preceden a la tioredoxina o un punto de escisión para enterocinas. Los restos de histidina facilitan la

purificación mediante IMAC (cromatografía de afinidad por iones metálicos inmovilizados) como se describe en Porath y col., Prot. Exp. Purif. 3:263-281 (1992), mientras que el sitio de escisión de la enterocinasa proporciona un medio de purificación del polipéptido deseado a partir de la proteína de fusión. Una discusión sobre vectores que contienen proteínas de fusión se proporciona en Kroll y col., DNA Cell Biol. 12:441-453(1993)).

5 TÉCNICAS DE LIBERACIÓN 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 células *in vivo*. Esto puede conseguirse usando cualquiera de diversos enfoques bien conocidos, varios de los cuales se describen a continuación con fines de ilustración.

1. ADENOVIRUS

10 Uno de los procedimientos preferidos para la liberación *in vivo* de una o más secuencias de ácido nucleico implica el uso de un vector de expresión adenovirus. Por "vector de expresión adenovirus" se entiende que incluye las construcciones que contienen secuencias de adenovirus suficientes para (a) soportar el empaquetamiento de la construcción y (b) expresar un polinucleótido que se ha clonado en el mismo en una orientación sentido o
15 antisentido. Por supuesto, en el contexto de una construcción antisentido, la expresión no requiere que el producto génico se sintetice.

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

25 El adenovirus es particularmente adecuado para su uso como vector de transferencia génica debido a su genoma de tamaño medio, facilidad de manipulación, título elevado, gama amplia de células diana y alta infectividad. Ambos extremos del genoma viral contienen repeticiones 100-200 pares de bases invertidas (ITR), que son elementos *cis* necesarios para la replicación del ADN viral y el empaquetamiento. Las regiones temprana (E) y tardía (L) del genoma contienen diferentes unidades de transcripción que están divididas por el inicio de la replicación del ADN viral. La región E1 (E1A y E1B) codifica proteínas responsables de la regulación de la transcripción del genoma viral
30 y unos pocos genes celulares. La expresión de la región E2 (E2A y E2B) da como resultado la síntesis de las proteínas para la replicación del ADN del virus. Estas proteínas están implicadas en la replicación del ADN, la expresión tardía del gen y la inactivación de la célula huésped (Renan, 1990). Los productos de los genes tardíos, incluyendo la mayoría de las proteínas de la cápside viral, se expresan solo después de un procesamiento significativo de un único transcrito primario producido por el promotor tardío principal (MLP). El MLP, (localizado en 16,8 mu) es particularmente eficaz durante la fase tardía de infección y todo el ARNm producido partir de este promotor posee una secuencia líder 5'-tripartita (TPL) que los convierte en ARNm preferidos para la traducción

35 En un sistema actual, el adenovirus recombinante se genera a partir de recombinación homóloga entre el vector lanzadera y el vector provirus. Debido a la posible recombinación entre dos vectores provirales, el adenovirus de tipo salvaje puede generarse a partir de este proceso. Por lo tanto, es crucial aislar un único clon del virus a partir de una placa individual y examinar su estructura genómica.

La generación y propagación de los vectores adenovirus actuales, que tienen una replicación deficiente, dependen de una línea celular colaboradora única, denominada 293, que se transformó a partir de células de riñón embrionario humano por fragmentos de ADN Ad5 y expresa de forma constitutiva proteínas E1 (Graham y col., 1977). Puesto
45 que la región E3 es prescindible del genoma de adenovirus (Jones y Shenk, 1978), los vectores de adenovirus actuales, con la ayuda de células 293, llevan ADN extraño en cualquiera de las regiones E1, D3 o ambas (Graham y Prevec, 1991). En la naturaleza, el adenovirus puede empaquetar aproximadamente 105 % del genoma de tipo salvaje (Ghosh-Choudhury y col., 1987), proporcionando capacidad para aproximadamente 2 kB extra de ADN. Combinado con los aproximadamente 5,5 kb de ADN que se pueden reemplazar en las regiones E1 y E3, la capacidad máxima del vector de adenovirus actual está por debajo de 7,5 kB, o aproximadamente el 15 % de la
50 longitud total del vector. Más del 80 % del genoma viral de adenovirus permanece en la estructura del vector y es la fuente de la citotoxicidad transmitida por vectores. Asimismo, la deficiencia en la replicación de los virus con delección de E1 es incompleta. Por ejemplo, se han observado fugas de expresión génica viral con los vectores disponibles actualmente a altas multiplicidades de infección (MOI) (Mulligan, 1993).

55 Las líneas celulares auxiliares pueden derivar de células humanas, tales como células de riñón embrionario humano, células musculares, células hematopoyéticas u otras células mesenquimatosas o epiteliales embrionarias humanas. Como alternativa, las células auxiliares pueden derivar de las células de otras especies de mamíferos que son permisivas para adenovirus humano. Dichas células incluyen, por ejemplo, células Vero u otras células embrionarias

mesenquimatosas o epiteliales de mono. Tal como se ha afirmado anteriormente, la línea celular auxiliar preferida actualmente es 293.

Racher y col., (1995) han divulgado procedimientos mejorados para cultivar células 293 y propagar adenovirus. En un formato, los agregados celulares naturales se cultivan mediante la inoculación de células individuales en matraces de agitación siliconados de 1 litro (Techne, Cambridge, Reino Unido) que contiene 100-200 ml de medio. Después de agitar a 40 rpm, la viabilidad celular se estima con azul tripán. En otro formato, se usan microportadores de Fibra-Cel (Bibby Sterlin, Stone, Reino Unido) (5 g/l) del siguiente modo. Se añade un inóculo celular, resuspendido en 5 ml de medio, al portador (50 ml) en un matraz Erlenmeyer de 250 ml y se deja inmóvil, con agitación ocasional, durante de 1 a 4 horas. A continuación, el medio se sustituye con 50 ml de medio fresco y se inicia la agitación. Para la producción de virus, se deja que las células crezcan hasta aproximadamente un 80 % de confluencia, tiempo tras el cual se sustituye el medio (hasta un 25 % del volumen final) y se añade el adenovirus a un MOI de 0,05. Los cultivos se dejan inmóviles durante la noche, tras lo cual se aumenta el volumen a 100 % y se comienza la agitación durante otras 72 horas.

Aparte del requisito de que el vector de adenovirus sea defectuoso en la replicación, o al menos condicionalmente defectuoso, no se cree que la naturaleza del vector de adenovirus sea crucial para la práctica exitosa de la invención. El adenovirus puede ser de cualquiera de los 42 serotipos o subgrupos AF diferentes conocidos. El adenovirus de tipo 5 del subgrupo C es el material de partida preferente con el fin de obtener un vector adenovirus de replicación defectuosa condicional para su uso en la presente invención, puesto que el adenovirus de tipo 5 es un adenovirus humano sobre el que se conoce una gran cantidad de información bioquímica y genética, y que se ha utilizado históricamente para la mayoría de las construcciones que emplean adenovirus como vector.

Tal como se ha afirmado anteriormente, el vector típico es defectuoso en cuanto a la replicación y no tendrá una región E1 de adenovirus. Por lo tanto, será más conveniente introducir el polinucleótido que codifica el gen de interés en la posición de la que se han eliminado las secuencias codificantes de E1. Sin embargo, la posición de inserción de la construcción dentro de las secuencias de adenovirus no es crucial para la invención. El polinucleótido que codifica el gen de interés también puede insertarse en lugar de la región E3 delecionada en los vectores de reemplazo de E3 como describen Karlsson y col., (1986) o en la región E4 en la que una línea celular auxiliar o el virus auxiliar complementa 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 puede obtenerse en títulos altos, por ejemplo, 10^9 - 10^{11} unidades formadoras de placas por ml y son altamente infecciosos. El ciclo de vida del adenovirus no requiere integración en el genoma de la célula huésped. Los genes extraños liberados por vectores de adenovirus son episomales y, por lo tanto, tienen una genotoxicidad baja para las células huésped. No se han notificado efectos secundarios en estudios de vacunación con adenovirus de tipo silvestre (Couch y col., 1963; Top y col., 1971), demostrando su seguridad y potencial terapéutico como vectores de transferencia de genes *in vivo*.

Los vectores de adenovirus se han usado en la expresión de genes eucariotas (Levrero y col., 1991; Gomez-Foix y col., 1992) y en el desarrollo de vacunas (Grunhaus y Horwitz, 1992; Graham y Prevec, 1992). Recientemente, estudios en animales han sugerido que el adenovirus recombinante podría usarse para terapia génica (Stratford-Perricaudet y Perricaudet, 1991; Stratford-Perricaudet y col., 1990; Rich y col., 1993). Los estudios realizados en la administración de adenovirus recombinante a diferentes tejidos incluyen instilación tráqueal (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 pueden proceder de adenovirus humano. Como alternativa, pueden originarse a partir de adenovirus de otras especies, por ejemplo chimpancé, que puede tener la ventaja de que los vectores virales no son neutralizados por los anticuerpos contra adenovirus humanos circulantes en muchos sujetos humanos (véase, por ejemplo: Tassis N y col. Gene Therapy 2006 13:421-429).

Adenovirus type 35, que es relativamente poco frecuente y, por lo tanto, hay niveles bajos de inmunidad preexistente al propio vector, se ha utilizado como sistema de liberación en ciertas vacunas contra la tuberculosis que se están desarrollando (véase, por ejemplo, Radosevic y col. Infection and Immunity 2007 75(8):4105-4115). El adenovirus de tipo 35 también puede ser de particular valor en la presente invención como vector de liberación.

2. RETROVIRUSS

Los retrovirus son un grupo de virus de ARN de cadena sencilla caracterizados por una capacidad de convertir su ARN en ADN de doble cadena en células infectadas por un proceso de transcripción inversa (Coffin, 1990). A continuación, el ADN resultante se integra de forma estable en cromosomas celulares como un provirus y dirige la síntesis de proteínas virales. La integración da como resultado la retención de las secuencias génicas virales en la célula receptora y sus descendientes. El genoma retroviral contiene tres genes, gag, pol y env, que codifican las proteínas de la cápsida, la enzima polimerasa y los componentes de la cubierta, respectivamente. Una secuencia que se encuentra corriente arriba del gen gag contiene una señal para el empaquetamiento del genoma en viriones. Dos secuencias de repetición terminal larga (LTR) están presentes en los extremos 5' y 3' del genoma viral. Estos

contienen un promotor fuerte y secuencias potenciadoras y también requiere la integración en el genoma de la células huésped (Coffin, 1990).

Con el fin de construir un vector retroviral, un ácido nucleico que codifica una o más secuencias de oligonucleótidos o polinucleótidos de interés se inserta en el genoma viral en el lugar de ciertas secuencias virales para producir un virus que es de replicación defectuosa. Con el fin de producir viriones, se construye una línea celular de empaquetamiento que contiene los genes gag, pol y env pero sin las LTR y los componentes de empaquetamiento (Mann y col., 1983). Cuando un plásmido recombinante que contiene un ADNc, junto con las secuencias LTR y de empaquetamiento se introduce en esta línea celular (mediante precipitación con fosfato de calcio por ejemplo), la secuencia de empaquetamiento permite que el transcrito de ARN del plásmido recombinante se empaquete en partículas virales, que a continuación, se secretan en los medios de cultivo (Nicolas y Rubenstein, 1988; Temin, 1986; Mann y col., 1983). A continuación, el medio que contiene los retrovirus recombinantes se recoge, opcionalmente se concentra, y se utiliza para la transferencia génica. Los vectores retrovirales son capaces de infectar una amplia variedad de tipos de células. Sin embargo, la integración y la expresión estable requieren la división de las células huésped (Paskind y col., 1975).

Recientemente se ha desarrollado un enfoque novedoso diseñado para permitir la dirección específica de vectores de retrovirus basándose en la modificación química de un retrovirus mediante la adición química de restos de lactosa a la envuelta viral. Esta modificación podría permitir la infección específica de hepatocitos a través de los receptores de sialoglicoproteína.

Se diseñó un enfoque diferente para la orientación de los retrovirus recombinantes en el que se usaron anticuerpos biotinilados contra una proteína de la envuelta retroviral y contra un receptor celular específico. Los anticuerpos se acoplaron a través de los componentes de biotina usando estreptavidina (Roux y col., 1989). Utilizando anticuerpos contra los antígenos del complejo mayor de histocompatibilidad de clase I y de clase II, se demostró la infección de diversas células humanas que portaban estos antígenos de superficie con un virus ecotrópico *in vitro* (Roux y col., 1989).

3. VIRUS ADENOASOCIADOS

AAV (Ridgeway, 1988; Hermonat y Muzyczka, 1984) es un parvovirus descubierto como contaminación de las reservas adenovirales. Es un virus ubicuo (los anticuerpos están presentes en el 85 % de la población humana de Estados Unidos) que no se ha relacionado con ninguna enfermedad. También se clasifica como un dependovirus, debido a que su replicación depende de la presencia de un virus auxiliar, tal como adenovirus. Se han aislado cinco serotipos, de los cuales el AAV-2 es el mejor caracterizado. El AAV tiene un ADN lineal monocatenario que está encapsidado en las proteínas de la cápsida VP1, VP2 y VP3 para formar un virión icosaédrico de 20 a 24 nm de diámetro (Muzyczka y McLaughlin, 1988).

El ADN de AAV tiene aproximadamente 4,7 kilobases de longitud. Contiene dos marcos de lectura abiertos y está flanqueado por dos ITR. Hay dos genes principales en el genoma del AAV: *rep* y *cap*. El gen *rep* codifica proteínas responsables de replications virales, mientras que *cap* codifica la proteína de la cápsida VP1-3. Cada ITR forma una estructura en horquilla con 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 utilizar como vector con todas las secuencias de codificación virales eliminadas y sustituidas por el casete de genes para la liberación. Se han identificado tres promotores virales y se han llamado p5, p19 y p40, en función de su posición en el mapa. La transcripción a partir p5 y p19 da como resultado la producción de proteínas rep y la transcripción de p40 produce las proteínas de la cápsida (Hermonat y Muzyczka, 1984).

Hay varios factores que llevaron a los investigadores a estudiar la posibilidad de usar rAAV como vector de expresión. Uno de ellos es que los requisitos para la liberación de un gen para integrarse en el cromosoma del huésped son sorprendentemente pocos. Es necesario tener las ITR de 145 pb, que son solo el 6 % del genoma del AAV. Esto deja espacio en el vector de montar una inserción de ADN de 4,5 kb. Mientras que esta capacidad portadora puede impedir que el AAV libere genes grandes, se adapta ampliamente para liberar construcciones antisentido.

El AAV es también una buena opción de vehículos de liberación por su seguridad. Existe un mecanismo de rescate relativamente complicado: no solo se requieren adenovirus de tipo salvaje, sino también los genes para movilizar rAAV. De forma análoga, el AAV no es patógeno y no está asociado con ninguna enfermedad. La eliminación de las secuencias de codificación virales minimiza las reacciones inmunes a la expresión génica viral, y por lo tanto, rAAV no evoca una respuesta inflamatoria.

4. OTROS VECTORES VIRALES COMO CONSTRUCCIONES DE EXPRESIÓN

Otros vectores virales pueden emplearse como construcciones de expresión en la presente invención para la liberación de secuencias de oligonucleótidos o polinucleótidos en una célula huésped. Se pueden usar vectores derivados de virus tales como virus vaccinia (Ridgeway, 1988; Coupar y col., 1988), lentivirus, poliovirus y virus del herpes. También se puede esperar que sean útiles otros vectores derivados de virus de viruela, tal como el virus de la viruela aviar. Ofrecen varias características atractivas para diversas células de mamífero (Friedmann, 1989;

Ridgeway, 1988; Coupar y col., 1988; Horwich y col., 1990).

Con el reciente reconocimiento de virus de hepatitis B defectuosos, se adquirió nueva información sobre la relación estructura-función de diferentes secuencias virales. Los estudios *in vitro* mostraron que el virus podría conservar la capacidad de empaquetamiento dependiente de auxiliar y la transcripción inversa a pesar de la delección de hasta el 80 % de su genoma (Horwich y col., 1990). Esto sugirió que grandes porciones del genoma podrían reemplazarse con material genético extraño. El hepatotropismo y la persistencia (integración) eran propiedades particularmente atractivas para la transferencia génica dirigida al hígado. Chang y col., (1991) introdujeron el gen de la cloranfenicol acetiltransferasa (CAT) en el genoma del virus de la hepatitis B de pato en el lugar de las secuencias de codificación de la polimerasa, de superficie y de presuperficie. Se cotransfectó con el virus de tipo salvaje en una línea celular de hepatoma aviar. Se utilizaron medios de cultivo con títulos elevados de virus recombinante para infectar hepatocitos primarios de pato. Se detectó expresión del gen de CAT estable durante al menos 24 días después de la transfección (Chang y col., 1991).

Vectores "virales" adicionales incluyen partículas similares a virus (VLP) y fagos

5. VECTORES NO VIRALES

Con el fin de efectuar la expresión de las secuencias de oligonucleótidos o polinucleótidos de la presente invención, la construcción de expresión debe liberarse en una célula. Esta liberación se puede realizar *in vitro*, como en procedimientos de laboratorio para transformar líneas de células, o *in vivo* o *ex vivo*, como en el tratamiento de ciertos estados de enfermedad. Tal como se ha descrito anteriormente, un mecanismo preferido para la liberación es a través de la infección viral, en la que la construcción de expresión está encapsulada en una partícula viral infecciosa.

Una vez que la construcción de expresión se ha liberado en la célula, el ácido nucleico que codifica las secuencias de oligonucleótidos o polinucleótidos deseadas se pueden colocar y expresaron en diferentes sitios. En ciertas realizaciones, el ácido nucleico que codifica la construcción puede integrarse de forma estable en el genoma de la célula. Esta integración puede realizarse en una ubicación y orientación específicas mediante recombinación homóloga (sustitución génica) o puede integrarse en una ubicación inespecífica aleatoria (aumento génico). En otras realizaciones adicionales, el ácido nucleico puede mantenerse de forma estable en la célula como un segmento episómico separado de ADN. Dichos segmentos de ácido nucleico o "episomas" codifican secuencias suficientes para permitir el mantenimiento y replicación independiente o sincronizada con el ciclo de la célula huésped. El modo en que la construcción de la expresión se libera en una célula y donde permanece en la célula el ácido nucleico depende del tipo de construcción de expresión usada.

En determinadas realizaciones de la invención, la construcción de expresión que comprende una o más secuencias de oligonucleótidos o polinucleótidos puede consistir simplemente en ADN recombinante desnudo o plásmidos. La transferencia de la construcción puede llevarse a cabo, por ejemplo, mediante cualquier procedimiento que permeabiliza física o químicamente la membrana celular. Esto es particularmente aplicable para la transferencia *in vitro*, pero puede aplicarse a su uso *in vivo* también. Dubensky y col., (1984) inyectaron con éxito ADN de poliomavirus en forma de precipitados en fosfato de calcio en el hígado y el bazo de ratones adultos y recién nacidos que demuestran replicación viral activa e infección aguda. Benvenisty y Reshef (1986) también demostraron que la inyección intraperitoneal directa de plásmidos precipitados en fosfato cálcico da como resultado la expresión de los genes transfectados. Se ha observado que el ADN que codifica un gen de interés también se puede transferir de forma similar *in vivo* y expresar el producto génico.

Otra realización de la invención para transferir una construcción de expresión de ADN desnudo en células puede implicar el bombardeo de partículas. Este procedimiento depende de la capacidad de acelerar los microproyectiles recubiertos con ADN a una velocidad alta, que les permite perforar las membranas celulares y entrar en las células sin matarlas (Klein y col., 1987). Se han desarrollado varios dispositivos para acelerar partículas pequeñas. Uno de tales dispositivos se basa en una descarga de alto voltaje para generar una corriente eléctrica, que, a su vez, proporciona la fuerza motriz (Yang y col., 1990). Los microproyectiles usados han consistido en sustancias biológicamente inertes, tales como perlas de tungsteno o de oro.

Órganos seleccionados, incluyendo el hígado, la piel y el tejido muscular de ratas y ratones han sido bombardeados *in vivo* (Yang y col., 1990; Zelenin y col., 1991). Esto puede requerir la exposición quirúrgica del tejido o las células para eliminar cualquier tejido intermedio entre la pistola y el órgano diana, es decir, tratamiento *ex vivo*. De nuevo, el ADN que codifica un gen particular puede liberarse a través de este procedimiento y todavía estar incorporado.

Las bacterias también pueden utilizarse como un procedimiento de liberación (por ejemplo, listeria, véase el documento WO2004/11048) y, en particular, BCG.

COMPOSICIONES POLIPEPTÍDICAS

La presente invención, en otros aspectos, proporciona composiciones polipeptídicas.

En general, un polipéptido de la invención será un polipéptido aislado (es decir, separado de dichos componentes

con los que normalmente se puede encontrar en la naturaleza).

Por ejemplo, una proteína de origen natural está aislado en caso de que esté separado de algunos o todos los materiales coexistentes en el sistema natural. Preferentemente, dichos polipéptidos tienen una pureza de al menos aproximadamente un 90 %, más preferentemente de al menos aproximadamente un 95 % y, lo más preferentemente de al menos aproximadamente un 99 %. Se considera que un polinucleótido está aislado si, por ejemplo, se clona en un vector que no forme parte de su ambiente natural.

Los polipéptidos se pueden preparar usando cualquiera de una variedad de técnicas bien conocidas. Los polipéptidos recombinantes codificados por secuencias de ADN como se ha descrito anteriormente se pueden preparar fácilmente a partir de las secuencias de ADN utilizando cualquiera de diversos vectores de expresión conocidos por los expertos en la materia. La expresión puede lograrse en cualquier célula huésped apropiada que se ha transformado o transfectado con un vector de expresión que contiene una molécula de ADN que codifica un polipéptido recombinante. Las células huésped adecuadas incluyen células procariotas, levaduras y eucariotas superiores, tales como células de mamífero y células vegetales. Preferentemente, las células huésped empleadas son *E. coli*, levadura o una línea celular de mamífero tal como COS o CHO. Los sobrenadantes de sistemas de huésped/vector adecuados que secretan proteína o polipéptido recombinante en el medio de cultivo se pueden concentrar primero usando un filtro comercialmente disponible. Tras la concentración, el concentrado se puede aplicar a una matriz de purificación adecuada, tal como una matriz de afinidad o una resina de intercambio iónico. Finalmente, se pueden usar una o más etapas de HPLC de fase inversa para purificar adicionalmente un polipéptido recombinante.

Los polipéptidos de la invención, fragmentos inmunogénicos de los mismos y otras variantes que tienen menos de aproximadamente 100 aminoácidos y, generalmente, menos de aproximadamente 50 aminoácidos, también se pueden generar por medios sintéticos, usando técnicas bien conocidas por los expertos en la materia. Por ejemplo, tales polipéptidos pueden sintetizarse utilizando cualquiera de las técnicas en fase sólida disponibles comercialmente, tales como el procedimiento de síntesis en fase sólida de Merrifield, en el que se añaden secuencialmente aminoácidos a una cadena de aminoácidos en crecimiento. Véase Merrifield, J. Am. Chem. Soc. 85:2149-2146 (1963). El equipo para la síntesis automática de polipéptidos está disponible comercialmente en proveedores tales como Perkin Elmer/Applied BioSystems Division (Foster City, CA) y se pueden manejar de acuerdo con las instrucciones del fabricante.

Dentro de ciertas realizaciones específicas, un polipéptido puede ser una proteína de fusión que comprende múltiples polipéptidos como se describe en el presente documento, o que comprende al menos un polipéptido como se describe en el presente documento y una secuencia no relacionada, ejemplos de tales proteínas incluyen proteínas del tétanos, la tuberculosis y la hepatitis (véase, por ejemplo, Stoute y col., New Engl. J. Med. 336:86-91(1997)). Una pareja de fusión puede, por ejemplo, ayudar a proporcionar epítopos de linfocitos T colaboradores (una pareja de fusión inmunológica), preferentemente epítopos de linfocitos T colaboradores por seres humanos, o ayudar en la expresión de la proteína (un potenciador de la expresión) con rendimientos más elevados que la proteína recombinante nativa. Determinados compañeros de fusión preferidos son compañeros de fusión tanto inmunológicos como potenciadores de la expresión. Se pueden seleccionar otros compañeros de fusión para aumentar la solubilidad de la proteína o para permitir dirigir la proteína a los compartimentos intracelulares deseados. Parejas de fusión todavía adicionales incluyen marcadores de afinidad, que facilita la purificación de la proteína.

Las proteínas de fusión se pueden preparar generalmente usando técnicas estándar, incluida conjugación química. Preferentemente, la proteína de fusión se expresa como una proteína recombinante, lo que permite la producción de niveles incrementados, con respecto a una proteína no fusionada, en un sistema de expresión. En resumen, las secuencias de ADN que codifican los componentes polipeptídicos pueden ensamblarse por separado y ligarse en un vector de expresión adecuado. El extremo 3' de la secuencia de ADN que codifica un componente polipeptídico está ligado, con o sin un enlazador peptídico, al extremo 5' de una secuencia de ADN que codifica el segundo componente polipeptídico de modo que los marcos de lectura de las secuencias estén en fase. Esto permite la traducción en una única proteína de fusión que conserva la actividad biológica de ambos polipéptidos componentes.

Se puede usar una secuencia enlazadora peptídica para separar el primero y el segundo componentes polipeptídicos una distancia suficiente para garantizar que cada polipéptido se pliega en sus estructuras secundarias y terciarias. Dicha secuencia de enlazador peptídico se incorpora a la proteína de fusión usando técnicas convencionales bien conocidas en la materia. Las secuencias de enlazador peptídico adecuadas pueden seleccionarse 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 los polipéptidos primero y segundo; y (3) la ausencia de restos hidrófobos o cargados que pudieran reaccionar con los epítopos funcionales del polipéptido. Las secuencias de enlazador peptídico preferidas contienen restos de Gly, Asn y Ser. Otros aminoácidos prácticamente neutros, tales como Thr y Ala, también pueden usarse en la secuencia enlazadora. Las secuencias de aminoácidos que pueden ser útiles empleadas como enlazadores incluyen aquellas desveladas en Maratea y col., Gene 40:39-46 (1985); Murphy y col., Proc. Natl. Acad. Sci. USA 83:8258-8262 (1986); la patente de Estados Unidos n.º 4.935.233 y la Patente de Estados Unidos n.º 4.751.180. La secuencia enlazadora puede tener, generalmente, de 1 a aproximadamente 50 aminoácidos de longitud. No son

necesarias secuencias enlazadoras cuando los polipéptidos primero y segundo tienen regiones de aminoácidos N-terminales no esenciales que pueden usarse para separar los dominios funcionales e impedir la interferencia estérica.

5 En realizaciones preferentes, una pareja de fusión inmunológica deriva de la proteína D, una proteína de superficie de la bacteria gramnegativa *Haemophilus influenzae* B (documento WO 91/18926). Preferentemente, un derivado de proteína D comprende aproximadamente el primer tercio de la proteína (por ejemplo, los primeros 100-110 aminoácidos del extremo N) y un derivado de proteína D puede estar lipidado. En determinadas realizaciones preferentes, los primeros 109 restos de una pareja de fusión de lipoproteína D se incluye en el extremo N para proporcionar al polipéptido con epítomos de células T exógenas adicionales y para aumentar el nivel de expresión en *E. coli* (funcionando de este modo como un potenciador de la expresión). La cola lipídica asegura una presentación óptima del antígeno a las células presentadoras de antígeno. Otras parejas de fusión incluyen la proteína no estructural del virus de la gripe, NS1 (hemaglutinina). Típicamente, se usan los 81 aminoácidos en N-terminal, aunque se pueden usar fragmentos diferentes que incluyen epítomos de linfocitos T colaboradores.

15 En otra realización, la pareja de fusión inmunológica es la proteína conocida como LYTA o una porción de la misma (preferentemente, una porción en el extremo C). LYTA deriva de *Streptococcus pneumoniae*, que sintetiza una N-acetil-L-alanina amidasa conocida como amidasa LYTA (codificada por el gen de *LytA*; Gene 43:265-292 (1986)). LYTA es una autolisina que degrada específicamente ciertos enlaces en el esqueleto de peptidoglicano. El dominio C-terminal de la proteína LYTA es responsable de la afinidad por la colina o por algunos análogos de la colina tales como el DEAE. Esta propiedad se ha explotado para el desarrollo de plásmidos de expresión C-LyTA en *E. coli* útiles para la expresión de proteínas de fusión. Se ha descrito la purificación de proteínas híbridas que contienen el fragmento C-LYTA en su término amino (véase *Biotechnology* 10:795-798 (1992)). En una realización preferente, una porción de repetición de LYTA se puede incorporar en una proteína de fusión. Una porción repetida se encuentra en la región C-terminal comenzando en el resto 178. Una porción repetida particularmente preferida incorpora los restos 188-305.

25 LINFOCITOS T

Las composiciones inmunoterapéuticas pueden también, o como alternativa, comprender linfocitos T específicos para un antígeno de *Mycobacterium*. Tales células pueden prepararse generalmente *in vitro* o *ex vivo*, usando procedimientos estándar. Por ejemplo, los linfocitos T pueden aislarse de médula ósea, sangre periférica o una fracción de médula ósea o sangre periférica de un paciente, utilizando un sistema de separación de células disponible comercialmente, tal como el Sistema Isolex™, disponible de Nexell Therapeutics, Inc. (Irvine, CA; véase también la patente de Estados Unidos N.º 5.240.856; la patente de Estados Unidos n.º 5.215.926; el documento WO 89/06280; los documentos WO 91/16116 y WO 92/07243). Como alternativa, los linfocitos T pueden derivar de seres humanos relacionados o no relacionados, mamíferos no humanos, líneas celulares o cultivos.

35 Los linfocitos T pueden estimularse con un polipéptido de la invención, polinucleótido que codifica un polipéptido tal, y/o una célula presentadora de antígeno (CPA) que expresa un polipéptido de este tipo. Dicha estimulación se realiza en condiciones y durante un tiempo suficiente para permitir la generación de linfocitos T que son específicos para el polipéptido. Preferentemente, el polipéptido o polinucleótido está presente dentro de un vehículo de liberación, tal como una microesfera, para facilitar la generación de linfocitos T específicos.

40 Los linfocitos T se consideran específicos de un polipéptido de la invención si los linfocitos T proliferan específicamente, secretan citocinas o matan las células diana recubiertas con el polipéptido o que expresan un gen que codifica el polipéptido. La especificidad de los linfocitos T puede evaluarse usando cualquiera de una diversidad de técnicas estándar. Por ejemplo, dentro de un ensayo de liberación de cromo o ensayo de proliferación, un índice de estimulación de más de dos veces mayor en la lisis y/o proliferación, en comparación con controles negativos, indica especificidad de linfocitos T. Tales ensayos pueden realizarse, por ejemplo, como se describe en Chen y col., *Cancer Res.* 54:1065-1070(1994)). Como alternativa, la detección de la proliferación de linfocitos T puede conseguirse mediante diversas técnicas conocidas. Por ejemplo, la proliferación de linfocitos T puede detectarse midiendo una tasa aumentada de la síntesis de ADN (por ejemplo, mediante cultivos de marcaje en pulsos de linfocitos T con timidina tritiada y midiendo la cantidad de timidina tritiada incorporada en el ADN). El contacto con un polipéptido de la invención (100 ng/ml - 100 µg/ml, preferentemente 200 ng/ml - 25 µg/ml) durante 3 - 7 días debería dar como resultado al menos un aumento de dos veces en la proliferación de los linfocitos T. El contacto como se ha descrito anteriormente durante 2-3 horas debería dar como resultado la activación de los linfocitos T, medido usando ensayos de citocinas estándar en el que un aumento de dos veces el nivel de liberación de citocinas (por ejemplo, TNF o IFN-γ) es indicativo de activación de linfocitos T (véase Coligan y col., *Current Protocols in Immunology*, vol. 1 (1998)). Los linfocitos T que se han activado en respuesta a un polipéptido, polinucleótido o una CPA que expresa un polipéptido pueden ser CD4+ y/o CD8+. Los linfocitos T específicos de la proteína pueden expandirse usando técnicas estándar. En realizaciones preferentes, los linfocitos T derivan de un paciente, un donante relacionado o un donante no relacionado, y se administran al paciente después de la estimulación y la expansión.

60 Con fines terapéuticos, los linfocitos T CD4+ o CD8+ que proliferan en respuesta a un polipéptido, polinucleótido o CPA pueden expandirse en número *in vitro* o *in vivo*. La proliferación de tales linfocitos T *in vitro* puede conseguirse de diversas maneras. Por ejemplo, los linfocitos T pueden volver a exponerse a un polipéptido, o un péptido corto

que corresponde a una porción inmunogénica de un polipéptido tal, con o sin la adición de factores de crecimiento de linfocitos T, tales como la interleucina-2, y/o células estimulantes que sintetizan un polipéptido. Como alternativa, uno o más linfocitos T que proliferan en presencia de la proteína se pueden expandir en número por clonación. Los procedimientos para clonar células son bien conocidos en la técnica, e incluyen dilución limitante.

5 COMPOSICIONES FARMACÉUTICAS

En las realizaciones adicionales, las composiciones de polinucleótido, polipéptido, linfocitos T y/o anticuerpos divulgados en el presente documento se formularán en soluciones farmacéuticamente aceptables o fisiológicamente aceptable para la administración a una célula o un animal, ya sea solo, o en combinación con uno o más de otras modalidades de tratamiento.

10 También se entenderá que, si se desea, el segmento de ácido nucleico (por ejemplo, ARN o ADN) que expresa un polipéptido como se divulga en el presente documento puede administrarse en combinación con otros agentes también, tal como, por ejemplo, otras proteínas o polipéptidos o diversos agentes farmacéuticamente activos, incluyendo agentes quimioterapéuticos eficaces contra una infección por *M. tuberculosis*. De hecho, prácticamente ningún límite a otros componentes que también pueden incluirse, dado que los agentes adicionales no causan un efecto adverso significativo tras el contacto con las células diana o tejidos del huésped. Por tanto, las composiciones se pueden administrar junto con diversos otros agentes según se requiere en el caso particular. Dichas composiciones pueden purificarse a partir de células huésped u otras fuentes biológicas o, como alternativa, pueden sintetizarse químicamente como se describe en el presente documento. De forma análoga, dichas composiciones pueden comprender además composiciones de ARN o ADN sustituido o derivado.

20 La formulación de excipientes farmacéuticamente aceptables y soluciones de vehículo es bien conocida en la materia, al igual que el desarrollo de regímenes de dosificación y tratamiento adecuados para usar las composiciones particulares descritas en el presente documento en diversos regímenes de tratamiento, incluyendo, por ejemplo, administración oral, parenteral, intravenosa, intranasal e intramuscular y formulación. Otras vías de administración incluyen a través de las superficies mucosas.

25 Típicamente, las formulaciones que comprenden una cantidad terapéuticamente eficaz administran de aproximadamente 0,1 ug a aproximadamente 1.000 ug de polipéptido por administración, más típicamente de aproximadamente 2,5 ug a aproximadamente 100 ug de polipéptido por administración. Con respecto a las composiciones de polinucleótidos, estas típicamente administran de 10 ug a aproximadamente 20 mg del polinucleótido de la invención por administración, más típicamente de aproximadamente 0,1 mg a aproximadamente 10 mg del polinucleótido de la invención por administración.

30 Naturalmente, la cantidad de compuesto o compuestos activos en cada composición terapéuticamente útil puede prepararse de una manera tal que se obtenga en cualquier dosis unitaria dada del compuesto. El experto en la materia de preparar dichas formulaciones farmacéuticas contemplará factores tales como solubilidad, la biodisponibilidad, la semivida biológica, la vía de administración, la vida de útil del producto, así como otras consideraciones farmacológicas y, como tales, pueden ser deseables diversas dosificaciones y regímenes terapéuticos.

1. ADMINISTRACIÓN ORAL

35 En ciertas aplicaciones, las composiciones farmacéuticas divulgadas en el presente documento pueden administrarse mediante administración oral a un animal. Como tales, estas composiciones pueden formularse con un diluyente inerte o con un vehículo comestible asimilable, o se pueden encerrar en cápsulas de gelatina dura o blanda o pueden comprimirse en comprimidos o pueden incorporarse directamente con los alimentos de la dieta.

Los compuestos activos pueden incluso estar incorporados con excipientes y utilizarse en forma de comprimidos, comprimidos bucales, trociscos, cápsulas, elixires, suspensiones, jarabes, obleas y similares (Mathiowitz y col., 1997; Hwang y col., 1998; patente de Estados Unidos 5.641.515; patente de Estados Unidos N.º 5.580.579 y patente de Estados Unidos N.º 5.792.451, cada uno incorporado específicamente en el presente documento por referencia en su totalidad). Los comprimidos, trociscos, píldoras, cápsulas y similares pueden también contener los siguientes: un aglutinante, goma de tragacanto, goma arábiga, almidón de maíz o gelatina; excipientes, tales como fosfato dicálcico; un agente disgregante, tal como almidón de maíz, almidón de patata, ácido algínico y similares; un lubricante, tal como estearato de magnesio; y un agente edulcorante, tal como sacarosa, lactosa o sacarina o un agente aromatizante, tal como menta, aceite de gaulteria o aromatizante de cereza. Cuando la forma unitaria de dosificación es una cápsula, puede contener, además de materiales del tipo anterior, un vehículo líquido. Pueden estar presentes otros materiales diversos como revestimientos o para modificar de otro modo la forma física de la unidad de dosificación. Por ejemplo, los comprimidos, las píldoras o las cápsulas se pueden recubrir con goma laca, azúcar o ambos. Un jarabe o elixir puede contener el compuesto activo, sacarosa como agente edulcorante, metilo y propilparabenos como conservantes, un pigmento y un aromatizante, tal como sabor a cereza o a naranja. Por supuesto, cualquier material usado en la preparación de cualquier forma farmacéutica unitaria debe ser farmacéuticamente puros y sustancialmente no tóxico a las cantidades usadas. Además, los componentes activos se pueden incorporar en la preparación y formulaciones de liberación sostenida.

Para la administración oral, las composiciones de la presente invención, de forma alternativa, se pueden incorporar con uno o más excipientes en forma de un colutorio, dentífrico, comprimido bucal, pulverización oral o formulación sublingual administrada por vía oral. Por ejemplo, se puede preparar un colutorio incorporando el ingrediente activo en la cantidad requerida en un disolvente apropiado, tal como una solución de borato de sodio (solución de Dobell).

5 Como alternativa, el ingrediente activo se puede incorporar en una solución oral, tal como una que contenga borato de sodio, glicerina y bicarbonato de potasio, o se dispersa en un dentífrico o añadir en una cantidad terapéuticamente eficaz a una composición que puede incluir agua, aglutinantes, abrasivos, agentes aromatizantes, agentes espumantes y humectantes. Como alternativa, las composiciones pueden conformarse en un comprimido o solución que se pueden colocar debajo de la lengua o disolver de otro modo en la boca.

10 2. ADMINISTRACIÓN INYECTABLE

En ciertas circunstancias será deseable administrar las composiciones farmacéuticas divulgadas en el presente documento por vía parenteral, por vía intravenosa, por vía intramuscular, por vía intradérmica, o incluso por vía intraperitoneal como se describe en la patente de Estados Unidos 5.543.158; la patente de Estados Unidos 5.641.515 y la patente de Estados Unidos 5.399.363 (cada una incorporada específicamente en el presente documento por referencia en su totalidad). Las soluciones de los compuestos activos como base libre o sales farmacológicamente aceptables pueden prepararse en agua mezclada adecuadamente con un tensioactivo, tal como hidroxipropilcelulosa. Las dispersiones también pueden prepararse en glicerol, polietilenglicoles líquidos, y mezclas de los mismos, en aceites. En condiciones normales de almacenamiento y uso, estas preparaciones contienen un conservante para evitar el crecimiento de microorganismos.

20 Las formas farmacéuticas adecuadas para su uso inyectable incluyen soluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de soluciones o dispersiones inyectables estériles (patente de Estados Unidos 5.466.468, incorporada específicamente en el presente documento por referencia en su totalidad). En todos los casos, la forma debe ser estéril y debe ser fluida de modo que se pueda introducir con facilidad en las jeringuillas. Debe ser estable en condiciones de fabricación y almacenamiento y tiene que conservarse contra la acción de microorganismos contaminantes, tales como bacterias y hongos. El vehículo puede ser un disolvente o medio de dispersión que contiene, 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 puede mantenerse, por ejemplo, mediante el uso de un recubrimiento tal como lecitina, mediante el mantenimiento del tamaño de partícula necesario y mediante el uso de tensioactivos. La prevención de la acción de microorganismos puede facilitarse mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorbutanol, fenol, ácido sórbico, timerosal, y similares. En muchos casos, será preferente incluir agentes isotónicos, por ejemplo, azúcares o cloruro sódico. La absorción prolongada de las composiciones inyectables se puede efectuar mediante el uso en las composiciones de agentes que retrasa la absorción, por ejemplo, monoestearato de aluminio y gelatina.

35 Para la administración parenteral en una solución acuosa, por ejemplo, la solución debe tamponarse de forma adecuada, si fuera necesario, y el diluyente líquido en primer lugar se hace isotónico con suficiente solución salina o glucosa. Estas soluciones acuosas particulares son especialmente adecuadas para administración intravenosa, intramuscular, subcutánea e intraperitoneal. A este respecto, a la luz de la presente divulgación, los expertos en la materia conocerán un medio acuoso estéril que se puede emplear. Por ejemplo, una dosificación puede disolverse en 1 ml de solución de NaCl isotónica y añadirse a 1000 ml de fluido de hipodermoclasia o inyectarse en el sitio propuesto de infusión (véase, por ejemplo, Remington's Pharmaceutical Sciences, 15ª edición, pág 1035/-1038 y 1570-1580). Necesariamente se producirá alguna variación en la dosis en función de la afección del sujeto que se esté tratando. La persona responsable de la administrará determinará, en cualquier caso, la dosis adecuada para el sujeto individual. Por otra parte, para la administración humana, las preparaciones deben satisfacer normas de esterilidad, pirogenicidad y de seguridad y pureza requeridos por la Oficina de la FDA de los patrones biológicos.

Las soluciones inyectables estériles se preparan incorporando compuestos activos en la cantidad requerida en el disolvente apropiado con diversos otros ingredientes como se ha enumerado anteriormente, según sea necesario, seguido de esterilización por filtración. En general, las dispersiones se preparan mediante la incorporación de los diversos ingredientes activos esterilizados en un vehículo estéril que contiene el medio de dispersión básico y los otros ingredientes requeridos de los enumerados en lo que antecede. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los procedimientos preferentes de preparación son técnicas de secado al vacío y liofilización, que dan un polvo del principio activo más cualquier ingrediente deseado adicional a partir de una solución del mismo previamente esterilizada mediante filtración.

55 Las composiciones divulgadas en el presente documento se pueden formular en una forma neutra o sal. Entre las sales farmacéuticamente aceptables se incluyen las sales de adición ácida (formadas con los grupos amino libre de la proteína) y que se forman con ácidos inorgánicos, tales como, por ejemplo, ácido clorhídrico o fosfórico, o dichos ácidos orgánicos como acético, oxálico, tartárico, mandélico o similares. Las sales formadas a partir de los grupos carboxilo libres también pueden derivar de bases inorgánicas, tales como, por ejemplo, los hidróxidos sódico, potásico, amónico, cálcico o férrico, y bases orgánicas tales como isopropilamina, trimetilamina, histidina, procaína y similares. Tras la formulación, las soluciones se pueden administrar de un modo compatible con la formulación de dosificación y una cantidad tal que sea terapéuticamente eficaz. Las formulaciones se administran con facilidad

mediante diversas formas de dosificación, tales como soluciones inyectables, cápsulas de liberación de fármacos y similares.

Tal como se usa en el presente documento, "vehículo" incluye todos y cada uno de los disolventes, medio de dispersión, vehículos, recubrimiento, diluyente, agentes antibacterianos y antifúngicos, agentes isotónicos y retardantes de la absorción, tampones, soluciones vehículo, suspensiones, coloides y similares. El uso de dichos medios y agentes para sustancias farmacéuticamente activas es bien conocido en la técnica. Excepto en la medida en que cualquier medio o agente convencional es incompatible con el principio activo, se contempla su uso en las composiciones terapéuticas. En las composiciones también pueden incorporarse principios activos suplementarios.

La frase "farmacéuticamente aceptable" se refiere a entidades moleculares y composiciones que no producen reacciones adversas, alérgicas o perjudiciales cuando se administran a un ser humano. La preparación de una composición acuosa que contiene una proteína como principio activo es bien entendida en la materia. Típicamente, dichas composiciones se preparan como inyectables, ya sea como soluciones líquidas o suspensiones; Las formas sólidas adecuadas para solución en, o suspensión en, líquido antes de la inyección también se pueden preparar. La preparación también puede emulsionarse.

3. ADMINISTRACIÓN NASAL Y BUCAL

En ciertas realizaciones, las composiciones farmacéuticas se pueden administrar mediante pulverizaciones intranasales, pulverizaciones bucales, inhalación, y/u otros vehículos de administración en aerosol. Los procedimientos para la liberación de genes, ácidos nucleicos, y composiciones peptídicas directamente en los pulmones por ejemplo, pulverizaciones en aerosol nasales y bucales se ha descrito, por ejemplo, en la patente de Estados Unidos 5.756.353 y la patente de Estados Unidos 5.804.212 (cada una incorporada específicamente en el presente documento por referencia en su totalidad). De forma análoga, 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, incorporada específicamente en el presente documento por referencia en su totalidad) también son bien conocidos en las técnicas farmacéuticas. De forma análoga, la administración de fármacos transmucosal en forma de una matriz de soporte de politetrafluoroetileno se describe en la patente de Estados Unidos 5.780.045 (incorporada específicamente en el presente documento por referencia en su totalidad).

4. ADMINISTRACIÓN MEDIADA POR LIPOSOMAS, NANOCÁPSULAS Y MICROPARTÍCULAS

En ciertas realizaciones, los inventores contemplan el uso de liposomas, nanocápsulas, micropartículas, microesferas, partículas lipídicas, vesículas y similares, para la introducción de las composiciones de la presente invención en células huésped adecuadas. En particular, las composiciones de la presente invención se pueden formular para la administración o bien encapsuladas en una partícula lipídica, un liposoma, una vesícula, una nanoesfera o una nanopartícula o similares.

Dichas formulaciones pueden preferirse para la introducción de formulaciones farmacéuticamente aceptables de los ácidos nucleicos o construcciones divulgadas en el presente documento. La formación y uso de liposomas es generalmente conocida por los expertos en la técnica (véase por ejemplo, Couvreur y col., 1977; Couvreur, 1988; Lasic, 1998; que describe el uso de liposomas y nanocápsulas en la terapia antibiótica dirigida a infecciones bacterianas intracelulares y enfermedades) Recientemente, se han desarrollado liposomas con mejor estabilidad en suero y mejores semividas en circulación (Gabizon y Papahadjopoulos, 1988; Allen y Choun, 1987; Patente de Estados Unidos 5.741.516. incorporada específicamente en el presente documento por referencia en su totalidad). Además, se han revisado diversos procedimientos de liposomas y preparaciones similares 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 N.º 5.738.868 y patente de Estados Unidos N.º 5.795.587, cada uno incorporado específicamente en el presente documento por referencia en su totalidad).

Los liposomas se han utilizado con éxito con un número de tipos celulares que normalmente son resistentes a la transfección mediante otros procedimientos, incluyendo suspensiones de linfocitos T, cultivos de hepatocitos primarios y células PC 12 (Renneisen y col., 1990; Muller y col., 1990). Además, los liposomas están libres de las restricciones de longitud del ADN que son típicas de sistemas de administración basados en virus. Los liposomas se han utilizado eficazmente 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 transcripción y efectores alostéricos (Nicolau y Gersonde, 1979) en diversas líneas celulares cultivadas y animales. Además, se han realizado varios ensayos clínicos exitosos que examinan la eficacia de la administración de fármacos mediada por liposomas (López-Berestein y col., 1985a; 1985b; Coune, 1988; Sculier y col., 1988). Por otro lado, varios estudios sugieren que el uso de liposomas no está asociado con respuestas autoinmunes, toxicidad o localización gonadal después de la administración sistémica (Mori y Fukatsu, 1992).

Los liposomas se forman a partir de fosfolípidos que se dispersan en un medio acuoso y forman espontáneamente vesículas bicapa concéntricas multilamelares (también denominadas vesículas multilamelares (VML). Las VML

generalmente tienen diámetros de 25 nm a 4 μ m. La aplicación de ultrasonidos a las VML da lugar a la formación de vesículas unilamelares pequeñas (VUP) con diámetros en el intervalo de 200 a 500 Å, que contienen una solución acuosa en el núcleo.

5 Los liposomas tienen semejanza con las membranas celulares y se contemplan para su uso en relación con la presente invención como vehículos para las composiciones peptídicas. Son ampliamente adecuados para atrapar sustancias tanto liposolubles como hidrosolubles atrapados, es decir en los espacios acuosos y dentro de la propia bicapa, respectivamente. Es posible que los liposomas portadores de fármacos incluso se pueden emplear para la administración específica de sitio de agentes activos modificando selectivamente la formulación liposómica.

10 Además de las enseñanzas de Couvreur y col., (1977; 1988), la siguiente información puede utilizarse en la generación de formulaciones liposómicas. Los fosfolípidos pueden formar diversas estructuras distintas de liposomas cuando se dispersan en agua, dependiendo de la relación molar entre el lípido y el agua. A relaciones bajas, el liposoma es la estructura preferida. Las características físicas de los liposomas dependen del pH, la fuerza iónica y la presencia de cationes divalentes. Los liposomas pueden mostrar baja permeabilidad a sustancias iónicas y polares, pero a temperaturas elevadas experimentan una transición de fase que altera notablemente su permeabilidad. La transición de fase implica un cambio de una estructura ordenada estrechamente empaquetada, conocida como el estado de gel, a una estructura menos ordenada débilmente empaquetada, conocida como el estado fluido. Esto se produce a una temperatura de transición de fase característica y da lugar a un aumento de la permeabilidad a los iones, azúcares y fármacos.

20 Además de la temperatura, la exposición a proteínas puede alterar la permeabilidad de los liposomas. Ciertas proteínas solubles, tales como el citocromo c, se unen, deforman y penetran en la bicapa, causando de este modo cambios en la permeabilidad. El colesterol inhibe esta penetración de proteínas, aparentemente empaquetando los fosfolípidos con más fuerza. Se contempla que las formaciones de liposomas más útiles para la administración de antibióticos e inhibidores contendrán colesterol.

25 La capacidad para atrapar solutos varía entre diferentes tipos de liposomas. Por ejemplo, las VML son moderadamente eficaces para atrapar solutos, pero las VUP son extremadamente ineficientes. Las VUP ofrecen la ventaja de homogeneidad y reproducibilidad en la distribución del tamaño, sin embargo, y un compromiso entre el tamaño y la eficiencia de captura lo ofrecen las vesículas unilamelares grandes (VUG). Estas se preparan mediante evaporación de éter y son de tres a cuatro veces más eficientes en el atrapamiento soluto que las VML.

30 Además de las características de los liposomas, un determinante importante en el atrapamiento de compuestos son las propiedades fisicoquímicas del propio compuesto. Los compuestos polares son atrapados en los espacios acuosos y los compuestos no polares se unen a la bicapa lipídica de la vesícula. Los compuestos polares se liberan a través de permeación o cuando la bicapa se rompe, pero los compuestos no polares permanecen afiliados a la bicapa a menos que se rompa por la temperatura o la exposición a lipoproteínas. Ambos tipos muestran velocidades máximas de flujo de salida a la temperatura de transición de fase.

35 Los liposomas interactúan con las células a través de cuatro mecanismos diferentes: endocitosis por células fagocíticas del sistema reticuloendotelial, tales como macrófagos y neutrófilos; adsorción a la superficie celular, ya sea por fuerzas hidrófobas o electrostáticas débiles no específicas, o mediante interacciones específicas con componentes de la superficie celular; fusión con la membrana plasmática celular por inserción de la bicapa lipídica del liposoma en la membrana plasmática, con liberación simultánea de contenidos liposomales en el citoplasma; y por transferencia de lípidos liposómicos a membranas celulares o subcelulares, o viceversa, sin ninguna asociación de los contenidos del liposoma. A menudo es difícil determinar qué mecanismo es operativo y más de uno pueden operar al mismo tiempo.

45 El destino y la disposición de los liposomas inyectados por vía intravenosa dependen de sus propiedades físicas, tales como el tamaño, la fluidez y la carga superficial. Pueden persistir en tejidos durante horas o días, dependiendo de su composición, y las semividas en el rango de sangre de minutos a varias horas.

50 Los liposomas más grandes, tales como las VML y las VUP, son absorbidos rápidamente por las células fagocíticas del sistema reticuloendotelial, pero la fisiología del sistema circulatorio restringe la salida de estas especies grandes en la mayoría de los sitios. Solo pueden salir en lugares en los que existen grandes aberturas o poros en el endotelio capilar, tales como los sinusoides del hígado o el bazo. Por lo tanto, Estos órganos son el sitio predominante de absorción. Por otro lado, las VUP muestran una distribución de tejido más amplia, pero aún son secuestradas considerablemente en el hígado y el bazo. En general, este comportamiento in vivo limita la segmentación potencial de los liposomas únicamente a aquellos órganos y tejidos accesibles para su gran tamaño. Estos incluyen la sangre, el hígado, el bazo, la médula ósea y los órganos linfoides.

55 La orientación no es generalmente una limitación en términos de la presente invención. Sin embargo, si se desea una orientación específica, se dispone de procedimientos para conseguirlo. Los anticuerpos se pueden usar para unirse a la superficie del liposoma y para dirigir el anticuerpo y su contenido de fármaco a receptores antigénicos específicos situados en una superficie del tipo de célula particular. También se pueden usar determinantes de carbohidratos (componentes de la superficie celular de glicoproteínas o glicolípidos que desempeñan un papel en el

reconocimiento célula-célula, interacción y adhesión) como sitios de reconocimiento, ya que tienen potencial para dirigir liposomas a tipos celulares particulares. Principalmente, se contempla el uso de la inyección intravenosa de preparaciones liposómicas, pero también son concebibles otras vías de administración.

5 Como alternativa, la invención proporciona formulaciones de nanocápsulas farmacéuticamente aceptables de las composiciones de la presente invención. Las nanocápsulas pueden atrapar generalmente compuestos de una manera estable y reproducible (Henry-Michelland *y col.*, 1987; Quintanar-Guerrero *y col.*, 1998; Douglas *y col.*, 1987). Para evitar efectos secundarios debido a la sobrecarga polimérica intracelular, tales partículas ultrafinas (de tamaño de aproximadamente 0,1 μm) deberían diseñarse usando polímeros capaces de degradarse *in vivo*. Las nanopartículas biodegradables de polialquil-cianoacrilato que cumplen estos requisitos se contemplan para su uso
10 en la presente invención. Tales partículas se pueden fabricar fácilmente, como se ha descrito (Couvreur *y col.*, 1980; 1988; zur Muhlen *y col.*, 1998; Zambaux *y col.*, 1998; Pinto-Alphandry *y col.*, 1995 y la patente de Estados Unidos N.º 5.145.684, incorporada específicamente en el presente documento por referencia en su totalidad).

También se pueden utilizar parches cutáneos para administración transcutánea.

COMPOSICIONES INMUNOGÉNICAS

15 En ciertas realizaciones preferentes de la presente invención, se proporcionan composiciones inmunogénicas. Las composiciones inmunogénicas comprenderán generalmente uno o más polipéptidos o polinucleótidos, tal como los descritos anteriormente, en combinación con un inmunestimulante. Un inmunestimulante puede ser cualquier sustancia que mejora o potencia una respuesta inmunitaria (de anticuerpos y/o mediada por células) a un antígeno exógeno. Los ejemplos de inmunestimulantes incluyen adyuvantes, microesferas biodegradables (por ejemplo, galactida poliláctica) y liposomas (en los que se incorpora el compuesto; véase, *por ejemplo*, Fullerton, la Patente de los Estados Unidos n.º 4.235.877).
20

La preparación de composiciones inmunogénicas se describe generalmente en, por ejemplo, Powell y Newman. eds., Vaccine Design (el enfoque de subunidades 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 partes inmunogénicas de otros antígenos de *M. tuberculosis*, ya sea incorporadas en un polipéptido de fusión o como un compuesto separado, dentro de la composición inmunogénica o farmacéutica.
25

Las composiciones inmunogénicas ilustrativas pueden contener un polinucleótido (por ejemplo ADN) que codifica uno o más de los polipéptidos como se ha descrito anteriormente, de manera que el polipéptido se genera *in situ* (provocando de este modo una respuesta inmunitaria). Como se ha indicado anteriormente, el ADN puede estar presente dentro de cualquiera de una variedad de sistemas de administración conocidos por los expertos en la materia, incluyendo sistemas de expresión de ácidos nucleicos, bacterias y sistemas de expresión en virus. En la materia se conocen bien numerosas vías de administración de genes, tal como las descritas por Rolland, Crit. Rev. Therap. Drug Carrier Systems 15:143-198 (1998), y referencias citadas en el mismo. Sistemas de expresión de ácido nucleico adecuados contienen las secuencias de ADN necesarias para la expresión en el paciente (tal como un promotor adecuado y una señal de terminación). Los sistemas de administración bacterianos implican la administración de una célula huésped bacteriana (por ejemplo, una cepa de *Mycobacterium*, *Bacillus* o *Lactobacillus*, incluyendo el bacilo de Calmette-Guerrin o *Lactococcus lactis*) que expresa el polipéptido (por ejemplo, en su superficie celular 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 introducirse usando un sistema de expresión viral (por ejemplo, vaccinia u otro virus de viruela, retrovirus, o adenovirus), que puede implicar el uso de un virus no patógeno (defectuoso), de replicación competente. Los sistemas adecuados se divulgan, 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. las Patentes de Estados Unidos n.º 4.603.112, 4.769.330 y 5.017.487; el documento WO 89/01973; la patente de Estados Unidos n.º 4.777.127; el documento GB 2,200,651; el documento EP 0345242; el documento 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 tales sistemas de expresión son bien conocidas por los expertos normales en la técnica. El ADN también puede ser "desnudo", como se ha descrito, por ejemplo, en Ulmer *y col.*, Science 259:1745-1749 (1993) y revisado por Cohen, Science 259:1691-1692 [1993]. La captación de ADN desnudo se puede aumentar recubriendo el ADN sobre perlas biodegradables, que se transportan eficazmente en las células. Será evidente que una composición inmunogénica puede comprender tanto un polinucleótido como un componente de polipéptido. Dicha composición inmunogénica puede proporcionar una respuesta inmunitaria mejorada.
30
35
40
45
50
55

Será evidente que una composición inmunogénica puede contener sales farmacéuticamente aceptables de los polinucleótidos y polipéptidos proporcionados en este documento. Tales sales se pueden preparar a partir de bases farmacéuticamente aceptables no tóxicas, 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)
60

Mientras que cualquier vehículo adecuado conocido por los expertos en la materia se puede emplear en las composiciones inmunogénicas de la presente invención, el tipo de vehículo variará dependiendo del modo de administración. Las composiciones de la presente invención pueden usarse para cualquier manera de administración adecuada, incluyendo, por ejemplo, tópica, oral, nasal, intravenosa, intracraneal, intraperitoneal, subcutánea o intramuscular. Para administración parenteral, tal como inyección subcutánea, el vehículo comprende, preferentemente, agua, solución salina, alcohol, una grasa, una cera o un tampón. Para administración oral, se puede usar cualquiera de los vehículos anteriores o un vehículo sólido, tales como manitol, lactosa, almidón, estearato de magnesio, sacarina sódica, talco, celulosa, glucosa, sacarosa y carbonato de magnesio. Las microesferas biodegradables (por ejemplo, polilactato poliglicolato) también se puede emplear como vehículos para las composiciones farmacéuticas de la presente invención. Las microesferas biodegradables adecuadas se divulgan, por ejemplo, en las patentes de Estados Unidos 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 partículas-proteínas descritos en la patente de Estados Unidos N.º 5.928.647, que son capaces de inducir respuestas de linfocitos T citotóxicos restringidos de clase I en un huésped.

Tales composiciones también pueden comprender tampones (por ejemplo, solución salina tamponada neutra o solución salina tamponada con fosfato), hidratos de carbono (por ejemplo, glucosa, manosa, sacarosa o dextranos), manitol, proteínas, polipéptidos o aminoácidos, tales como glicina, antioxidantes, bacteriostáticos, agentes quelantes, tales como EDTA o glutatión, adyuvantes por ejemplo, hidróxido de aluminio), solutos que hacen que la formulación sea isotónica, hipotónica o débilmente hipertónica con la sangre del receptor, agentes suspensores, agentes espesantes y/o disolventes. Como alternativa, las composiciones de la presente invención pueden formularse como un liofilizado. Los compuestos también se pueden encapsular dentro de liposomas usando tecnología bien conocida.

Cualquiera de diversos inmunoestimulantes se puede emplear en las composiciones inmunogénicas de esta invención. Por ejemplo, puede incluirse un adyuvante. La mayoría de los adyuvantes contienen una sustancia diseñada para proteger el antígeno del catabolismo rápido, tal como hidróxido de aluminio o aceite mineral, y un estimulador de respuestas inmunitarias, tales como lípido A, *Bordetella pertussis* o especies de *Mycobacterium* o proteínas derivadas de *Mycobacterium*. Por ejemplo, se puede usar *M. vaccae* ("pVac") deslipidada y desglicolipidada. Los adyuvantes adecuados están disponibles comercialmente como, por ejemplo, adyuvante Incompleto de Freund y adyuvante completo de Freund (Difco Laboratories, Detroit, MI); adyuvante 65 de Merck (Merck and Company, Inc., Rahway, NJ); AS01 B, AS02A, AS15, AS-2 y derivados de los mismos (GlaxoSmithKline, Philadelphia, PA); CWS (esqueleto de pared celular de un bacilo tuberculoso), TDM (dicorinomicolato de trehalosa), Leif (factor de iniciación de la elongación de Leishmania), sales de aluminio tales como gel de hidróxido de aluminio (alumbre) o fosfato de aluminio; sales de calcio, hierro o cinc; una suspensión insoluble de tirosina acilada; azúcares acilados; polisacáridos derivatizados catiónicamente o aniónicamente; polifosfacenos; microesferas biodegradables; monofosforil lípido A (MPL®); y Quil A (por ejemplo, QS-21). Las citocinas, tales como GM-CSF o interleucina-2, -7 o -12, también se pueden usar como adyuvantes.

Un adyuvante se refiere a los componentes en una vacuna o composición terapéutica que aumentan la respuesta inmunitaria específica al antígeno (véase, por ejemplo, Edelman, AIDS Res. Hum Retroviruses 8:1409-1411 (1992)). Los adyuvantes inducen respuestas inmunitarias del tipo Th1 y de tipo Th-2. Las citocinas de tipo Th1 (por ejemplo, IFN- γ , IL-2 e IL-12) tienden a favorecer la inducción de la respuesta inmunitaria mediada por células a un antígeno administrado, mientras que las citocinas tipo Th-2 (por ejemplo, IL-4, IL-5, IL-6, IL-10) tienden a favorecer la inducción de respuestas inmunitarias humorales. Los adyuvantes capaces de estimulación preferencial de una respuesta inmunitaria mediada por células Th 1 se describen en el documento WO 94/00153 y WO 95/17209.

Dentro de las composiciones inmunogénicas proporcionadas en el presente documento, a composición adyuvante se diseña preferentemente para inducir una respuesta inmune predominantemente del tipo Th1. Después de la aplicación de una composición inmunogénica como se proporciona en el presente documento, un paciente típicamente soportará una respuesta inmunitaria que incluye respuestas de tipo Th1 y Th2. En una realización preferente, en la que una respuesta es predominantemente de tipo Th1, el nivel de citocinas de tipo Th1 aumentará a un grado mayor que el nivel de citocinas de tipo Th2. Los niveles de estas citocinas pueden evaluarse fácilmente mediante ensayos convencionales. Para una revisión de las familias de citocinas, véase Janeway, y col., Immunobiology, 5ª Edición, 2001.

Las composiciones Rv1753c normalmente comprenden uno o más adyuvantes, por ejemplo, AS01 B (3-de-O-acilado monofosforil lípido A (3D-MPL®) y QS21 en una formulación de liposomas, véase, la Publicación de Patente de Estados Unidos N.º 2003/0143240); AS02A (3D-MPL® y QS21 y una emulsión de aceite en agua; véase, Bojang, y col., Lancet (2001) 358:1927); ENHANZYN® (Detox); 3D-MPL®; saponinas, incluyendo Quil A y sus componentes, por ejemplo, miméticos de QS21 y de saponina; CWS (pared esqueleto celular de un bacilo de la tuberculosis); TDM (dicorinomicolato de trehalosa); aminoalquilo glucosaminida 4-fosfatos (AGPS); oligonucleopéptidos inmunoestimulantes, por ejemplo, CPG; Leif (Leishmania factor de iniciación de elongación); y derivados de los mismos. En una realización preferida, un polipéptido Rv1753c se administra con uno o más adyuvantes seleccionados del grupo que consiste en 3D-MPL® y QS21 en una formulación de liposomas, por ejemplo, AS01 B y 3D-MPL® y QS21, y una emulsión de aceite en agua (por ejemplo AS02A). Los sistemas adyuvantes AS01 B y AS02A se describen adicionalmente en Pichyangkul, y col., Vaccine (2004) 22:3831-40.

Al administrar el antígeno Rv1753c como ácido nucleico, se puede administrar, por ejemplo, en un vector viral (es decir, un vector de adenovirus) o en una célula huésped de bacteria mutante (es decir, un mutante, *Mycobacterium avirulentum*, célula huésped de *Lactobacillus* o *Bacillus*, incluyendo bacilo de Calmette-Guerin (BCG) y *Lactococcus lactis*).

5 Los adyuvantes preferidos para su uso en provocar una respuesta predominantemente de tipo Th1 incluyen, por ejemplo, una combinación de monofosforil lípido A (MPL®), preferentemente 3-O-desacilado monofosforil lípido A (3D-MPL®), opcionalmente con una sal de aluminio (véase, por ejemplo, Ribí, y col., 1986, Immunology and Immunopharmacology of Bacterial Endotoxins, Plenum Publ. Corp., NY, pág. 407-419; GB 2122204B; GB 2220211; y US 4,912,094). Una forma preferida de 3D-MPL® está en forma de una emulsión que tiene un tamaño de partícula pequeño inferior a 0,2 µm de diámetro y su procedimiento de fabricación se divulga en el documento WO 94/21292. Las formulaciones acuosas que comprenden monofosforil lípido A y un tensioactivo se han descrito en el documento WO 98/43670. Adyuvantes preferidos de ejemplo incluyen AS01 B (MPL® y QS21 en una formulación de liposomas), 3D-MPL® y QS21 en una formulación de liposomas, AS02A (MPL® y QS21 y una emulsión de aceite en agua), 3D-MPL® y QS21 y una emulsión de aceite en agua y AS15, disponibles en GlaxoSmithKline. Los adyuvante MPL® están disponibles en GlaxoSmithKline (véanse las patentes de Estados Unidos n.º 4.436.727; 4.877.611; 4.866.034 y 4.912.094).

Los oligonucleótidos que contienen CpG (en los que el dinucleótido CpG no está metilado) también inducen una respuesta predominante Th1- CpG es una abreviatura para motivos de dinucleótido de citosina-guanosina presentes en el ADN. Dichos oligonucleótidos son bien conocidos y se describen, por ejemplo, en el documento WO 96/02555, el documento WO 99/33488 y las patentes de Estados Unidos n.º N.º 6,008,200 y 5.856.462. Las secuencias de ADN inmunoestimulantes también se describen, por ejemplo, en Sato y col., Science 273:352 (1996). CpG cuando se formula en composiciones inmunogénicas, se administra generalmente en solución libre junto con antígeno libre (documento WO 96/02555; McCluskie y Davis, citado anteriormente) o conjugado covalentemente a un antígeno (documento WO 98/16247), o formulado con un vehículo, tal como hidróxido de aluminio (antígeno de la superficie de hepatitis) Davis y col., citado anteriormente; Brazolot-Millan y col., Proc.Natl.Acad.Sci., EE.UU., 1998, 95(26), 15553-8). Se sabe que CpG se conoce en la técnica como adyuvante cuando se administra por las rutas tanto sistémicas como mucosales (documento WO 96/02555, documento EP 468520, Davis y col, J.Immunol, 1998, 160(2):870-876; McCluskie y Davis, J.Immunol., 1998, 161(9):4463-6).

Otro adyuvante preferido es una saponina o miméticos o derivados de saponina, tales como Quil A, preferentemente QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), que se puede usar solo o en combinación con otros adyuvantes. Por ejemplo, un sistema potenciado implica la combinación de un monofosforil lípido A (MPL®) y derivado de saponina, tal como la combinación de QS21 y 3D-MPL®, como se describe en el documento WO 94/00153, o una composición menos reactogénica en la que QS21 se inactiva con colesterol, como se describe en el documento WO 96/33739. Otras formulaciones preferidas comprenden una emulsión de aceite en agua y tocoferol. Una formulación adyuvante particularmente potente que implica QS21, 3D-MPL® y tocoferol en una emulsión de aceite en agua se describe en el documento WO 95/17210. Adyuvantes de saponina adicionales de uso en la presente invención incluyen QS7 (descrita en los documentos WO 96/33739 y WO 96/11711) y QS17 (descritos en la patente de Estados Unidos n.º 5.057.540 y el documento EP 0 362 279 B1).

Como alternativa, las formulaciones de saponina pueden combinarse con vehículos de vacunas compuestos de quitosano u otros polímeros policatiónicos, partículas de polilactida y polilactida-co-glicólido, matriz polimérica a base de poli-N-acetil glucosamina, partículas compuestas por polisacáridos o polisacáridos modificados químicamente, partículas basadas en lípidos y liposomas, partículas compuestas por monoésteres de glicerol, etc. las saponinas también se pueden formular en presencia de colesterol para formar estructuras particuladas tales como liposomas o ISCOM®. Adicionalmente, las saponinas pueden formularse junto con un éter o éster de polioxietileno, ya sea en una solución no particulada o suspensión, o en una estructura particulada, tal como un liposoma paucilamelar o ISCOM®. Las saponinas también se pueden formular con excipientes tales como CARBOPOL® para aumentar la viscosidad, o puede formularse en una forma de polvo seco con un excipiente en polvo tal como lactosa.

En una realización, el sistema adyuvante incluye la combinación de un monofosforil lípido A y un derivado de saponina, tal como la combinación de adyuvante QS21 y 3D-MPL®, como se describe en el documento WO 94/00153, o una composición menos reactogénica en la que QS21 se inactiva con liposomas que contienen colesterol, como se describe en el documento WO 96/33739. Otras formulaciones adecuadas comprenden una emulsión de aceite en agua y tocoferol. Otra formulación de adyuvante adecuada que usa QS21, adyuvante 3D-MPL® y tocoferol en una emulsión de aceite en agua se describe en el documento WO 95/17210.

Otro sistema adyuvante potenciado implica la combinación de un oligonucleótido que contiene CpG y un derivado de saponina, particularmente la combinación de CpG y QS21 como se describe en el documento WO 00/09159. Adecuadamente, la formulación comprende adicionalmente una emulsión de aceite en agua y tocoferol.

Otros adyuvantes adecuados incluyen MONTANIDE® ISA 720 (Seppic, Francia), SAF (Chiron, California, Estados Unidos), ISCOMS® (CSL), MF-59 (Chiron), la serie SBAS de adyuvantes (SmithKline Beecham, Rixensart, Bélgica), Detox (Corixa), RC-529 (Corixa) y otros aminoalquilglucosaminida 4-fosfatos (AGP), como os descritos en la solicitud de patente de Estados Unidos pendiente de tramitación n.º 8/853,826 y 09/074,720, cuyas divulgaciones son

incorporadas por referencia en el presente documento en su totalidad, y adyuvantes de polioxietiléneter, como los descritos en el documento WO 99/52549A1. SmithKline Beecham y Corixa Corporation ahora forman parte de GlaxoSmithKline.

Otros adyuvantes adecuados incluyen moléculas adyuvantes de la fórmula general (I):



en la que, n es 1-50, A es un enlace o -C(O)-, R es alquilo C₁₋₅₀ o fenilalquilo C₁₋₅₀.

Un adyuvante adicional de interés es la cadena b de la toxina Shiga, que se utiliza por ejemplo como se describe en el documento WO2005/112991.

10 Una realización de la presente invención consiste en una composición inmunogénica que comprende un éter de polioxietileno de fórmula general (I), en la que n es entre 1 y 50, preferentemente de 4-24, lo más preferentemente 9; el componente R es C₁₋₅₀, preferentemente alquilo C_{4-C20} y, más preferentemente alquilo C₁₂, y A es un enlace. La concentración de los éteres de polioxietileno debe estar en el intervalo de 0,1-20 %, preferentemente de 0,1-10 %, y lo más preferentemente en el intervalo de 0,1-1 %. Los polioxietiléneteres preferidos se seleccionan del siguiente grupo: polioxietilen-9-lauriléter, polioxietilen-9-esteoriléter, polioxietilen-8-esteoriléter, polioxietilen-4-lauriléter, polioxietilen-35-lauriléter y polioxietilen-23-lauriléter. Los polioxietiléneteres, tales como polioxietilenlauriléter, se describen en el índice de Merck (12 edición: entrada 7717). Estas moléculas adyuvantes se describen en el documento WO 99/52549.

20 Cualquier composición inmunogénica proporcionada en este documento puede prepararse usando procedimientos bien conocidos que dan como resultado una combinación de antígeno, potenciador de la respuesta inmunitaria y un vehículo o excipiente adecuado. Las composiciones descritas en el presente documento se pueden administrar como parte de una formulación de liberación sostenida (es decir, una formulación tal como una cápsula, esponja o gel (compuesta de polisacáridos, por ejemplo) que efectúa una liberación lenta del compuesto tras la administración). Tales formulaciones pueden prepararse generalmente usando tecnología bien conocida (véase, por ejemplo, Coombes y col., Vaccine 14: 1429-1438 (1996)) y administrado por, por ejemplo, vía oral, rectal o implantación subcutánea, o mediante implantación en el sitio de destino deseado. Las formulaciones de liberación sostenida pueden contener un polipéptido, polinucleótido o anticuerpo disperso en una matriz vehículo y/o contenido dentro de un depósito rodeado por una membrana de control de velocidad.

30 Los vehículos para su uso dentro de tales formulaciones son biocompatibles, y también pueden ser biodegradables; preferentemente la formulación proporciona un nivel relativamente constante de liberación del componente activo. Tales vehículos incluyen micropartículas de poli(lactida-co-glicolida), poliacrilato, látex, almidón, celulosa, dextrano y similares. Otros vehículos de liberación retardada incluyen biovectores supramoleculares, que comprenden un núcleo hidrófilo no líquido (por ejemplo, un polisacárido u oligosacárido reticulado) y, opcionalmente, una capa externa que comprende un compuesto anfífilo, tal como un fosfolípido (véase, por ejemplo, la patente de Estados Unidos Nº 5.151.254 y las solicitudes PCR WO 94/20078, WO 94/23701 y WO 96/06638). La cantidad de compuesto activo contenido dentro de una formulación de liberación sostenida depende del sitio de implantación, la velocidad y duración esperada de la liberación.

40 Se puede usar cualquiera de una variedad de vehículos de administración dentro de las composiciones farmacéuticas y composiciones inmunogénicas para facilitar la producción de una respuesta inmunitaria específica de antígeno. Los vehículos de administración incluyen células presentadoras de antígeno (CPA), tales como células dendríticas, macrófagos, linfocitos B, monocitos y otras células que pueden modificarse por ingeniería genética para que sean CPA eficaces. Tales células pueden pero no lo necesitan, modificarse genéticamente para aumentar la capacidad de presentar el antígeno, para mejorar la activación y/o el mantenimiento de la respuesta de linfocitos T y/o para que sean inmunológicamente compatibles con el receptor (es decir, de haplotipo HLA equivalente). Las CPA pueden generalmente aislarse de cualquiera de diversos fluidos biológicos y órganos, y pueden ser células autólogas, alogénicas, singénicas o xenogénicas.

50 Ciertas realizaciones preferidas de la presente invención usan células dendríticas o progenitoras de las mismas como células presentadoras de antígeno. Las células dendríticas son CPA muy potentes (Banchereau y Steinman, Nature 392: 245-251 (1998)) y se ha demostrado que son eficaces como adyuvante fisiológico para inducir inmunidad profiláctica o terapéutica (véase Timmerman y Levy, Ann. Rev. Med. 50:507-529(1999)). En general, las células dendríticas pueden identificarse basándose en su forma típica (estrellada *in situ*, con marcados procesos citoplásmicos (dendritas) visibles *in vitro*), su capacidad para captar, procesar y presentar antígenos con alta eficacia y su capacidad para activar respuestas de linfocitos T no expuestos previamente. Las células dendríticas pueden, por supuesto, modificarse por ingeniería genética para expresar receptores de superficie celular específicos o ligandos que no se encuentran habitualmente en las células dendríticas *in vivo* o *ex vivo*, y tales células dendríticas modificadas se contemplan en la presente invención. Como alternativa a las células dendríticas, las vesículas secretadas de células dendríticas cargadas con antígeno (llamadas exosomas) pueden usarse dentro de una composición inmunogénica (véase Zitvogel y col., Nature Med. 4:594-600(1998)).

Las células dendríticas y los progenitores pueden obtenerse de sangre periférica, médula ósea, ganglios linfáticos,

bazo, piel, sangre de cordón umbilical o cualquier otro tejido adecuado o fluido. Por ejemplo, las células dendríticas pueden diferenciarse *ex vivo* añadiendo una combinación de citocinas tales como GM-CSF, IL-4, IL-13 y/o TNF a cultivos de monocitos recogidos de sangre periférica. Como alternativa, las células CD34 positivas recolectadas de sangre periférica, sangre del cordón umbilical o médula ósea pueden diferenciarse en células dendríticas añadiendo al medio de cultivo combinaciones de GM-CSF, IL-3, TNF α , ligando de CD40, LPS, ligando de flt3 y/o otro u otros compuestos que inducen la diferenciación, maduración y proliferación de células dendríticas.

Las células dendríticas se clasifican convenientemente como células "inmaduras" y "maduras", que permite una forma sencilla de discriminar entre dos fenotipos bien caracterizados. Sin embargo, esta nomenclatura no debe interpretarse que excluye todas las posibles etapas intermedias de diferenciación. Las células dendríticas inmaduras se caracterizan como CPA con una alta capacidad de captación y procesamiento de antígenos, que se correlaciona con la alta expresión del receptor Fc y el receptor de manosa. El fenotipo maduro se caracteriza típicamente por una menor expresión de estos marcadores, pero una alta expresión de moléculas de superficie celular responsables de la activación de los linfocitos T tales como el MHC de clase I y de clase II, moléculas de adhesión (por ejemplo, CD54 y CD11) y moléculas coestimuladoras (por ejemplo, CD40, CD80, CD86 y 4-1 BB).

Las CPA pueden generalmente transfectarse con un polinucleótido que codifica una proteína (o porción u otra variante de la misma) de modo que el polipéptido, se expresa en la superficie celular. Dicha transfección puede tener lugar *ex vivo*, y entonces puede usarse una composición farmacéutica o composición inmunogénica que comprende dichas células transfectadas, tal como se describe en el presente documento. Como alternativa, un vehículo de liberación génica cuyo objetivo es una célula presentadora de antígeno dendrítica o agente puede administrarse a un paciente, dando como resultado transfección que se produce *in vivo*. La transfección *in vivo* y *ex vivo* de las células dendríticas, por ejemplo, puede generalmente realizarse usando cualquiera de los procedimientos conocidos en la técnica, tales como los descritos en el documento WO 97/24447, o el enfoque de pistola génica descrito por Mahvi y col., *Immunology and Cell Biology* 75:456-460 (1997). La carga de antígenos de las células dendríticas puede conseguirse incubando células dendríticas o células progenitoras con el polipéptido, ADN (desnudo o dentro de un vector plasmídico) o ARN; o con bacteria o virus recombinante que expresa el antígeno (por ejemplo, vaccinia, viruela aviar, adenovirus o lentivirus). Antes de la carga, el polipéptido se puede conjugar covalentemente a una pareja inmunológica que proporcione ayuda de linfocitos T (por ejemplo, una molécula vehículo). Como alternativa, una célula dendrítica puede pulsarse con una pareja inmunológica no conjugada, separadamente o en presencia del polipéptido.

Las composiciones inmunogénicas y las composiciones farmacéuticas pueden presentarse en recipientes monodosis o multidosis, tales como ampollas o viales sellados. Tales recipientes están preferentemente sellados herméticamente para conservar la esterilidad de la formulación hasta su uso. En general, las formulaciones pueden almacenarse como suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos. Como alternativa, una composición farmacéutica o composición inmunogénica se pueden almacenar en un estado liofilizado que requiere solo la adición de un vehículo líquido estéril inmediatamente antes de su uso.

En el presente documento se divulga a una "sensibilización" o primera administración de un polipéptido Rv1753c (incluyendo las variantes, fragmentos inmunogénicos o proteínas de fusión), o polinucleótido que codifica un polipéptido, le sigue uno o más "refuerzos" o administraciones posteriores de un polipéptido Rv1753c (incluyendo variantes, fragmentos inmunogénicos o proteínas de fusión) o polinucleótido que codifica dicho polipéptido (procedimiento de "sensibilización y refuerzo"). Por ejemplo, una primera administración con un polipéptido Rv1753c (incluyendo las variantes, fragmentos inmunogénicos o proteínas de fusión) o un polinucleótido que codifica dicho polipéptido es seguida por una o más administraciones posteriores de un polipéptido Rv1753c (incluyendo variantes, fragmentos inmunogénicos o proteínas de fusión) o un polinucleótido que codifica dicho polipéptido.

En el presente documento se divulga una primera administración con un polipéptido o polinucleótido Rv1753c es seguida por una o más administraciones posteriores de un polipéptido Rv1753c. En una realización, una primera administración con un polipéptido o polinucleótido Rv1753c es seguida por una o más administraciones posteriores de un polinucleótido Rv1753c. Por lo general, la primera administración o "sensibilización" y la segunda administración o "refuerzo" se proporcionan con aproximadamente 2-12 semanas de diferencia, o con hasta 4-6 meses de diferencia. Las posteriores administraciones de "refuerzo" se dan con aproximadamente 6 meses de diferencia, o con tanto como 1, 2, 3, 4 o 5 años de diferencia. El tratamiento de refuerzo convencional (por ejemplo, una administración de sensibilización a proteína seguida por una administración de refuerzo de proteína) puede ser también útil en la prevención o tratamiento de la tuberculosis (por ejemplo, la prevención o el tratamiento de la tuberculosis latente, en particular la prevención o retraso de la reactivación de la tuberculosis).

ANTICUERPOS

"Anticuerpo" se refiere a un polipéptido que comprende una región marco de un gen de inmunoglobulina o fragmentos del mismo que se une específicamente y reconoce un antígeno. Los genes de inmunoglobulina reconocidos incluyen los genes de región constante kappa, lambda, alfa, gamma delta, epsilon y mu, así como la miríada de genes de la región variable de las inmunoglobulinas. Las cadenas ligeras se clasifican como cadenas kappa o lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta o epsilon, que, a su vez, definen

las clases de inmunoglobulina IgG, IgM, IgA, IgD e IgE, respectivamente.

Una unidad estructural de inmunoglobulina (anticuerpo) de ejemplo comprende un tetrámero. Cada tetrámero está compuesto por dos pares idénticos de cadenas polipeptídicas, en el que cada par tiene una cadena "ligera" (aproximadamente 25 kDa) y una cadena "pesada" (aproximadamente 50-70 kDa). El extremo N de cada cadena define una región variable de aproximadamente 100 a 110 o más aminoácidos principalmente responsables del reconocimiento del antígeno. Las expresiones "cadena ligera variable" (V_L) y cadena pesada variable (V_H) se refiere a estas cadenas ligeras y pesadas, respectivamente.

Los anticuerpos existen como inmunoglobulinas intactas o como una serie de fragmentos bien caracterizados producidos mediante digestión con varias peptidasas. Por lo tanto, por ejemplo, la pepsina digiere un anticuerpo debajo de los puentes disulfuro en la región bisagra para producir $F(ab)'_2$, un dímero de Fab que en sí mismo es una cadena ligera unida a V_H-C_H1 mediante un puente disulfuro. El $r F(ab)'_2$, puede reducirse en condiciones suaves para romper el enlace disulfuro en la región bisagra, de modo que el dímero $r F(ab)'_2$, se convierte en un monómero Fab' . El monómero Fab' es esencialmente Fab con parte de la región bisagra (véase *Fundamental Immunology* (Paul ed., 3ª ed. 1993). Aunque varios fragmentos de anticuerpo se definen en términos de la digestión de un anticuerpo intacto, un experto en la técnica apreciará que dichos fragmentos pueden sintetizarse de novo bien químicamente o usando metodología de ADN recombinante. Por tanto, el término anticuerpo, como se usa en el presente documento, también incluyen fragmentos de anticuerpos bien producidos mediante la modificación de los anticuerpos enteros o los sintetizados *de novo* usando metodologías de ADN recombinante (por ejemplo, Fv de cadena sencilla) (scFv)) o los identificados usando bibliotecas de expresión en fase (véase, por ejemplo, McCafferty y col., *Nature* 348:552-554 (1990)).

Para la preparación anticuerpos monoclonales o policlonales, se puede usar cualquier técnica conocida en la materia (véase, por ejemplo, Kohler y Milstein, *Nature* 256:495-497 (1975); Kozbor y col., *Immunology Today* 4: 72 (1983); Cole y col., pp. 77-96 en *Monoclonal Antibodies and Cancer Therapy* (1985)). Se pueden adaptar técnicas para la producción de anticuerpos de cadena sencilla (patente de Estados Unidos n.º 4.946.778) para producir anticuerpos frente a polipéptidos de la presente invención. Asimismo, se pueden usar ratones transgénicos, u otros organismos, tales como otros mamíferos, para expresar anticuerpos humanizados. Como alternativa, se puede usar tecnología de expresión en fagos para identificar anticuerpos y fragmentos Fab heteroméricos que se unen específicamente a antígenos seleccionados (véase, por ejemplo, McCafferty y col., *Nature* 348:552-554 (1990); Marks y col., *Biotechnology* 10:779-783 (1992)).

La frase "se une específicamente (o selectivamente) a un anticuerpo" o "es inmunorreactivo específicamente (o selectivamente)" con, cuando hace referencia a una proteína o péptido, se refiere a una reacción de unión que es determinativa de la presencia de la proteína en una población heterogénea de proteínas y otros productos biológicos. Por lo tanto, en las condiciones de inmunoensayo designadas, los anticuerpos especificados se unen a una proteína concreta, al menos dos veces la basal y no se unen sustancialmente en una cantidad significativa a otras proteínas presentes en la muestra. La unión específica a un anticuerpo en dichas condiciones puede requerir un anticuerpo que se selecciona según su especificidad para una proteína concreta. Por ejemplo, se pueden seleccionar anticuerpos policlonales producidos frente a las proteínas de fusión para obtener solo aquellos anticuerpos policlonales que son específicamente inmunorreactivos con la proteína de fusión y no con componentes individuales de las proteínas de fusión. Esta selección se puede efectuar restando los anticuerpos que sufren reacción cruzada con los antígenos individuales. Se pueden usar diversos formatos de inmunoensayo para seleccionar anticuerpos específicamente inmunorreactivos con una proteína particular. Por ejemplo, de forma rutinaria se usan inmunoensayos ELISA en fase sólida para seleccionar anticuerpos específicamente inmunorreactivos con una proteína (véase, por ejemplo, Harlow y Lane, *Antibodies, A Laboratory Manual* (1988) y *Using Antibodies: A Laboratory Manual* (1998), para una descripción de formatos y condiciones de inmunoensayo que se pueden utilizar para determinar la inmunorreactividad específica). Típicamente, una reacción selectiva o específica será al menos dos veces la señal o ruido de fondo y, más típicamente, más de 10, 20 o 100 veces el fondo (por ejemplo, la unión a otras proteínas de *Mycobacterium*, tales como otras proteínas de *Mycobacterium tuberculosis*).

DIAGNÓSTICO

En el presente documento se divulgan procedimientos para usar uno o más de los polipéptidos descritos anteriormente para diagnosticar tuberculosis (por ejemplo, utilizando ensayos basados en la respuesta de linfocitos T o ensayos basados en anticuerpos de formato convencional).

Por ejemplo, se proporciona un procedimiento para determinar infección previa por *M. tuberculosis* en un individuo que comprende:

- a) obtener una muestra del individuo;
- (b) poner en contacto dicha muestra con un polipéptido aislado que comprende:
 - (i) una secuencia de la proteína Rv1753c;
 - (ii) una variante de una secuencia de la proteína Rv1753c; o

(iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c;

(c) cuantificar la respuesta de muestra.

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

Idóneamente, el individuo no habrá sido previamente vacunado contra la infección por *M. tuberculosis* (por ejemplo, idóneamente, el individuo no han sido previamente vacunados con BCG).

10 La respuesta de la muestra se puede cuantificar mediante una serie de medios conocidos por los expertos en la materia, incluyendo la monitorización de la proliferación de linfocitos o la producción de citocinas o anticuerpos específicos. Por ejemplo, se puede usar ELISPOT de linfocitos T para monitorizar las citocinas tales como interferón gamma (IFN), interleucina 2 (IL2) e interleucina 5 (IL5). se puede usar ELISPOT de linfocitos B-cell para monitorizar la estimulación de antígenos específicos de *M. tuberculosis*. La respuesta celular puede caracterizarse también por el uso de tinción intra y extracelular y análisis mediante un citómetro de flujo.

Los procedimientos de cuantificación de una respuesta de proliferación de la muestra incluyen:

15 (i) pulsar células cultivadas con un radiomarcador (por ejemplo timidina tritiada) y monitorización de la captación de tritio (por ejemplo, centelleo gas);
(ii) marcaje con éster de carboxifluoresceína succinimidil diacetato (CFSE) y monitorización de la fluorescencia de la división celular mediante citometría de flujo.

20 La cuantificación de una respuesta de citocinas de ejemplo incluye, en particular, la monitorización de la producción de interferón gamma.

Cuando se usan tales procedimientos de cuantificación, una respuesta positiva a un antígeno puede estar definida por una relación de señal a ruido (relación S/N) de al menos 2:1 (por ejemplo, al menos 3:1 o al menos 5:1).

25 En un aspecto adicional de la presente divulgación, se proporcionan procedimientos para diagnosticar la infección por *M. tuberculosis* usando una prueba cutánea. Tal como se usa en el presente documento, una "prueba cutánea" es cualquier ensayo realizado directamente en un paciente en el que se mide una reacción de hipersensibilidad de tipo retardado (DTH) (tal como hinchazón, enrojecimiento o dermatitis) después de la inyección intradérmica de un polipéptido Rv153c como se ha descrito anteriormente (o variante, fragmentos inmunogénicos de la misma o nucleótidos que los codifican). Dicha inyección se puede lograr usando cualquier dispositivo adecuado suficiente para poner en contacto las combinaciones de antígeno con células dérmicas del paciente, tal como una jeringa de tuberculina o jeringuilla de 1 ml. La reacción se mide después de un período de tiempo, por ejemplo al menos 48 horas después de la inyección, especialmente 48-72 horas.

30 La reacción DTH es una respuesta inmunitaria mediada por células, que es mayor en pacientes que han sido expuestos previamente al antígeno de ensayo. La respuesta se puede medir visualmente, usando una regla. En general, una respuesta que es mayor que aproximadamente 0,5 cm de diámetro, especialmente mayor que aproximadamente 1,0 cm de diámetro, es una respuesta positiva, indicativa de infección previa por *M. tuberculosis*, que puede o no puede manifestarse como enfermedad activa.

35 Para el uso en una prueba cutánea, el componente Rv1753c se formula adecuadamente como una composición farmacéutica que contiene un vehículo fisiológicamente aceptable. De manera adecuada, el vehículo empleado en tales composiciones farmacéuticas es una solución salina con conservantes apropiados, tales como fenol y/o Tween 80™.

Ejemplos

Los ejemplos siguientes se proporcionan a modo de ilustración únicamente y no como limitación. Los expertos en la técnica reconocerán fácilmente varios parámetros no críticos que podrían modificarse o cambiarse para dar esencialmente resultados similares.

45 Ejemplo 1 - Identificación de Rv1753c como objetivo de vacuna de TB latente

El gen Rv1753c, también conocido como PPE24, codifica para un miembro de proteína de la familia PPE de *Mycobacterium tuberculosis* de proteínas ricas en Gly, Asn.

50 Rv1753c se seleccionó basándose en un análisis de todo el genoma de los genes de *Mycobacterium tuberculosis* asociados con el mantenimiento de la fase de latencia y la infectividad como en Murphy y Brown BMC.Infect. Dis. 2007 7:84-99. Las dianas del gen de fase de latencia potencial en *Mycobacterium tuberculosis* se priorizaron a través de un meta-análisis de bioinformática de conjuntos de datos de micromatriz de ADN de todo el genoma publicados de la expresión de genes bacterianos en condiciones de latencia simuladas. La localización subcelular de las proteínas de *M. tuberculosis* codificadas por genes se llevó a cabo posteriormente en todo el genoma para

identificar los objetivos de la vacuna.

En resumen, las condiciones experimentales en los modelos de latencia fueron bastante variadas, de modo que se desarrolló un sistema de puntuación de cero a cinco para normalizar estos datos según dos criterios: 1) la pertinencia de las condiciones experimentales al estado de latencia y 2) el orden de rango de expresión. La puntuación máxima para un conjunto de datos experimentales en particular se ajustó basado en la potencial relevancia para la aparición clínica de las infecciones de *M. tuberculosis* en fase de latencia. La Tabla 1 muestra los conjuntos de datos recogidos para la etapa 1 junto con las puntuaciones máximas ajustadas para cada conjunto de datos. Los conjuntos de datos adicionales sobre la esencialidad del gen para el crecimiento se obtuvieron de los estudios publicados usando experimentos de inactivación basada en transposones (TraSH). Los genes que no tenían ningún efecto sobre el crecimiento recibieron una puntuación de cero.

Tabla 1- Puntuaciones modelos experimentales y criterios de puntuación para la expresión génica en micromatriz de ADN de *M. tuberculosis* e inactivación génica de todo el genoma (esencialidad de la fase de crecimiento).

Referencia	Modelo experimental	Punto de tiempo: Puntuación máxima ^a
Betts JC y col., Mol. Microbiol. 2002 43:717-731	Inanición en O ₂ controlado	96 h: 3 24 h: 2 4 h: 1
Hampshire T y col., Tuberculosis.(Edinb.) 2004 84:228-238	Depleción de nutrientes en O ₂ controlado	62 y 75 d: 5 49 d: 4 18 d: 2
Muttucumaru DG y col., Tuberculosis.(Edinb.) 2004 84:239-246	Modelo Wayne de hipoxia [#]	14 d (NRP-2): 4 7 d (NRP-1): 2
Voskuil MI y col., Tuberculosis.(Edinb.) 2004 84:218-227	Modelo Wayne de hipoxia [#]	30 y 80 d: 5 14 y 20 d: 4 10 y 12 d: 3 6 y 8 d: 2
Schnappinger D y col., J. Exp. Med. 2003 198:693-704	Infección de macrófagos de ratón, +/- γ -INF	24 y 48 h: 5
Karakousis PC y col., J. Exp. Med. 2004 200:647-657	Implante subcutáneo de fibras huescas en ratones	10 d: 3
Talaat AM y col., Proc. Natl. Acad. Sci. U.S.A 2004, 101:4602-4607	Infección de ratones. MTB recogido de pulmón ^b	28 d: 3
Sasseti CM y col., Mol. Microbiol. 2003 48:77-84	Bibliotecas mutadas de TraSH cultivadas en medio sólido	14 d: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 <i>M. tuberculosis</i>	7 d:5
Sasseti CM y col., Proc.Natl.Acad.Sci.U.S.A 2003 100:12989-12994	Ratones C57BL/6J infectados con bibliotecas mutadas de TraSH de <i>M. tuberculosis</i>	7, 14, 28 y 56 d: 5
^a Puntuación máxima según la relevancia como modelo de latencia; h = horas; d = día. ^b Relación de <i>M. tuberculosis</i> de pulmón Balb/c respecto a MTB en cultivo aireado durante 28 d. [#] Wayne LG y Hayes LG Infect. Immun. 1996 64:2062-2069		

Etapa 2 - Aplicando el segundo criterio, el orden de rangos de la expresión génica, las puntuaciones de genes de cada conjunto de datos se ordenaron de mayor según la relación de expresión (expresión en la condición experimental frente a las células en cultivo líquido en fase log). El gen de puntuación más alta recibió la puntuación máxima para ese conjunto de datos en particular (indicado en la columna 3 de la Tabla 1. (por ejemplo 5, 4,..., 1 punto)). La puntuación se redujo en 0,005 puntos por 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 clasificados en orden 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 los experimentos en los que se recogieron datos de múltiples puntos de tiempo, la puntuación máxima a través de todos los puntos de tiempo se usó como la puntuación final.

En la etapa 3, las puntuaciones para cada gen en cada una de las condiciones experimentales se recogieron en una base de datos Microsoft Access. Se agregaron campos de referencia para facilitar el establecimiento de prioridades, tal como el ID de Refseq, la función de Genbank, la nota GenBank, la clasificación en Tuberculist y los enlaces KEGG y Sanger Centre. Mediante la combinación de los datos de diferentes estudios y fuentes, se alcanzó una

visión de consenso acerca de los genes y las vías particulares más cruciales para la supervivencia en el estado latente.

5 En la etapa 4, se obtuvo una lista priorizada de dianas terapéuticas mediante la utilización de los 400 genes de puntuación superiores (-10 % del genoma) complementada con una análisis computacional y manual experto de las vías bioquímicas, enzimología, trazabilidad de fármacos, homología con los genes humanos y otros conocimientos previos. La gran mayoría de los genes de alta puntuación vienen del subconjunto donde dos o tres de los grupos se cruzan.

10 En la etapa 5, la identificación de la localización subcelular de las proteínas de *M. tuberculosis* codificadas por genes, se llevó a cabo con todo el genoma. La heurística utilizado para la predicción de las proteínas de membrana se describe en Chalker y col., J. Bacteriol. 2001 183:1259-1268. Se generaron perfiles de hidropatía promedio (H) (von Heijne G J. Mol. Biol. 1992 225:487-494) usando valores de hidropatía GES (Engelman DM y col., Annu. Rev. Biophys. Chem. 1986 15:321-353) ponderados usando una ventana trapezoidal. Usando un proceso similar a las etapas iniciales del algoritmo TopPred II (Claros MG y col. Comput Appl. Biosci. 1994 10:685-686), se predijeron segmentos transmembrana helicoidales (TMS) para cada secuencia de péptido mediante la selección de 19 aminoácidos centrado en el valor de H más alto (MaxH), enmascarando estos de una consideración adicional y repitiendo el proceso hasta que quedaran picos con una H de > 0,5. Las localizaciones subcelulares se asignaron según el pico del valor MaxH, el número de segmentos con una H de > 1,0, y valores de la distribución y el pico de H de la putativo TMS. Se eligió un punto de corte MaxH de 1,15 para maximizar la discriminación entre dos conjuntos de datos de ensayo de administración 34 SwissProtein contienen proteínas de transmembrana y citoplásmicas, respectivamente (Boyd D y col., Protein Sci. 1998, 7:201-205). Las proteínas con un MaxH de <1,15 se clasificaron como citoplásmica, mientras que aquellas con una MaxH de > 1,15 y al menos tres TMS posibles fueron clasificados como proteínas de membrana. Las proteínas ancladas se definieron como que tenían exactamente dos TMS, uno de partida antes del aminoácido (aa) 35 y uno que tiene un H de > 1,15 con el otro que tiene un H no inferior a 0,5. SignalP con ajustes Grampositivos se usó específicamente para *M. bacterium* para identificar las proteínas secretadas entre aquellas clasificadas como citoplasmáticas o "desconocidas" en el análisis heurístico (Nielsen H y col., Protein Eng. 1997 10: 1-6).

Rv1753c se clasificó en un puesto muy alto como antígeno de vacuna de acuerdo con varios criterios:

30 (i) Rv1753c está consistentemente refulada por aumento s a través de todos los modelos de latencia. Entre todo el conjunto de 3999 genes putuados en el meta-análisis, Rv1753c se clasificó en el puesto 116 como uno de los genes sobreexpresados del 10 % superior en todos los modelos de latencia. La puntuación de regulación por aumento para Rv1753c fue 13,29, que se compara favorablemente con la puntuación de gen superior de 22,28. Rv1753c no se había regulado por disminución en ningún modelo de la latencia, con una puntuación de 0 (en comparación con -18,13 para el gen más regulado por disminución).

35 (ii) Rv1753cse clasificó como esencial para el crecimiento de acuerdo con modelos de crecimiento *in vitro* de supervivencia de *M. tuberculosis* (puntuación 2,07, de una puntuación posible de 5).

(iii) la localización subcelular predijo que la proteína Rv1753c se secreta y, por tanto, tiene una exposición extracelular significativa, lo que indica idoneidad como diana de la vacuna.

Ejemplo 2 - Identificación del epítipo de Rv1753c

Procedimiento

40 La predicción del epítipo de linfocitos T se basa en los siguientes enfoques:

Predicción	Nombre	URL/Referencias
CD4 y CD8	Multipred	sitio web: antigen.i2r.a-star.edu.sg/Multipred/ Zhang,G.L., Khan,A.M., Srinivasan,K.N., August,J.T. y Brusica,V. (2005) "MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides" Nucleic Acids Res. 33, W172 - W179.
	SVMHC	sitio web: www.bs.informatik.uni-tuebingen.de/SVMHC "Prediction of MHC class I binding peptides, using SVMHC." Pierre Dönnès y Arne Elofsson in: BMC Bioinformatics 2002 3:25
CD4	ProPred	sitio web: 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 interno basado en: H. Bian, J. Hammer (2004) "Discovery of promiscuous HLA-II-restricted T cell epitopes with TEPITOPE." Methods 34: 468-75

(continuación)

Predicción	Nombre	URL/Referencias
CD8	nHLA	sitio web: www.imtech.res.in/raghava/nhlaped/ 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	sitio web: www.cbs.dtu.dk/services/NetCTL/ "An integrative approach to CTL epitope prediction. A combined algorithm integrating MHC-I binding, TAP transport efficiency, and 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	sitio web: 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	sitio web: www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm Hans-Georg Rammensee, Jutta Bachmann, Niels Nikolaus Emmerich, Oskar Alexander Bachor, Stefan Stevanovic: "SYFPEITHI: base de datos para ligandos y motivos peptídicos del MHC." Immunogenetics (1999) 50: 213-219
	PredTAP	sitio web: 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	website www.paproc2.de/paproc1/paproc1.html C. Kuttler, A.K. Nussbaum, T.P. Dick, H.-G. Rammensee, H. Schild, K.P. Haderler, "An algorithm for the prediction of proteasomal cleavages", J. Mol. Biol. 298 (2000), 417-429 A.K. Nussbaum, C. Kuttler, K.P. Haderler, H.-G. Rammensee, H. Schild, "PAProC: A Prediction Algorithm for Proteasomal Cleavages available on the WWW", Immunogenetics 53 (2001), 87-94

Resultados

Tabla 2 - Epítomos de células T CD4+ humanas Rv1753c putativos

Número de epítomo de CD4 putativo	Posición de aminoácidos	Secuencia del epítomo	SEQ ID NO:	Alelo de HLA
1	57	WQGASSSAM	SEQ ID NO: 30	DRB1_0401
2	81	VQAEQTAAQ	SEQ ID NO: 31	DRB1_0401
3	100	VKTAVVQPM	SEQ ID NO: 32	DRB1_0301, DRB1_1301
4	105	VQPMLVAAN	SEQ ID NO: 33	DRB1_1301
5	109	LVAANRADL	SEQ ID NO: 34	DRB1_0301, DRB1_0801, DRB1_1101, DRB1_1301, DRB1_1501
6	117	LVSLVMSNL	SEQ ID NO: 35	DRB1_1501
7	120	LVMSNLFGQ	SEQ ID NO: 36	DRB1_0401, DRB1_1301
8	140	YEQMWAADV	SEQ ID NO: 37	DRB1_0101
9	144	WAADVSAMS	SEQ ID NO: 38	DRB1_0401
10	172	LQNLAGLPA	SEQ ID NO: 39	DRB1_0101, DRB1_1101, DRB1_1501
11	261	FGNLGSNNV	SEQ ID NO: 40	DRB1_0401
12	291	FGNTGNNNI	SEQ ID NO: 41	DRB1_0401

(continuación)

Número de epítipo de CD4 putativo	Posición de aminoácidos	Secuencia del epítipo	SEQ ID NO:	Alelo de HLA
13	413	FLNAGNINT	SEQ ID NO: 42	DRB1_0401
14	453	LQFSITTPD	SEQ ID NO: 43	DRB1_0401
15	673	LTIPAGITI	SEQ ID NO: 44	DRB1_1501
16	725	FGIPFTLQF	SEQ ID NO: 45	DRB1_0401, DRB1_1101
17	731	LQFQTNVPA	SEQ ID NO: 46	DRB1_0401
18	733	FQTNVPALQ	SEQ ID NO: 47	DRB1_0401, DRB1_0801, DRB1_1101
19	770	YTLTGPIVI	SEQ ID NO: 48	DRB1_0101, DRB1_0401, DRB1_1101
20	782	FLPAFNIPG	SEQ ID NO: 49	DRB1_0401
21	866	LTIDPINLT	SEQ ID NO: 50	DRB1_0401
22	891	LTIDPINLT	SEQ ID NO: 51	DRB1_0301, DRB1_1501
23	954	YFNSSTAPS	SEQ ID NO: 52	DRB1_0401, DRB1_101
24	955	FNSSTAPSS	SEQ ID NO: 53	DRB1_0401
25	976	FGNNGSGLS	SEQ ID NO: 54	DRB1_0401
26	1000	YQNFGLSS	SEQ ID NO: 55	DRB1_0101, DRB1_0801, DRB1_1101, DRB1_1501
27	1003	FGGLSSGFS	SEQ ID NO: 56	DRB1_0401
28	1020	FANRGILPF	SEQ ID NO: 57	DRB1_0801
29	1025	ILPFSVASV	SEQ ID NO: 58	DRB1_1301
30	1037	FANIGTNLA	SEQ ID NO: 59	DRB1_0401, DRB1_1101

Tabla 3 - Epítipos de células T CD8+ humanas Rv1753c putativos

Número de epítipo de CD8 putativo	Posición de aminoácidos	Secuencia del epítipo	SEQ ID NO:	Alelo de HLA
1	2	NFSVLPPEI	SEQ ID NO: 60	A24
2	5	VLPPEINSA	SEQ ID NO: 61	A2, A_0201
3	6	LPPEINSAL	SEQ ID NO: 62	B7, B8, B_3501, B51
4	8	PEINSALIF	SEQ ID NO: 63	B44
5	9	EINSALIFA	SEQ ID NO: 64	A_0201, A_0301
6	18	GAGPEPMAA	SEQ ID NO: 65	A_0101, B_3501,

ES 2 647 321 T3

(continuación)

Número de epítipo de CD8 putativo	Posición de aminoácidos	Secuencia del epítipo	SEQ ID NO:	Alelo de HLA
7	20	GPEPMAAAA	SEQ ID NO: 66	B7, B_3501
8	22	EPMAAAATA	SEQ ID NO: 67	B7, B_0702, B8, B_3501, B51
9	26	AAATAWDGL	SEQ ID NO: 68	A1, B8, B_3501
10	28	ATAWDGLAM	SEQ ID NO: 69	B7
11	30	AWDGLAMEL	SEQ ID NO: 70	A1, A_2402, B44, Cw_0602
12	33	GLAMELASA	SEQ ID NO: 71	A_0101, A_0301, A2, A_0201
13	34	LAMELASAA	SEQ ID NO: 72	A3, A_0301, B51
14	48	VTSGLVGGA	SEQ ID NO: 73	A_0101, A_0301
15	64	AMAAAAAPY	SEQ ID NO: 74	A1, A3, A_0301, A_0101, B_4403
16	66	AAAAAPYAA	SEQ ID NO: 75	A_0301, B_3501
17	68	AAAPYAOWL	SEQ ID NO: 76	A1, A24, B_3501, B51
18	69	AAPYAOWLA	SEQ ID NO: 77	A1, A_0301, B_3501
19	70	APYAOWLAA	SEQ ID NO: 78	A3, A_0301, B7, B_0702, B8, B_3501
20	72	YAAWLAAAA	SEQ ID NO: 79	A_0301, B8, B_3501
21	73	AAWLAAAAV	SEQ ID NO: 80	A2, A_0201, B7, B51
22	75	WLAAAAVQA	SEQ ID NO: 81	A2, A3, A_0201
23	82	QAEQTAAQA	SEQ ID NO: 82	A1, A_0301
24	83	AEQTAAQAA	SEQ ID NO: 83	B44, B_4403
25	86	TAAQAAAMI	SEQ ID NO: 84	A3, B8, B51
26	91	AAMIAEFEA	SEQ ID NO: 85	A_0201, A_0301, B_3501
27	92	AMIAEFEAV	SEQ ID NO: 86	A2, A_0201
28	95	AEFEAVKTA	SEQ ID NO: 87	B44
29	97	FEAVKTAVV	SEQ ID NO: 88	B8, B44
30	98	EAVKTAVVQ	SEQ ID NO: 89	B8, B_3501
31	101	KTAWQPML	SEQ ID NO: 90	A_0101, A_0201
32	106	QPMLVAANR	SEQ ID NO: 91	A3, B7, B_0702, B_3501, B51

ES 2 647 321 T3

(continuación)

Número de epítipo de CD8 putativo	Posición de aminoácidos	Secuencia del epítipo	SEQ ID NO:	Alelo de HLA
33	107	PMLVAANRA	SEQ ID NO: 92	A2, A_0201, B8
34	109	LVAANRADL	SEQ ID NO: 93	B7
35	112	ANRADLVSL	SEQ ID NO: 94	B7, B44
36	114	RADLVSLVM	SEQ ID NO: 95	B7, B_3501
37	118	VSLVMSNLF	SEQ ID NO: 96	A24, A_0101
38	124	NLFGQNAPA	SEQ ID NO: 97	A2
39	130	APAIAAIEA	SEQ ID NO: 98	B7, B_3501
40	132	AIAAIEATY	SEQ ID NO: 99	A1, A_0101, A3, A_0301
41	138	ATYEQMWAA	SEQ ID NO: 100	A_0101, A2, A_0301
42	142	QMWAADVSA	SEQ ID NO: 101	A2, A_0201
43	150	AMSAYHAGA	SEQ ID NO: 102	A2, A_0201
44	152	SAYHAGASA	SEQ ID NO: 103	B7, B_3501
45	153	AYHAGASAI	SEQ ID NO: 104	A1, A_0201, A3, A_2402, A_0301, A24
46	157	GASAIASAL	SEQ ID NO: 105	B7, B_3501
47	160	AIASALSPF	SEQ ID NO: 106	A_0301, B7
48	164	ALSPFSKPL	SEQ ID NO: 107	A_0101, A2, A_0201
49	167	PFSKPLQNL	SEQ ID NO: 108	A24, A_2402, Cw_0401
50	170	KPLQNLAGL	SEQ ID NO: 109	B7, B_3501, B51
51	174	NLAGLPAWL	SEQ ID NO: 110	A2, A_0201, B7, Cw_0602
52	175	LAGLPAWLA	SEQ ID NO: 111	A_0101, A_0301
53	178	LPAWLASGA	SEQ ID NO: 112	B7, B_3501
54	181	WLASGAPAA	SEQ ID NO: 113	A_0201
55	185	GAPAAAMTA	SEQ ID NO: 114	A3, A_0301, B8
56	186	APAAAMTAA	SEQ ID NO: 115	A3, B_3501, B7
57	189	AAMTAAAGI	SEQ ID NO: 116	A1, A_2402, B51
58	192	TAAAGIPAL	SEQ ID NO: 117	B7, B51, Cw_0602
59	193	AAAGIPALA	SEQ ID	A_0101, A_0301

ES 2 647 321 T3

(continuación)

Número de epítopo de CD8 putativo	Posición de aminoácidos	Secuencia del epítopo	SEQ ID NO:	Alelo de HLA
			NO: 118	
60	199	ALAGGPTAI	SEQ ID NO: 119	A1, A_0101, A2, A_0201, A_0301
61	201	AGGPTAINL	SEQ ID NO: 120	A1, A24, B51
62	203	GPTAINLGI	SEQ ID NO: 121	A_2402, B7, B_0702, B8, B_3501, B51
63	206	AINLGIANV	SEQ ID NO: 122	A2, A_0201
64	231	NANLGNYNF	SEQ ID NO: 123	A24, B_3501
65	236	NYNFGSGNF	SEQ ID NO: 124	A24
66	263	NLGSNNVGV	SEQ ID NO: 125	A2, A_0201
67	383	SLNTGSYNM	SEQ ID NO: 126	A2
68	408	NANTGFLNA	SEQ ID NO: 127	A_0101, A_0301
69	413	FLNAGNINT	SEQ ID NO: 128	A2
70	418	NINTGVFNI	SEQ ID NO: 129	A_0201, A_0301
71	447	GVGQGS LQF	SEQ ID NO: 130	B7, B_3501
72	456	SITTPDLTL	SEQ ID NO: 131	A_0101, A_0201, A_0301
73	459	TPDLTL PPL	SEQ ID NO: 132	B7, B_3501, B51
74	461	DLTL PPLQI	SEQ ID NO: 133	A_0101, A_0201
75	466	PLQIPGISV	SEQ ID NO: 134	A_0201
76	469	IPGISVPAF	SEQ ID NO: 135	B7, B_3501
77	471	GISVPAFSL	SEQ ID NO: 136	A_0101, A_0201, A_0301, B44
78	474	VPAFSLPAI	SEQ ID NO: 137	B7, B51
79	476	AFSLPAITL	SEQ ID NO: 138	A_0201, A24, B7
80	479	LPAITLPSL	SEQ ID NO: 139	A24, B7, B_3501, B51, B_0702, Cw_0401, Cw_0602
81	481	AITLPSLNI	SEQ ID NO: 140	A_0101, A_0301
82	483	TLPSLNIPA	SEQ ID NO: 141	A2, A_0201, A_0301
83	484	LPSLNIPAA	SEQ ID NO: 142	B7, B_3501, B51
84	492	ATTPANITV	SEQ ID NO: 143	A1, A_0101, A2, A_0201
85	494	TPANITVGA	SEQ ID NO: 144	B7, B_3501

ES 2 647 321 T3

(continuación)

Número de epítipo de CD8 putativo	Posición de aminoácidos	Secuencia del epítipo	SEQ ID NO:	Alelo de HLA
86	497	NITVGAFSL	SEQ ID NO: 145	A2, A_0201, A_0301, A24
87	502	AFSLPGLTL	SEQ ID NO: 146	A24, A_2402, B7
88	505	LPGLTLPSL	SEQ ID NO: 147	B7, B_3501, B51, B_0702, Cw_0602
89	509	TLPSLNIPA	SEQ ID NO: 148	A2
90	518	ATTPANITV	SEQ ID NO: 149	A1, A2
91	523	NITVGAFSL	SEQ ID NO: 150	A2, A24, A_0201
92	528	AFSLPGLTL	SEQ ID NO: 151	A_2402, B7
93	531	LPGLTLPSL	SEQ ID NO: 152	B7, B_0702, B_3501, B51, Cw_0602
94	535	TLPSLNIPA	SEQ ID NO: 153	A2
95	544	ATTPANITV	SEQ ID NO: 154	A1, A2
96	549	NITVGAFSL	SEQ ID NO: 155	A2, A24, A_0201
97	554	AFSLPGLTL	SEQ ID NO: 156	A_2402, B7
98	557	LPGLTLPSL	SEQ ID NO: 157	B7, B_0702, B_3501, B51, Cw_0602
99	561	TLPSLNIPA	SEQ ID NO: 158	A2
100	570	ATTPANITV	SEQ ID NO: 159	A1, A2
101	575	NITVGAFSL	SEQ ID NO: 160	A2, A24, A_0201
102	580	AFSLPGLTL	SEQ ID NO: 161	A_2402, B7
103	583	LPGLTLPSL	SEQ ID NO: 162	B7, B_0702, B_3501, B51, Cw_0602
104	587	TLPSLNIPA	SEQ ID NO: 163	A2
105	596	ATTPANITV	SEQ ID NO: 164	A1, A2
106	601	NITVGAFSL	SEQ ID NO: 165	A2, A24
107	609	LPGLTLPSL	SEQ ID NO: 166	B7, B_0702, B_3501, B51, Cw_0602
108	622	ATTPANITV	SEQ ID NO: 167	A1, A2
109	625	PANITVSGF	SEQ ID NO: 168	A24, B_3501
110	627	NITVSGFQL	SEQ ID NO: 169	A_0201, A_0301
111	635	LPPLSIPSV	SEQ ID NO: 170	B7, B51
112	636	PPLSIPVA	SEQ ID	B7, B_3501

ES 2 647 321 T3

(continuación)

Número de epítipo de CD8 putativo	Posición de aminoácidos	Secuencia del epítipo	SEQ ID NO:	Alelo de HLA
			NO: 171	
113	640	IPSVAIPPV	SEQ ID NO: 172	B7, B51
114	645	IPPVTVPPI	SEQ ID NO: 173	B7, B51
115	650	VPPITVGAF	SEQ ID NO: 174	B7, B_3501
116	662	PLQIPEVTI	SEQ ID NO: 175	A_0201
117	665	IPEVTIPQL	SEQ ID NO: 176	B7, B_3501, B51
118	669	TIPQLTIPA	SEQ ID NO: 177	A_0201, A_0301
119	673	LTIPAGITI	SEQ ID NO: 178	A_0101, A_0201, B51
120	678	GITIGGFSL	SEQ ID NO: 179	A_0101, A_0201, A_0301
121	686	LPAIHTQPI	SEQ ID NO: 180	B7, B8, B51
122	688	AIHTQPITV	SEQ ID NO: 181	A_0101, A_0201, A_0301
123	693	PITVGQIGV	SEQ ID NO: 182	A_0201
124	698	QIGVGQFGL	SEQ ID NO: 183	A_0201
125	705	GLPSIGWDV	SEQ ID NO: 184	A2, A_0201
126	706	LPSIGWDVF	SEQ ID NO: 185	B7, B_3501
127	712	DVFLSTPRI	SEQ ID NO: 186	A_0201
128	714	FLSTPRITV	SEQ ID NO: 187	A_0101, A2, A_0201, B8
129	717	TPRITVPAF	SEQ ID NO: 188	B7, B8, B_3501
130	725	FGIPFTLQF	SEQ ID NO: 189	A_0201, B8, B_3501
131	729	FTLQFQTNV	SEQ ID NO: 190	A_0101, A2, A_0201
132	732	QFQTNVPAL	SEQ ID NO: 191	A24
133	739	ALQPPGGGL	SEQ ID NO: 192	A_0101, A_0201
134	747	LSTFTNGAL	SEQ ID NO: 193	A_0101, B7
135	748	STFTNGALI	SEQ ID NO: 194	A_0101, A_0201, A_0301, A24
136	749	TFTNGALIF	SEQ ID NO: 195	A24
137	754	ALIFGEFDL	SEQ ID NO: 196	A_0101, A2, A_0201
138	758	GEFDLPQLV	SEQ ID NO: 197	B44

ES 2 647 321 T3

(continuación)

Número de epítipo de CD8 putativo	Posición de aminoácidos	Secuencia del epítipo	SEQ ID NO:	Alelo de HLA
139	762	LPQLWHPY	SEQ ID NO: 198	B7, B51
140	764	QLVVHPYTL	SEQ ID NO: 199	A_0101, A2, A_0201, B8
141	768	HPYTLTGPI	SEQ ID NO: 200	B7, B8, B51
142	770	YTLTGPIVI	SEQ ID NO: 201	A_0101, A2, A_0201
143	774	GPIVIGSFF	SEQ ID NO: 202	A24, B7, B_3501
144	775	PIVIGSFFL	SEQ ID NO: 203	A_0101, A_0301
145	780	SFFLPAFNI	SEQ ID NO: 204	A24
146	783	LPAFNIPGI	SEQ ID NO: 205	B7, B51
147	788	IPGIDVPAI	SEQ ID NO: 206	B7, B51
148	790	GIDVPAINV	SEQ ID NO: 207	A_0101, A_0201, A_0301
149	793	VPAINVDGF	SEQ ID NO: 208	B7, B_3501
150	795	AINVDGFTL	SEQ ID NO: 209	A_0101, A_0201, A_0301, B44
151	802	TLPQITTPA	SEQ ID NO: 210	A2
152	803	LPQITTPAI	SEQ ID NO: 211	B7, B8, B51
153	808	TPAITTPEF	SEQ ID NO: 212	B7, B_3501
154	810	AITTPEFAI	SEQ ID NO: 213	A_0101, A_0201, A_0301
155	813	TPEFAIPPI	SEQ ID NO: 214	B7, B51
156	818	IPPIGVGGF	SEQ ID NO: 215	B7, B_3501
157	820	PIGVGGFTL	SEQ ID NO: 216	A_0101, A_0201, A_0301
158	828	LPQITTQEI	SEQ ID NO: 217	B7, B51
159	829	PQITTQEII	SEQ ID NO: 218	A_0101
160	835	EIITPELTI	SEQ ID NO: 219	A_0101, A_0201, A_0301
161	838	TPELTINSI	SEQ ID NO: 220	B7, B51
162	840	ELTINSIGV	SEQ ID NO: 221	A_0201
163	845	SIGVGGFTL	SEQ ID NO: 222	A_0201, A_0301
164	853	LPQITTPPI	SEQ ID NO: 223	B7, B51
165	858	TPPITTPPL	SEQ ID	B7, B_3501, B51

ES 2 647 321 T3

(continuación)

Número de epítipo de CD8 putativo	Posición de aminoácidos	Secuencia del epítipo	SEQ ID NO:	Alelo de HLA
			NO: 224	
166	860	PITTPPLTI	SEQ ID NO: 225	A_0101, A_0201, A_0301
167	863	TPPLTIDPI	SEQ ID NO: 226	B7, B51
168	870	PINLTGFTL	SEQ ID NO: 227	A_0101, A_0301
169	913	TPPLTIEPI	SEQ ID NO: 228	B7, B51
170	915	PLTIEPIGV	SEQ ID NO: 229	A_0101, A_0201
171	918	IEPIGVGGF	SEQ ID NO: 230	B44
172	929	PPLTVPGIH	SEQ ID NO: 231	B_3501
173	930	PLTVPGIHL	SEQ ID NO: 232	A_0101, A_0201
174	935	GIHLPSTTI	SEQ ID NO: 233	A_0101, A_0301
175	937	HLPSTTIGA	SEQ ID NO: 234	A2
176	938	LPSTTIGAF	SEQ ID NO: 235	B7, B8, B_3501
177	946	FAIPGGPGY	SEQ ID NO: 236	A_0101, A_0301, B_3501
178	958	STAPSSGFF	SEQ ID NO: 237	A1, A_0101, A24
179	986	WFNTNPAGL	SEQ ID NO: 238	A24
180	1002	NFGGLSSGF	SEQ ID NO: 239	A24
181	1005	GLSSGFSNL	SEQ ID NO: 240	A_0101, A2, A_0201
182	1012	NLGSGVSGF	SEQ ID NO: 241	A_0201
183	1020	FANRGILPF	SEQ ID NO: 242	B8, B_3501
184	1022	NRGILPFSV	SEQ ID NO: 243	A2, A24, B51
185	1025	ILPFSVASV	SEQ ID NO: 244	A2, A_0201
186	1026	LPFSVASVV	SEQ ID NO: 245	B7, B51
187	1029	SVASVVSGF	SEQ ID NO: 246	A24, B7
188	1036	GFANIGTNL	SEQ ID NO: 247	A24, Cw_0401, Cw_0602

Como se puede ver en las Tablas 2 y 3, Rv1753c contiene un número de epítipos de células T CD4+ y CD8 T predichos. Por otro lado, esta información sugiere que la proteína lleva epítipos que pueden ser reconocidos por los HLA existentes en todo el mundo (es decir HLA de individuos caucásicos, africanos, asiáticos o latinoamericanos, véase el sitio web en www.allelefrequencies.net).

Ejemplo 3 - Homólogos de H37Rv

Se identificaron las secuencias de Rv1753c de un número de cepas de *M. tuberculosis* y BCG utilizando búsquedas BLASTP de GenBank (número de registro de referencia de H37Rv YP_177830.1):

Cepa	Número de acceso	% de identidad
CDC1551	NP_336255.1	95
F11	YP_001287714.1	80
Haarlem	ZP_02247061.1	82
C	ZP_00878894.1	96
BCG	YP_977884.1	83

5 La alineación de las secuencias homólogas indica un alto nivel de identidad en las regiones N-terminal y C-terminal, produciéndose la mayoría de la variación en una región central de unión.

ENSAYOS BIOLÓGICOS**Cuantificación de las respuestas de linfocitos T a Rv1753c**

10 Los polipéptidos pueden examinarse para determinar su capacidad para activar los linfocitos T (inducción de la proliferación y/o la producción de citocinas) en células mononucleares de sangre periférica (PBMC) o en preparaciones de sangre completa de individuos infectados (tal como infectados de forma latente).

Los individuos infectados de forma latente se identifican generalmente por una prueba cutánea que tiene un diámetro superior a 10 mm y sin síntomas, con ningún cultivo positivo para Mtb, con un esputo negativo negativo y sin lesión (detectado mediante radiografía de tórax).

15 Una gama de ensayos *in vitro* se puede utilizar sobre la base de muestras de PBMC o sangre entera: después de la reestimulación en presencia del antígeno (o variante/fragmento inmunogénico del mismo según sea apropiado) la proliferación de las células puede determinarse (medido por CFSE/citometría de flujo) o la producción de citocinas cuantificadas (presente en el sobrenadante de las células cultivadas y medido mediante ELISA, o, después de la tinción intracelular de los linfocitos T CD4 y CD8 y análisis mediante citometría de flujo).

20 Por ejemplo, las muestras de PBMC se pueden obtener a partir de sangre entera heparinizada mediante centrifugación en gradiente de densidad Ficoll-Hypaque siguiendo procedimientos estándar. A continuación, las células pueden lavarse criopreservarse en nitrógeno líquido hasta su análisis (para más detalles véase Lalvani A y col. J. Infect. Dis. 1999, 180:1656-1664).

Proliferación de linfocitos T

25 La respuesta inmunitaria específica se puede caracterizar por la realización de análisis de proliferación de linfocitos usando timidina tritiada. Esta técnica evalúa la expansión celular tras 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.

30 Más adecuadamente, la proliferación de linfocitos se puede realizar usando el éster de succinimidilo de diacetato de carboxyfluoresceína (CFSE). El CFSE se acopla de forma espontánea e irreversible en las proteínas intracelulares y de superficie celular por reacción con cadenas laterales de lisina y otros grupos amina disponibles. Cuando las células de linfocitos se dividen, el marcaje con CFSE se distribuye por igual entre las células hijas, que son, por lo tanto, la mitad de fluorescentes que los padres. Como resultado, la reducción a la mitad de la intensidad de fluorescencia celular marca cada generación sucesiva en una población de células proliferantes y es seguida fácilmente por citometría de flujo (para más detalles véase Hodgkins, PD y col. J. Exp. Med. 1996, 184:277-281).

35 En la práctica, después de la descongelación, las PBMC se pueden lavar y se tiñen con CFSE antes de cultivar (2×10^5 células) durante 72 horas con 10 ug/ml de antígeno en medios de cultivo (RPMI-1640 suplementado con glutamina, aminoácidos no esenciales, piruvato y suero AB inactivado por calor humano). A continuación, se pueden recoger las células y su fenotipo se caracteriza mediante tinción de superficie para identificar linfocitos T de memoria CD8 y CD4+. Posteriormente, el análisis de citometría de flujo puede usarse para indicar la extensión de la proliferación de linfocitos en respuesta a cada antígeno (proporción de células con una disminución de la intensidad de CFSE a la estimulación *in vitro*).

Producción de citocinas

45 La producción de IFN- γ (o la producción de otras citocinas tales como por ejemplo, IL2, TNF- α , IL5, IL12, etc.) se pueden medir utilizando un ensayo de inmunoabsorción ligado a enzimas (ELISA). Las placas de ELISA se pueden recubrir con un anticuerpo monoclonal de ratón dirigido a IFN- γ humano (PharMingen, San Diego, CA) en

PBS durante cuatro horas a temperatura ambiente. A continuación se bloquearon los pocillos con PBS que contenía 5 % de leche desgrasada desecada durante 1 hora a temperatura ambiente. Las placas se lavan a continuación, por ejemplo, seis veces en PBS/0,2% de Tween-20 y las muestras se diluyeron a 1: 2 en medio de cultivo en las placas de ELISA se incuban durante la noche a temperatura ambiente. Las placas se lavan de nuevo y a cada pocillo se puede añadir un suero de IFN- γ anti-humano policlonal de conejo, por ejemplo, diluido a 1:3000 en/10% de suero normal de cabra. A continuación, las placas se incuban durante dos horas a temperatura ambiente, se lavan y se puede añadir IgG anti-conejo acoplada a peroxidasa de rábano (Sigma Chemical So., St. Louis, MO), por ejemplo, a una dilución de 1:2000 en PBS/5 % de leche desgrasada desecada. Después de una incubación de dos horas adicionales a temperatura ambiente, las placas se lavan y se añade sustrato de TMB. La reacción se puede detener tras 20 minutos con ácido sulfúrico 1N. La densidad óptica se puede determinar a continuación a 450 nm usando 570 nm como longitud de onda de referencia. Típicamente, las fracciones que resultan en ambos duplicados que dan una DO dos veces mayor que la DO media de las células cultivadas en medio solo pueden considerarse positivas.

Ejemplo 4 - Immunogenicidad en ratones CB6F1

La inmunogenicidad del antígeno se evaluó en ratones CB6F1 (primer cruce de generaciones de ratones BALB/c y ratones C57BL/6).

Se inmunizó a los ratones CB6F1 por vía intramuscular tres veces (el día 0, día 14 y día 28) con 0,5 ug o 2 ug de antígeno de proteína en combinación con el AS01 E sistema adyuvante (una formulación liposomal que comprende adyuvante 3D-MPL y QS21).

El diseño experimental fue el siguiente:

Grupo	Día 0	Día 14	Día 28
1	2 ug de Rv1753c/AS01E	2 ug de Rv1753c/AS01E	2 ug de Rv1753c/AS01E
2	0,5 ug de Rv1753c/AS01E	0,5 ug de Rv1753c/AS01E	0,5 ug de Rv1753c/AS01E

En cada grupo de protocolo se usó un total de 24 ratones.

Se recogieron linfocitos de sangre periférica (PBL) y se combinaron el día 21 (es decir, 7 días después de la segunda inmunización) y el día 35 (es decir, 7 días después de la tercera inmunización) y se midieron las respuestas de los linfocitos T CD4 y CD8 específicas de antígeno (determinadas mediante la producción de los linfocitos T CD4 o CD8 de IL-2 y/o IFN gamma y/o TNF-alfa) mediante citometría de flujo después la reestimulación *in vitro* durante la noche con grupos de péptidos de 15 aminoácidos que abarcan las secuencias de interés.

La detección de células T de ratón que expresan IL-2 y/o IFN-gamma y/o TNF-alfa se realizó mediante el uso de amplificación *in vitro* dirigida por antígeno a corto plazo de la expresión de citocinas.

En resumen, se añadió solución PharmLyse (BD-Pharmingen) a la sangre periférica de ratón heparinizada con el fin de lisar los glóbulos rojos. Los PBL (linfocitos de sangre periférica) obtenidos se lavaron y después se incubaron en presencia de un grupo de péptidos 15 aminoácidos, solapados por 11 aminoácidos, que cubrían la secuencia del antígeno de interés y de 1 ug/ml de anticuerpos frente a CD28 y CD49d (BD-Pharmingen). Cada péptido de 15 aminoácidos se utilizó a una concentración final de 1 ug/ml. Los controles de medio también se estimularon con anticuerpos frente a CD28 y CD49d.

El compuesto de bloqueo de la secreción de citocinas brefeldina-A (BD-Pharmingen) se añadió 2 horas después del inicio de los cultivos a 37 °C, 5 % de CO₂ y las células se mantuvieron a 37 °C, 5 % de CO₂ durante 4 horas adicionales seguidas por incubación durante la noche a + 4 °C.

Las células se recogieron a continuación y se tiñeron con anti-CD4 acoplado a anticuerpos Pacific Blue (BD - clon RM4-5, BD-Pharmingen) y alfa anti-CD8 acopladas a peridina clorofila A proteína (PerCp) o cianina 5,5 (Cy 5.5.) (clon 53-6.7, BD-Pharmingen).

A continuación se lavaron las células, se fijaron, se permeabilizaron (kit Cytofix-cytoperm, BD-Pharmingen) y se tiñeron con anticuerpos anti-IFN-g acoplados con alofocianina (clon XMG1.2, BDPharmingen), anticuerpos anti-IL-2 acoplado a isotiocianato de fluoresceína (FITC) (clon JES 6-5H4, Beckman Coulter) y anticuerpos anti-TNF-alfa acoplados a ficoeritrina (PE) (clon MP6-XT22, BDPharmingen). 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 la subpoblación de CD8 +.

Para más 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 control negativo, algunas células también se cultivaron durante la noche *in vitro* en medio de cultivo (sin estimular). Las respuestas específicas de antígeno se calcularon restando la respuesta citocina promedio producida por células no estimuladas de la respuesta promedio de citocinas producida por las células estimuladas con

péptidos.

En cada punto de tiempo y para cada grupo, se recogieron los datos a partir de 4 grupos de 6 ratones cada uno. Los datos siguientes se presentan como el % de linfocitos T CD4 o CD8 productoras de IL-2 y/o IFN-gamma y/o TNF-alfa. Cada grupo individual de ratones se representa gráficamente (triángulos), así como el valor promedio del grupo (barra).

La Figura 1 muestra que el día 21 (es decir, 7 días después de la segunda inmunización), las respuestas de linfocitos T CD4 y CD8 específicas de Rv1753c detectadas en ratones inmunizados con cualquiera de las dosis de Rv1753c/AS01 E. Los niveles de respuestas de linfocitos T específicos de Rv1753c son más altos en ratones inmunizados con 2 ug de Rv1753c/AS01 E que en los ratones inmunizados con 0,5 ug de Rv1753c/AS01 E.

La Figura 2 muestra el perfil de citocinas de respuesta de linfocitos T CD4 de los PBL estimulados por conjuntos de péptidos Rv1753c (no se elimina medio) el día 21 (es decir, 7 días después de la segunda inmunización).

La Figura 3 muestra el perfil de citocinas de respuesta de linfocitos T CD8 de los PBL estimulados por conjuntos de péptidos Rv1753c (no se elimina medio) el día 21 (es decir, 7 días después de la segunda inmunización).

La Figura 4 muestra que el día 35 (es decir, 7 días después de la tercera inmunización), las respuestas de linfocitos T CD4 y CD8 específicas de Rv1753c detectadas en ratones inmunizados con cualquiera de las dosis de Rv1753c/AS01 E. Los niveles de respuestas de linfocitos T específicos de Rv1753c son más altos en ratones inmunizados con 2 ug de Rv1753c/AS01 E que en los ratones inmunizados con 0,5 ug de Rv1753c/AS01 E. La tercera inyección aumenta la respuesta de linfocitos T CD4 observada el día 21 a la dosis menor de 0,5 ug pero no a la dosis más alta de 2 ug de antígeno. La respuesta de células T CD8 específicas de antígeno es menor el día 35 que el día 21.

La Figura 5 muestra el perfil de citocinas de respuesta de linfocitos T CD4 de los PBL estimulados por conjuntos de péptidos Rv1753c (no se elimina medio) el día 35 (es decir, 7 días después de la tercera inmunización). Debido a dificultades técnicas, los datos para el tercer grupo de células a la dosis de 2 ug no están disponibles.

La Figura 6 muestra el perfil de citocinas de respuesta de linfocitos T CD8 de los PBL estimulados por conjuntos de péptidos Rv1753c (no se elimina medio) el día 35 (es decir, 7 días después de la tercera inmunización). Debido a dificultades técnicas, los datos para el tercer grupo de células a la dosis de 2 ug no están disponibles.

Ejemplo 5 - Immunogenicidad en ratones C57BL/6

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

Se inmunizó a ratones C57BL/6 por vía intramuscular tres veces (el día 0, el día 14 y el día 28) con 1 ug o 4 ug de antígeno de proteína en combinación con un el sistema adyuvante AS01 E (una formulación de adyuvante liposómica que comprende 3D-MPL y QS21).

El diseño experimental fue el siguiente:

Grupo	Día 0	Día 14	Día 28
1	4 ug de Rv1753c/AS01E	4 ug de Rv1753c/AS01E	4 ug de Rv1753c/AS01E
2	1 ug de Rv1753c/AS01E	1 ug de Rv1753c/AS01E	1 ug de Rv1753c/AS01E

Se recogieron linfocitos de sangre periférica (PBL) y se combinaron el día 21 (es decir, 7 días después de la segunda inmunización) y el día 35 (es decir, 7 días después de la tercera inmunización) y se midieron las respuestas de los linfocitos T CD4 y CD8 específicas de antígeno (determinadas mediante la producción de los linfocitos T CD4 o CD8 de IL-2 y/o IFN gamma y/o TNF-alfa) mediante citometría de flujo después la reestimulación *in vitro* durante la noche con grupos de péptidos de 15 aminoácidos que abarcan las secuencias de interés. El procedimiento seguido fue como se ha descrito anteriormente.

Como control negativo, algunas células también se cultivaron durante la noche *in vitro* en medio de cultivo (sin estimular). Las respuestas específicas de antígeno se calcularon restando la respuesta citocina promedio producida por células no estimuladas de la respuesta promedio de citocinas producida por las células estimuladas con péptidos.

En cada punto de tiempo y para cada grupo, se recogieron los datos a partir de 4 grupos de 6 ratones cada uno. Los datos siguientes se presentan como el % de linfocitos T CD4 o CD8 productoras de IL-2 y/o IFN-gamma y/o TNF-alfa. Cada grupo individual de ratones se representa gráficamente (triángulos), así como el valor promedio del grupo (barra).

La Figura 7 muestra que el día 21 (es decir, 7 días después de la segunda inmunización), las respuestas de linfocitos T CD4 y CD8 específicas de Rv1753c detectadas en ratones inmunizados con cualquiera de las dosis de Rv1753c/AS01 E. Los niveles de respuestas de linfocitos T CD4 específicos de Rv1753c son similares con respecto

a la dosis inmunizante de Rv1753c. Por el contrario, los ratones inmunizados con 1 ug de Rv1753c/AS01 E mostraron respuestas de células T CD8-Rv1753c específicas más altas que los ratones inmunizados con 4 ug Rv1753c/AS01 E.

5 La Figura 8 muestra el perfil de citocinas de respuesta de linfocitos T CD4 de los PBL estimulados por conjuntos de péptidos Rv1753c (no se elimina medio) el día 21 (es decir, 7 días después de la segunda inmunización). Debido a dificultades técnicas, los datos para el tercer y cuarto grupos de células a la dosis de 1 ug no están disponibles.

La Figura 9 muestra el perfil de citocinas de respuesta de linfocitos T CD8 de los PBL estimulados por conjuntos de péptidos Rv1753c (no se elimina medio) el día 21 (es decir, 7 días después de la segunda inmunización). Debido a dificultades técnicas, los datos para el tercer y cuarto grupos de células a la dosis de 1 ug no están disponibles.

10 Los datos inmunológicos para el día 35 aún no estaban disponibles en el momento de elaboración del presente documento.

Ejemplo 6 Reconocimiento in vitro por PBMC de seres humanos con TB latente

15 Se realizaron experimentos con el fin de evaluar la respuesta de células T de sangre periférica específico para el antígeno al antígeno de la invención en 4 adultos sanos no expuestos anteriormente a 8 TB (prueba cutánea de PPD = 0 mm y adultos sanos infectados de forma latente con TB (prueba cutánea de PPD= 15 mm) o superior, de Sudáfrica

Datos de prueba cutánea de PPD

Número ID del individuo	Diámetro de la 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 en células aisladas mononucleares de sangre periférica (PBMC) mediante ensayo de tinción intracelular de citocinas (ICS).

20 La ICS realizada fue una adaptación de la metodología descrita previamente (véase Von Eschen y col., Hum. Vaccin. 2009 5(7)). Se estimularon los PBMC in vitro mediante un grupo de péptidos de 15 aminoácidos, solapando con 11 aminoácidos, que cubre la totalidad de la secuencia del antígeno de interés. Las células se estimularon con péptidos durante 2 horas, se cultivaron adicionalmente durante la noche en presencia de Brefeldina A, se procesaron para determinar la ICS y se analizaron usando citometría de flujo. Se midieron las frecuencias de los linfocitos T
 25 CD3+CD4+ o CD3+CD8+ específicos de antígeno que expresan IFN-gamma y/o TNF-alfa y/o IL-17. Las respuestas de células estimuladas con medio se restaron de las respuestas obtenidas en las células estimuladas por grupos de péptidos.

ICS: anticuerpos

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

35 Los resultados se presentan como el número de células T CD3 + CD4 + específicas de antígeno que expresan TNF-alfa e IFN-gamma, por millón de linfocitos T CD3 + CD4 + ya que estas células representan la principal población de las células CD4 T específicas de antígeno (el fondo se elimina nivel de respuesta debido a la media). No se detectaron linfocitos T CD8 + CD3 + específicos de antígeno. La Figura 10 muestra que una respuesta de células T

CD4 específicas de antígeno se mide en 7 de los 8 individuos con infección latente (no en número individual 60) en comparación con la respuesta de linfocitos T CD4 no específica medida en los individuos no expuestos previamente.

5 En conclusión, puede señalarse que el antígeno Rv1753c es capaz de provocar una respuesta inmunitaria tanto en ratones como CB6F1 como en ratones C57BL/6. Por otro lado, Además, el perfil de la producción de citocinas indica que una gran proporción de células T específicas de antígeno expresan una pluralidad de Th1 citocinas asociadas (es decir, una respuesta de células T polifuncional se provoca). e. Es importante destacar que los linfocitos T específicos de antígeno tanto CD4 y CD8 están presentes después de la inmunización, las células CD8 pueden ser particularmente importantes en un escenario de TB latente. La relevancia si Rv1753c para la infección humana es confirmada por el alto nivel de reconocimiento en los individuos con infección latente de Sudáfrica y la ausencia de respuestas en los sujetos no expuestos previamente. Por lo tanto, Rv1753c se puede esperar a ser fundamental en la prevención, el tratamiento y el diagnóstico de la infección tuberculosa latente.

10 A lo largo de esta memoria descriptiva y de las reivindicaciones que siguen, a menos que el contexto requiera lo contrario, se entenderá que la palabra "comprenden" y variaciones tales como "comprende" y "que comprende" implica la inclusión de un número entero o etapa o grupo de números enteros o grupo de etapas indicados, pero no la exclusión de ningún otro número entero o etapa o grupo de números enteros o de etapas.

15 LISTADO DE SECUENCIAS

20 <110> GlaxoSmithKline Biologicals S.A.
Glaxo Group Limited
Mettens, Pascal
Brown, James
Murphy, Dennis

<120> COMPOSICIONES Y PROCEDIMIENTOS NOVEDOSOS

<130> VB63087PCT

25 <150> US61/083692
<151> 25/07/2008

<160> 249

<170> PatentIn versión 3.5

30 <210> 1
<211> 1053
<212> PRT
<213> Mycobacterium tuberculosis

<220>

<221> MISC_FEATURE

<223> Cepa H37Rv

35 <400> 1

ES 2 647 321 T3

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

ES 2 647 321 T3

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		

ES 2 647 321 T3

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			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						

ES 2 647 321 T3

				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	

ES 2 647 321 T3

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> 2
 <211> 3162
 <212> ADN
 <213> Mycobacterium tuberculosis

5

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

<400> 2

atgaattttt ctgtactgcc gccggagatc aattcagcgc tgatattcgc cggggcaggg 60
 ccggaaccga tggcggcgcc cgcgacggcc tgggacgggt tggccatgga attggcctcg 120
 gccgcagcct ctttcggctc agtgacatcc ggactcgtgg gcggggcgtg gcagggcgcg 180
 tcgtcgtcgg cgatggcggc agcggcagcc ccctatgcgg cgtggcttgc cgcggcgggc 240
 gtccaggccg agcagacggc cgctcaggct gcggcgatga tagccgagtt tgaagcggtc 300
 aagacggcgg tgggtgcagcc gatgctgggt gcggccaacc gtgccgacct ggtgtcgtg 360
 gtgatgtcga acctgtttgg acagaacgct ccggcgatcg ctgccattga agccacgtac 420

10

ES 2 647 321 T3

gagcaaatgt gggctgccga tgtgtcggcg atgtctgcct accatgccgg ggcacggcg	480
atcgcctcgg cgctgtcccc gttcagtaaa ccgctgcaga acctggctgg cttgccggct	540
tggttgcca gcggcgcc tgcggccgc atgaccgcag ccgcaggcat accggcgctt	600
gcgggcggac ccaccgccat caacctgggc atagccaacg tcggcggctgg caacgtcggc	660
aacgccaaca acggccttgc caacatcggc aacgccaacc ttggcaacta caatttcggg	720
tccggaaatt tcggtaactc caatatcggc tcagcaagcc tgggtaataa caacatcggc	780
ttcgggaacc tcggcagcaa caatgtcggc gtgggaaacc ttggcaatct caacaccggg	840
tttgccaaca ccggccttggg caacttcggc tttggcaaca ctggcaacaa caacatcggc	900
atcggcttta ccggcaacaa ccagatcggc atcggcgggc tcaactcggg caccgggaat	960
ttcggattgt tcaactcggg cagcggaaac gtcggcttct tcaactccgg caatggaaac	1020
tttggcatcg gaaactcggg taatttcaac accggtggct ggaattctgg acacgggaac	1080
acgggcttct tcaatcggg ctcgtttaac accggtatgt tggacgtcgg caacgcgaac	1140
acaggcagcc tgaacaccgg cagttataac atgggcgact tcaatccggg gtcgtccaac	1200
accggcacgt tcaacaccgg aaatgctaac acgggtttcc tcaacccggg aaatatcaac	1260
actggtgtct tcaatattgg ccacatgaat aatgggctgt tcaacaccgg tgacatgaac	1320
aatggcgtct tctaccgggg cgtggggcag ggcagcctgc agttcagtat tacgacacct	1380
gatctgactc tgccgccgct gcaaataaccg gggatatcgg tccccgctt cagtctgccg	1440
gcaataacgc tgccgtcgt gaacatcccg gccgccacca caccggccaa catcacctc	1500
ggcgccttca gcctgcccgg gttgacgttg ccgctcgttga acatcccggc gccaccaca	1560
ccagccaaca tcaccgtggg tgccctcagc ctgcccgggt tgacgttgcc gtcgttgaac	1620
atcccggccg ccaccacacc agccaacatc accgtcggcg ccttcagcct gcccggttg	1680
acgttgccgt cgttgaacat cccggccgc accacaccag ccaacatcac cgtcggcgc	1740
ttcagcctgc ccgggttgac gttgccgtcg ttgaacatcc cggccgccac cacaccagcc	1800
aacatcacg tcggcgcctt cagcctgcc gggttgacgt tgccgtcgtt gaacatcccg	1860
gccgccacca caccgcca catcacctga agcggccttc agttgcctcc gctgagtatt	1920
ccttcgtag caattccgcc ggtgacggtc ccgccatta cgtggtgtc ttttaatttg	1980
ccgccattgc agattccgga agtaactatt ccgcagctga cgatacccgc gggtatcaca	2040
atcggtggt ttagtctacc tgcgatacat actcaaccga taacggtcgg ccagattggc	2100
gtgggccaat ttggcctgcc ctccatagc tgggatgtt tctaagcac acctaggata	2160
acagtaccgg cttttggaat accctttacc ctacaattcc agaccaatgt gcctgcgctt	2220
cagccgccg gcggcgggct tagtactttc accaatggcg ccctcatctt cggtgagttt	2280
gacttaccac aattggtggt tcaccatac acattgaccg gccctattgt catcgttca	2340

ES 2 647 321 T3

```

ttctttctgc cgccttcaa catacccggg atcgaˆtgtcc ccgˆctatcaa cgtcgatggc      2400
ttcacocctgc cgcagatcac caccˆccagct atcaccaccc cggagttcgc gatccctccg      2460
atcggcgtgg gcggttcac tctgccgcag atcaccaccc agˆgaaatcat caccˆccggag      2520
ctaaccatca actcgatcgg cgtcggcggg ttcaccctgc cgˆaaatcac caccˆccaccc      2580
atcaccaccc caccgctgac catcgacccc atcaacctca ccgˆgcttcac cctcccccaa      2640
atcaccaccc caccatcac caccˆccaccg ctgaccatcg accccatcaa cctcaccggc      2700
ttcacocctcc cccaaatcac caccˆccaccc atcaccaccc caccgctcac catcgagccg      2760
atcggcgtgg ggggcttcac cacgcccccg ctcaccgttc ccgˆgcatcca cctgcccagc      2820
accacgatcg gggccttcgc gatccccggg gggccgggct acttcaactc gagcaccgcg      2880
ccttcgctgg gcttcttcaa ttccgggtgcg ggcggcaact cggˆgcttcg caacaacggc      2940
tcgggcctct cgggttggtt caacaccaac ccgˆgccgggc tgttgggˆgcg ctcgggctat      3000
cagaacttcg gcgggctatc ctcgggcttt tccaaccttg gcagcggcgt ctcaggcttc      3060
gccaacaggg gcacˆctgcc gttctcggtg gccagcgtcg tttccggctt tgccaatˆatc      3120
ggcaccaacc tggcgggttt cttccaaggc accacgtcct aa                          3162

```

```

<210> 3
<211> 1105
<212> PRT
<213> Mycobacterium tuberculosis

```

5

```

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

```

```

<400> 3

```

```

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

```

10

ES 2 647 321 T3

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

ES 2 647 321 T3

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 Thr 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

ES 2 647 321 T3

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 Gly Ala Phe Ser Leu Pro Gly Leu Thr Leu
 625 630 635 640

Pro Ser Leu Asn Ile Pro Ala Ala Thr Thr Pro Ala Asn Ile Thr Val
 645 650 655

Gly Ala Phe Ser Leu Pro Gly Leu Thr Leu Pro Ser Leu Asn Ile Pro
 660 665 670

Ala Ala Thr Thr Pro Ala Asn Ile Thr Val Ser Gly Phe Gln Leu Pro
 675 680 685

Pro Leu Ser Ile Pro Ser Val Ala Ile Pro Pro Val Thr Val Pro Pro
 690 695 700

Ile Thr Val Gly Ala Phe Asn Leu Pro Pro Leu Gln Ile Pro Glu Val
 705 710 715 720

Thr Ile Pro Gln Leu Thr Ile Pro Ala Gly Ile Thr Ile Gly Gly Phe
 725 730 735

Ser Leu Pro Ala Ile His Thr Gln Pro Ile Thr Val Gly Gln Ile Gly
 740 745 750

Val Gly Gln Phe Gly Leu Pro Ser Ile Gly Trp Asp Val Phe Leu Ser
 755 760 765

Thr Pro Arg Ile Thr Val Pro Ala Phe Gly Ile Pro Phe Thr Leu Gln
 770 775 780

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

Thr Phe Thr Asn Gly Ala Leu Ile Phe Gly Glu Phe Asp Leu Pro Gln
 805 810 815

Leu Val Val His Pro Tyr Thr Leu Thr Gly Pro Ile Val Ile Gly Ser
 820 825 830

Phe Phe Leu Pro Ala Phe Asn Ile Pro Gly Ile Asp Val Pro Ala Ile
 835 840 845

Asn Val Asp Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Ala Ile Thr
 850 855 860

ES 2 647 321 T3

Thr Pro Glu Phe Ala Ile Pro Pro Ile Gly Val Gly Gly Phe Thr Leu
 865 870 875 880

Pro Gln Ile Thr Thr Gln Glu Ile Ile Thr Pro Glu Leu Thr Ile Asn
 885 890 895

Ser Ile Gly Val Gly Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro
 900 905 910

Ile Thr Thr Pro Pro Leu Thr Ile Asp Pro Ile Asn Leu Thr Gly Phe
 915 920 925

Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr
 930 935 940

Ile Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr Thr
 945 950 955 960

Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile Glu Pro Ile Gly Val Gly
 965 970 975

Gly Phe Thr Thr Pro Pro Leu Thr Val Pro Gly Ile His Leu Pro Ser
 980 985 990

Thr Thr Ile Gly Ala Phe Ala Ile Pro Gly Gly Pro Gly Tyr Phe Asn
 995 1000 1005

Ser Ser Thr Ala Pro Ser Ser Gly Phe Phe Asn Ser Gly Ala Gly
 1010 1015 1020

Gly Asn Ser Gly Phe Gly Asn Asn Gly Ser Gly Leu Ser Gly Trp
 1025 1030 1035

Phe Asn Thr Asn Pro Ala Gly Leu Leu Gly Gly Ser Gly Tyr Gln
 1040 1045 1050

Asn Phe Gly Gly Leu Ser Ser Gly Phe Ser Asn Leu Gly Ser Gly
 1055 1060 1065

Val Ser Gly Phe Ala Asn Arg Gly Ile Leu Pro Phe Ser Val Ala
 1070 1075 1080

Ser Val Val Ser Gly Phe Ala Asn Ile Gly Thr Asn Leu Ala Gly
 1085 1090 1095

Phe Phe Gln Gly Thr Thr Ser
 1100 1105

ES 2 647 321 T3

<210> 4
 <211> 975
 <212> PRT
 <213> Mycobacterium tuberculosis

5

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

<400> 4

```

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
  
```

ES 2 647 321 T3

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

ES 2 647 321 T3

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 Thr 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 Ser Gly Phe Gln Leu Pro Pro Leu
545 550 555 560

Ser Ile Pro Ser Val Ala Ile Pro Pro Val Thr Val Pro Pro Ile Thr
565 570 575

Val Gly Ala Phe Asn Leu Pro Pro Leu Gln Ile Pro Glu Val Thr Ile
580 585 590

Pro Gln Leu Thr Ile Pro Ala Gly Ile Thr Ile Gly Gly Phe Ser Leu
595 600 605

Pro Ala Ile His Thr Gln Pro Ile Thr Val Gly Gln Ile Gly Val Gly
610 615 620

Gln Phe Gly Leu Pro Ser Ile Gly Trp Asp Val Phe Leu Ser Thr Pro
625 630 635 640

Arg Ile Thr Val Pro Ala Phe Gly Ile Pro Phe Thr Leu Gln Phe Gln
645 650 655

Thr Asn Val Pro Ala Leu Gln Pro Pro Gly Gly Gly Leu Ser Thr Phe
660 665 670

Thr Asn Gly Ala Leu Ile Phe Gly Glu Phe Asp Leu Pro Gln Leu Val
675 680 685

Val His Pro Tyr Thr Leu Thr Gly Pro Ile Val Ile Gly Ser Phe Phe
690 695 700

Leu Pro Ala Phe Asn Ile Pro Gly Ile Asp Val Pro Ala Ile Asn Val
705 710 715 720

ES 2 647 321 T3

Asp Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Ala Ile Thr Thr Pro
725 730 735

Glu Phe Ala Ile Pro Pro Ile Gly Val Gly Gly Phe Thr Leu Pro Gln
740 745 750

Ile Thr Thr Gln Glu Ile Ile Thr Pro Glu Leu Thr Ile Asn Ser Ile
755 760 765

Gly Val Gly Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr
770 775 780

Thr Pro Pro Leu Thr Ile Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu
785 790 795 800

Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile Asp
805 810 815

Pro Ile Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro
820 825 830

Ile Thr Thr Pro Pro Leu Thr Ile Glu Pro Ile Gly Val Gly Gly Phe
835 840 845

Thr Thr Pro Pro Leu Thr Val Pro Gly Ile His Leu Pro Ser Thr Thr
850 855 860

Ile Gly Ala Phe Ala Ile Pro Gly Gly Pro Gly Tyr Phe Asn Ser Ser
865 870 875 880

Thr Ala Pro Ser Ser Gly Phe Phe Asn Ser Gly Ala Gly Gly Asn Ser
885 890 895

Gly Phe Gly Asn Asn Gly Ser Gly Leu Ser Gly Trp Phe Asn Thr Asn
900 905 910

Pro Ala Gly Leu Leu Gly Gly Ser Gly Tyr Gln Asn Phe Gly Gly Leu
915 920 925

Ser Ser Gly Phe Ser Asn Leu Gly Ser Gly Val Ser Gly Phe Ala Asn
930 935 940

Arg Gly Ile Leu Pro Phe Ser Val Ala Ser Val Val Ser Gly Phe Ala
945 950 955 960

Asn Ile Gly Thr Asn Leu Ala Gly Phe Phe Gln Gly Thr Thr Ser
965 970 975

ES 2 647 321 T3

<210> 5
 <211> 1050
 <212> PRT
 <213> Mycobacterium tuberculosis

5

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

<400> 5

```

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
    
```

ES 2 647 321 T3

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

ES 2 647 321 T3

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 Thr 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 Ser Gly Phe Gln Leu Pro Pro Leu
545 550 555 560

Ser Ile Pro Ser Val Ala Ile Pro Pro Val Thr Val Pro Pro Ile Thr
565 570 575

Val Gly Ala Phe Asn Leu Pro Pro Leu Gln Ile Pro Glu Val Thr Ile
580 585 590

Pro Gln Leu Thr Ile Pro Ala Gly Ile Thr Ile Gly Gly Phe Ser Leu
595 600 605

Pro Ala Ile His Thr Gln Pro Ile Thr Val Gly Gln Ile Gly Val Gly
610 615 620

Gln Phe Gly Leu Pro Ser Ile Gly Trp Asp Val Phe Leu Ser Thr Pro
625 630 635 640

Arg Ile Thr Val Pro Ala Phe Gly Ile Pro Phe Thr Leu Gln Phe Gln
645 650 655

Thr Asn Val Pro Ala Leu Gln Pro Pro Gly Gly Gly Leu Ser Thr Phe
660 665 670

Thr Asn Gly Ala Leu Ile Phe Gly Glu Phe Asp Leu Pro Gln Leu Val
675 680 685

Val His Pro Tyr Thr Leu Thr Gly Pro Ile Val Ile Gly Ser Phe Phe
690 695 700

Leu Pro Ala Phe Asn Ile Pro Gly Ile Asp Val Pro Ala Ile Asn Val
705 710 715 720

ES 2 647 321 T3

Asp Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Ala Ile Thr Thr Pro
725 730 735

Glu Phe Ala Ile Pro Pro Ile Gly Val Gly Gly Phe Thr Leu Pro Gln
740 745 750

Ile Thr Thr Gln Glu Ile Ile Thr Pro Glu Leu Thr Ile Asn Ser Ile
755 760 765

Gly Val Gly Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr
770 775 780

Thr Pro Pro Leu Thr Ile Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu
785 790 795 800

Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile Asp
805 810 815

Pro Ile Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro
820 825 830

Ile Thr Thr Pro Pro Leu Thr Ile Asp Pro Ile Asn Leu Thr Gly Phe
835 840 845

Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr
850 855 860

Ile Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr Thr
865 870 875 880

Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile Asp Pro Ile Asn Leu Thr
885 890 895

Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro
900 905 910

Leu Thr Ile Glu Pro Ile Gly Val Gly Gly Phe Thr Thr Pro Pro Leu
915 920 925

Thr Val Pro Gly Ile His Leu Pro Ser Thr Thr Ile Gly Ala Phe Ala
930 935 940

Ile Pro Gly Gly Pro Gly Tyr Phe Asn Ser Ser Thr Ala Pro Ser Ser
945 950 955 960

Gly Phe Phe Asn Ser Gly Ala Gly Gly Asn Ser Gly Phe Gly Asn Asn
965 970 975

ES 2 647 321 T3

Gly Ser Gly Leu Ser Gly Trp Phe Asn Thr Asn Pro Ala Gly Leu Leu
 980 985 990

Gly Gly Ser Gly Tyr Gln Asn Phe Gly Gly Leu Ser Ser Gly Phe Ser
 995 1000 1005

Asn Leu Gly Ser Gly Val Ser Gly Phe Ala Asn Arg Gly Ile Leu
 1010 1015 1020

Pro Phe Ser Val Ala Ser Val Val Ser Gly Phe Ala Asn Ile Gly
 1025 1030 1035

Thr Asn Leu Ala Gly Phe Phe Gln Gly Thr Thr Ser
 1040 1045 1050

<210> 6
 <211> 1078
 <212> PRT
 <213> Mycobacterium tuberculosis

5

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

<400> 6

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

10

ES 2 647 321 T3

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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 Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu
 915 920 925

Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile Glu
 930 935 940

Pro Ile Gly Val Gly Gly Phe Thr Thr Pro Pro Leu Thr Val Pro Gly
 945 950 955 960

Ile His Leu Pro Ser Thr Thr Ile Gly Ala Phe Ala Ile Pro Gly Gly
 965 970 975

Pro Gly Tyr Phe Asn Ser Ser Thr Ala Pro Ser Ser Gly Phe Phe Asn
 980 985 990

Ser Gly Ala Gly Gly Asn Ser Gly Phe Gly Asn Asn Gly Ser Gly Leu
 995 1000 1005

Ser Gly Trp Phe Asn Thr Asn Pro Ala Gly Leu Leu Gly Gly Ser
 1010 1015 1020

Gly Tyr Gln Asn Phe Gly Gly Leu Ser Ser Gly Phe Ser Asn Leu
 1025 1030 1035

Gly Ser Gly Val Ser Gly Phe Ala Asn Arg Gly Ile Leu Pro Phe
 1040 1045 1050

Ser Val Ala Ser Val Val Ser Gly Phe Ala Asn Ile Gly Thr Asn
 1055 1060 1065

Leu Ala Gly Phe Phe Gln Gly Thr Thr Ser
 1070 1075

<210> 7

<211> 1026

<212> PRT

5 <213> Mycobacterium bovis

<220>

<221> MISC_FEATURE

<223> Cepa BCG

<400> 7

10 Met Asn Phe Ser Val Leu Pro Pro Glu Ile Asn Ser Ala Leu Ile Phe
 1 5 10 15

ES 2 647 321 T3

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

ES 2 647 321 T3

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 Thr 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

ES 2 647 321 T3

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 Ser Gly Phe Gln Leu Pro Pro Leu Ser Ile Pro Ser Val Ala
 580 585 590

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

Pro Pro Leu Gln Ile Pro Glu Val Thr Ile Pro Gln Leu Thr Ile Pro
 610 615 620

Ala Gly Ile Thr Ile Gly Gly Phe Ser Leu Pro Ala Ile His Thr Gln
 625 630 635 640

Pro Ile Thr Val Gly Gln Ile Gly Val Gly Gln Phe Gly Leu Pro Ser
 645 650 655

Ile Gly Trp Asp Val Phe Leu Ser Thr Pro Arg Ile Thr Val Pro Ala
 660 665 670

Phe Gly Ile Pro Phe Thr Leu Gln Phe Gln Thr Asn Val Pro Ala Leu
 675 680 685

Gln Pro Pro Gly Gly Gly Leu Ser Thr Phe Thr Asn Gly Ala Leu Ile
 690 695 700

Phe Gly Glu Phe Asp Leu Pro Gln Leu Val Val His Pro Tyr Thr Leu
 705 710 715 720

Thr Gly Pro Ile Val Ile Gly Ser Phe Phe Leu Pro Ala Phe Asn Ile
 725 730 735

Pro Gly Ile Asp Val Pro Ala Ile Asn Val Asp Gly Phe Thr Leu Pro
 740 745 750

Gln Ile Thr Thr Pro Ala Ile Thr Thr Pro Glu Phe Ala Ile Pro Pro
 755 760 765

Ile Gly Val Gly Gly Phe Thr Leu Pro Gln Ile Thr Thr Gln Glu Ile
 770 775 780

ES 2 647 321 T3

Ile Thr Pro Glu Leu Thr Ile Asn Ser Ile Gly Val Gly Gly Phe Thr
785 790 795 800

Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile
805 810 815

Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro
820 825 830

Pro Ile Thr Thr Pro Pro Leu Thr Ile Asp Pro Ile Asn Leu Thr Gly
835 840 845

Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu
850 855 860

Thr Ile Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr
865 870 875 880

Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile Glu Pro Ile Gly Val
885 890 895

Gly Gly Phe Thr Thr Pro Pro Leu Thr Val Pro Gly Ile His Leu Pro
900 905 910

Ser Thr Thr Ile Gly Ala Phe Ala Ile Pro Gly Gly Pro Gly Tyr Phe
915 920 925

Asn Ser Ser Thr Ala Pro Ser Ser Gly Phe Phe Asn Ser Gly Ala Gly
930 935 940

Gly Asn Ser Gly Phe Gly Asn Asn Gly Ser Gly Leu Ser Gly Trp Phe
945 950 955 960

Asn Thr Asn Pro Ala Gly Leu Leu Gly Gly Ser Gly Tyr Gln Asn Phe
965 970 975

Gly Gly Leu Ser Ser Gly Phe Ser Asn Leu Gly Ser Gly Val Ser Gly
980 985 990

Phe Ala Asn Arg Gly Ile Leu Pro Phe Ser Val Ala Ser Val Val Ser
995 1000 1005

Gly Phe Ala Asn Ile Gly Thr Asn Leu Ala Gly Phe Phe Gln Gly
1010 1015 1020

Thr Thr Ser
1025

<210> 8
<211> 110

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis

5 <220>
 <221> péptido mat
 <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
    
```

10 <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
    
```

ES 2 647 321 T3

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

5 <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

10 <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

ES 2 647 321 T3

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

5 <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

ES 2 647 321 T3

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

5 <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

ES 2 647 321 T3

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> 392

<212> PRT

<213> Mycobacterium tuberculosis

5

<400> 14

ES 2 647 321 T3

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

ES 2 647 321 T3

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> 15
 <211> 423
 <212> PRT
 <213> Mycobacterium tuberculosis

5

<400> 15

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

ES 2 647 321 T3

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 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

ES 2 647 321 T3

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> 16
 <211> 95
 <212> PRT
 <213> Mycobacterium tuberculosis

5

<220>
 <221> INIT_MET
 <222>(1).(1)

10

<220>
 <221> péptido mat
 <222>(2).(95)

<400> 16

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

15

<210> 17
 <211> 338
 <212> PRT
 <213> Mycobacterium tuberculosis

<220>

ES 2 647 321 T3

<221> péptido mat
 <222>(43).(338)

<400> 17

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	

ES 2 647 321 T3

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> 18
 <211> 325
 <212> PRT
 <213> Mycobacterium tuberculosis

5

<220>
 <221> péptido mat
 <222>(41).(325)

<400> 18

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

10

ES 2 647 321 T3

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

ES 2 647 321 T3

Ser Leu Gly Ala Gly
285

5 <210> 19
<211> 144
<212> PRT
<213> Mycobacterium tuberculosis

<220>
<221> INIT_MET
<222>(1).(1)

10 <220>
<221> péptido mat
<222>(2).(144)

<400> 19

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

15 <210> 20
<211> 228
<212> PRT
<213> Mycobacterium tuberculosis

<220>
<221> péptido mat
<222>(24).(228)

20 <400> 20

ES 2 647 321 T3

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> 21
 <211> 355
 <212> PRT
 <213> Mycobacterium tuberculosis
 <220>

ES 2 647 321 T3

<221> péptido mat
 <222>(33).(355)

<400> 21

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		

ES 2 647 321 T3

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

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> 323
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> Mutante Ser/Ala de Mtb32A madura

<400> 22

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

ES 2 647 321 T3

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

ES 2 647 321 T3

Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
 305 310 315 320

Pro Pro Ala

<210> 23

<211> 96

<212> PRT

5 <213> Mycobacterium tuberculosis

<400> 23

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> 24

<211> 723

<212> PRT

10 <213> Secuencia artificial

<220>

<223> Mtb72f

<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

15

ES 2 647 321 T3

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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> 25
- <211> 723
- <212> PRT
- <213> Secuencia artificial

5

- <220>
- <223> M72
- <400> 25

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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> 26

<211> 702

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Mtb71f

<400> 26

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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
 580 585 590 595

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> 27
 <211> 920
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> M72-Mtb9.9-Mtb9.8
 <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

ES 2 647 321 T3

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

ES 2 647 321 T3

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> 28

<211> 1010

<212> PRT

5 <213> Secuencia artificial

<220>

<223> M103

<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

ES 2 647 321 T3

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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> 29
 <211> 1148
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> M114
 <400> 29

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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

ES 2 647 321 T3

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

5 <210> 30
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 30

Trp Gln Gly Ala Ser Ser Ser Ala Met
 1 5

10 <210> 31
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 31

Val Gln Ala Glu Gln Thr Ala Ala Gln
 1 5

15 <210> 32
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 32

Val Lys Thr Ala Val Val Gln Pro Met
 1 5

20 <210> 33
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 33

Val Gln Pro Met Leu Val Ala Ala Asn
 1 5

25 <210> 34
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 30 <400> 34

Leu Val Ala Ala Asn Arg Ala Asp Leu
 1 5

35 <210> 35
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis

ES 2 647 321 T3

<400> 35
 Leu Val Ser Leu Val Met Ser Asn Leu
 1 5

<210> 36
 <211> 9
 5 <212> PRT
 <213> Mycobacterium tuberculosis

<400> 36
 Leu Val Met Ser Asn Leu Phe Gly Gln
 1 5

<210> 37
 <211> 9
 10 <212> PRT
 <213> Mycobacterium tuberculosis

<400> 37
 Tyr Glu Gln Met Trp Ala Ala Asp Val
 1 5

<210> 38
 <211> 9
 15 <212> PRT
 <213> Mycobacterium tuberculosis

<400> 38
 Trp Ala Ala Asp Val Ser Ala Met Ser
 1 5

<210> 39
 <211> 9
 20 <212> PRT
 <213> Mycobacterium tuberculosis

<400> 39
 Leu Gln Asn Leu Ala Gly Leu Pro Ala
 1 5

<210> 40
 <211> 9
 25 <212> PRT
 <213> Mycobacterium tuberculosis

<400> 40
 Phe Gly Asn Leu Gly Ser Asn Asn Val
 1 5

<210> 41
 <211> 9
 30 <212> PRT
 <213> Mycobacterium tuberculosis

<400> 41
 Phe Gly Asn Thr Gly Asn Asn Asn Ile
 1 5

<210> 42
 <211> 9
 35 <212> PRT
 <213> Mycobacterium tuberculosis

<400> 42
 Phe Gly Asn Thr Gly Asn Asn Asn Ile
 1 5

<210> 43
 <211> 9
 40 <212> PRT
 <213> Mycobacterium tuberculosis

ES 2 647 321 T3

<400> 42
Phe Leu Asn Ala Gly Asn Ile Asn Thr
1 5

<210> 43
<211> 9
5 <212> PRT
<213> Mycobacterium tuberculosis

<400> 43
Leu Gln Phe Ser Ile Thr Thr Pro Asp
1 5

10 <210> 44
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

<400> 44
Leu Thr Ile Pro Ala Gly Ile Thr Ile
1 5

15 <210> 45
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

<400> 45
Phe Gly Ile Pro Phe Thr Leu Gln Phe
1 5

20 <210> 46
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

25 <400> 46
Leu Gln Phe Gln Thr Asn Val Pro Ala
1 5

30 <210> 47
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

<400> 47
Phe Gln Thr Asn Val Pro Ala Leu Gln
1 5

35 <210> 48
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

<400> 48
Tyr Thr Leu Thr Gly Pro Ile Val Ile
1 5

40 <210> 49
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

ES 2 647 321 T3

<400> 49
Phe Leu Pro Ala Phe Asn Ile Pro Gly
1 5

<210> 50
<211> 9
5 <212> PRT
<213> Mycobacterium tuberculosis

<400> 50
Leu Thr Ile Asp Pro Ile Asn Leu Thr
1 5

10 <210> 51
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

<400> 51
Leu Thr Ile Asp Pro Ile Asn Leu Thr
1 5

15 <210> 52
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

<400> 52
Tyr Phe Asn Ser Ser Thr Ala Pro Ser
20 1 5

<210> 53
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

25 <400> 53
Phe Asn Ser Ser Thr Ala Pro Ser Ser
1 5

<210> 54
<211> 9
30 <212> PRT
<213> Mycobacterium tuberculosis

<400> 54
Phe Gly Asn Asn Gly Ser Gly Leu Ser
1 5

<210> 55
<211> 9
35 <212> PRT
<213> Mycobacterium tuberculosis

<400> 55
Tyr Gln Asn Phe Gly Gly Leu Ser Ser
1 5

40 <210> 56
<211> 9
<212> PRT
<213> Mycobacterium tuberculosis

ES 2 647 321 T3

<400> 56

Phe Gly Gly Leu Ser Ser Gly Phe Ser
1 5

<210> 57

<211> 9

5 <212> PRT

<213> Mycobacterium tuberculosis

<400> 57

Phe Ala Asn Arg Gly Ile Leu Pro Phe

1

5

<210> 58

10 <211> 9

<212> PRT

<213> Mycobacterium tuberculosis

<400> 58

Ile Leu Pro Phe Ser Val Ala Ser Val
1 5

<210> 59

15 <211> 9

<212> PRT

<213> Mycobacterium tuberculosis

<400> 59

Phe Ala Asn Ile Gly Thr Asn Leu Ala
1 5

<210> 60

20 <211> 9

<212> PRT

<213> Mycobacterium tuberculosis

<400> 60

Asn Phe Ser Val Leu Pro Pro Glu Ile
1 5

<210> 61

30 <211> 9

<212> PRT

<213> Mycobacterium tuberculosis

<400> 61

Val Leu Pro Pro Glu Ile Asn Ser Ala
1 5

<210> 62

35 <211> 9

<212> PRT

<213> Mycobacterium tuberculosis

<400> 62

Leu Pro Pro Glu Ile Asn Ser Ala Leu
1 5

<210> 63

40 <211> 9

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis
 <400> 63
 Pro Glu Ile Asn Ser Ala Leu Ile Phe
 1 5
 5 <210> 64
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 64
 Glu Ile Asn Ser Ala Leu Ile Phe Ala
 1 5
 10 <210> 65
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 65
 Gly Ala Gly Pro Glu Pro Met Ala Ala
 1 5
 15 <210> 66
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 66
 Gly Pro Glu Pro Met Ala Ala Ala Ala
 1 5
 20 <210> 67
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 67
 Glu Pro Met Ala Ala Ala Ala Thr Ala
 1 5
 25 <210> 68
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 68
 Ala Ala Ala Thr Ala Trp Asp Gly Leu
 1 5
 30 <210> 69
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 69
 Ala Thr Ala Trp Asp Gly Leu Ala Met
 1 5
 35 <210> 70
 <211> 9

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis
 <400> 70
 Ala Trp Asp Gly Leu Ala Met Glu Leu
 1 5
 5 <210> 71
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 71
 Gly Leu Ala Met Glu Leu Ala Ser Ala
 1 5
 10 <210> 72
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 72
 Leu Ala Met Glu Leu Ala Ser Ala Ala
 1 5
 15 <210> 73
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 73
 Val Thr Ser Gly Leu Val Gly Gly Ala
 1 5
 20 <210> 74
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 74
 Ala Met Ala Ala Ala Ala Ala Pro Tyr
 1 5
 25 <210> 75
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 75
 Ala Ala Ala Ala Ala Pro Tyr Ala Ala
 1 5
 30 <210> 76
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 76
 Ala Ala Ala Pro Tyr Ala Ala Trp Leu
 1 5
 35 <210> 77
 <211> 9

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis
 <400> 77
 Ala Ala Pro Tyr Ala Ala Trp Leu Ala
 1 5
 5 <210> 78
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 78
 Ala Pro Tyr Ala Ala Trp Leu Ala Ala
 1 5
 10 <210> 79
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 79
 Tyr Ala Ala Trp Leu Ala Ala Ala Ala
 1 5
 15 <210> 80
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 80
 Ala Ala Trp Leu Ala Ala Ala Ala Val
 1 5
 20 <210> 81
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 81
 Trp Leu Ala Ala Ala Ala Val Gln Ala
 1 5
 25 <210> 82
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 82
 Gln Ala Glu Gln Thr Ala Ala Gln Ala
 1 5
 30 <210> 83
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 83
 Ala Glu Gln Thr Ala Ala Gln Ala Ala
 1 5
 35 <210> 84
 <211> 9

ES 2 647 321 T3

```

<212> PRT
<213> Mycobacterium tuberculosis

<400> 84

          Thr Ala Ala Gln Ala Ala Ala Met Ile
          1                               5

5  <210> 85
    <211> 9
    <212> PRT
    <213> Mycobacterium tuberculosis

    <400> 85

          Ala Ala Met Ile Ala Glu Phe Glu Ala
          1                               5

10 <210> 86
    <211> 9
    <212> PRT
    <213> Mycobacterium tuberculosis

    <400> 86

          Ala Met Ile Ala Glu Phe Glu Ala Val
          1                               5

15 <210> 87
    <211> 9
    <212> PRT
    <213> Mycobacterium tuberculosis

    <400> 87

          Ala Glu Phe Glu Ala Val Lys Thr Ala
          1                               5

20 <210> 88
    <211> 9
    <212> PRT
    <213> Mycobacterium tuberculosis

    <400> 88

          Phe Glu Ala Val Lys Thr Ala Val Val
          1                               5

25 <210> 89
    <211> 9
    <212> PRT
    <213> Mycobacterium tuberculosis

    <400> 89

          Glu Ala Val Lys Thr Ala Val Val Gln
          1                               5

30 <210> 90
    <211> 9
    <212> PRT
    <213> Mycobacterium tuberculosis

    <400> 90

          Lys Thr Ala Val Val Gln Pro Met Leu
          1                               5

35 <210> 91
    <211> 9
  
```

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis
 <400> 91
 Gln Pro Met Leu Val Ala Ala Asn Arg
 1 5
 5 <210> 92
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 92
 Pro Met Leu Val Ala Ala Asn Arg Ala
 1 5
 10 <210> 93
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 93
 Leu Val Ala Ala Asn Arg Ala Asp Leu
 1 5
 <210> 94
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 94
 Ala Asn Arg Ala Asp Leu Val Ser Leu
 1 5
 <210> 95
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 95
 Arg Ala Asp Leu Val Ser Leu Val Met
 1 5
 <210> 96
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 96
 Val Ser Leu Val Met Ser Asn Leu Phe
 1 5
 <210> 97
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 97
 Asn Leu Phe Gly Gln Asn Ala Pro Ala
 1 5
 <210> 98
 <211> 9

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis
 <400> 98
 Ala Pro Ala Ile Ala Ala Ile Glu Ala
 1 5
 5 <210> 99
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 99
 Ala Ile Ala Ala Ile Glu Ala Thr Tyr
 1 5
 10 <210> 100
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 100
 Ala Thr Tyr Glu Gln Met Trp Ala Ala
 1 5
 15 <210> 101
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 101
 Gln Met Trp Ala Ala Asp Val Ser Ala
 1 5
 20 <210> 102
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 102
 Ala Met Ser Ala Tyr His Ala Gly Ala
 1 5
 25 <210> 103
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 103
 Ser Ala Tyr His Ala Gly Ala Ser Ala
 1 5
 30 <210> 104
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 104
 Ala Tyr His Ala Gly Ala Ser Ala Ile
 1 5
 35 <210> 105
 <211> 9

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis
 <400> 105
 Gly Ala Ser Ala Ile Ala Ser Ala Leu
 1 5
 5 <210> 106
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 106
 Ala Ile Ala Ser Ala Leu Ser Pro Phe
 10 1 5
 <210> 107
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 15 <400> 107
 Ala Leu Ser Pro Phe Ser Lys Pro Leu
 1 5
 <210> 108
 <211> 9
 <212> PRT
 20 <213> Mycobacterium tuberculosis
 <400> 108
 Pro Phe Ser Lys Pro Leu Gln Asn Leu
 1 5
 <210> 109
 <211> 9
 25 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 109
 Lys Pro Leu Gln Asn Leu Ala Gly Leu
 1 5
 <210> 110
 <211> 9
 30 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 110
 Asn Leu Ala Gly Leu Pro Ala Trp Leu
 1 5
 <210> 111
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 111
 Leu Ala Gly Leu Pro Ala Trp Leu Ala
 40 1 5
 <210> 112
 <211> 9

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis
 <400> 112
 Leu Pro Ala Trp Leu Ala Ser Gly Ala
 1 5
 5 <210> 113
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 113
 Trp Leu Ala Ser Gly Ala Pro Ala Ala
 1 5
 10 <210> 114
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 114
 Gly Ala Pro Ala Ala Ala Met Thr Ala
 1 5
 15 <210> 115
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 115
 Ala Pro Ala Ala Ala Met Thr Ala Ala
 1 5
 20 <210> 116
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 116
 Ala Ala Met Thr Ala Ala Ala Gly Ile
 1 5
 25 <210> 117
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 117
 Thr Ala Ala Ala Gly Ile Pro Ala Leu
 1 5
 30 <210> 118
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 118
 Ala Ala Ala Gly Ile Pro Ala Leu Ala
 1 5
 35 <210> 119
 <211> 9

ES 2 647 321 T3

<212> PRT
 <213> Mycobacterium tuberculosis
 <400> 119
 Ala Leu Ala Gly Gly Pro Thr Ala Ile
 1 5
 5 <210> 120
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 120
 Ala Gly Gly Pro Thr Ala Ile Asn Leu
 1 5
 10 <210> 121
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 121
 Gly Pro Thr Ala Ile Asn Leu Gly Ile
 1 5
 1 5
 <210> 122
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 122
 Ala Ile Asn Leu Gly Ile Ala Asn Val
 1 5
 1 5
 20 <210> 123
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 123
 Asn Ala Asn Leu Gly Asn Tyr Asn Phe
 1 5
 1 5
 25 <210> 124
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 124
 Asn Tyr Asn Phe Gly Ser Gly Asn Phe
 1 5
 1 5
 30 <210> 125
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 125
 Asn Leu Gly Ser Asn Asn Val Gly Val
 1 5
 1 5
 35 <210> 126
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 126
 Asn Leu Gly Ser Asn Asn Val Gly Val
 1 5
 1 5
 40 <210> 127
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 127
 Asn Leu Gly Ser Asn Asn Val Gly Val
 1 5
 1 5

ES 2 647 321 T3

<210> 126
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 5 <400> 126
 Ser Leu Asn Thr Gly Ser Tyr Asn Met
 1 5

 <210> 127
 <211> 9
 <212> PRT
 10 <213> Mycobacterium tuberculosis
 <400> 127
 Asn Ala Asn Thr Gly Phe Leu Asn Ala
 1 5

 <210> 128
 <211> 9
 <212> PRT
 15 <213> Mycobacterium tuberculosis
 <400> 128
 Phe Leu Asn Ala Gly Asn Ile Asn Thr
 1 5

 <210> 129
 <211> 9
 <212> PRT
 20 <213> Mycobacterium tuberculosis
 <400> 129
 Asn Ile Asn Thr Gly Val Phe Asn Ile
 1 5

 <210> 130
 <211> 9
 <212> PRT
 25 <213> Mycobacterium tuberculosis
 <400> 130
 Gly Val Gly Gln Gly Ser Leu Gln Phe
 1 5

 <210> 131
 <211> 9
 <212> PRT
 30 <213> Mycobacterium tuberculosis
 <400> 131
 Ser Ile Thr Thr Pro Asp Leu Thr Leu
 1 5

 <210> 132
 <211> 9
 <212> PRT
 35 <213> Mycobacterium tuberculosis
 <400> 132
 Thr Pro Asp Leu Thr Leu Pro Pro Leu
 1 5

ES 2 647 321 T3

<210> 133
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 5 <400> 133
 Asp Leu Thr Leu Pro Pro Leu Gln Ile
 1 5

 <210> 134
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 10 <400> 134
 Pro Leu Gln Ile Pro Gly Ile Ser Val
 1 5

 <210> 135
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 15 <400> 135
 Ile Pro Gly Ile Ser Val Pro Ala Phe
 1 5

 <210> 136
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 20 <400> 136
 Gly Ile Ser Val Pro Ala Phe Ser Leu
 1 5

 <210> 137
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 25 <400> 137
 Val Pro Ala Phe Ser Leu Pro Ala Ile
 1 5

 <210> 138
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 30 <400> 138
 Ala Phe Ser Leu Pro Ala Ile Thr Leu
 1 5

 <210> 139
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 35 <400> 139
 Leu Pro Ala Ile Thr Leu Pro Ser Leu
 1 5

ES 2 647 321 T3

<210> 140
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 5 <400> 140
 Ala Ile Thr Leu Pro Ser Leu Asn Ile
 1 5

 <210> 141
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 10 <400> 141
 Thr Leu Pro Ser Leu Asn Ile Pro Ala
 1 5

 <210> 142
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 15 <400> 142
 Leu Pro Ser Leu Asn Ile Pro Ala Ala
 1 5

 <210> 143
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 20 <400> 143
 Ala Thr Thr Pro Ala Asn Ile Thr Val
 1 5

 <210> 144
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 25 <400> 144
 Thr Pro Ala Asn Ile Thr Val Gly Ala
 1 5

 <210> 145
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 30 <400> 145
 Asn Ile Thr Val Gly Ala Phe Ser Leu
 1 5

 <210> 146
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 35 <400> 146
 Ala Phe Ser Leu Pro Gly Leu Thr Leu
 1 5

ES 2 647 321 T3

<210> 147
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 5 <400> 147
 Leu Pro Gly Leu Thr Leu Pro Ser Leu
 1 5

 <210> 148
 <211> 9
 <212> PRT
 10 <213> Mycobacterium tuberculosis
 <400> 148
 Thr Leu Pro Ser Leu Asn Ile Pro Ala
 1 5

 <210> 149
 <211> 9
 <212> PRT
 15 <213> Mycobacterium tuberculosis
 <400> 149
 Ala Thr Thr Pro Ala Asn Ile Thr Val
 1 5

 <210> 150
 <211> 9
 <212> PRT
 20 <213> Mycobacterium tuberculosis
 <400> 150
 Asn Ile Thr Val Gly Ala Phe Ser Leu
 1 5

 <210> 151
 <211> 9
 <212> PRT
 25 <213> Mycobacterium tuberculosis
 <400> 151

 <210> 152
 <211> 9
 <212> PRT
 30 <213> Mycobacterium tuberculosis
 <400> 152
 Ala Phe Ser Leu Pro Gly Leu Thr Leu
 1 5

 <210> 153
 <211> 9
 <212> PRT
 35 <213> Mycobacterium tuberculosis
 <400> 153
 Leu Pro Gly Leu Thr Leu Pro Ser Leu
 1 5

 <210> 153
 <211> 9
 <212> PRT
 40 <213> Mycobacterium tuberculosis
 <400> 153
 Thr Leu Pro Ser Leu Asn Ile Pro Ala
 1 5

ES 2 647 321 T3

<210> 154
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 5 <400> 154
 Ala Thr Thr Pro Ala Asn Ile Thr Val
 1 5

 <210> 155
 <211> 9
 <212> PRT
 10 <213> Mycobacterium tuberculosis
 <400> 155
 Asn Ile Thr Val Gly Ala Phe Ser Leu
 1 5

 <210> 156
 <211> 9
 <212> PRT
 15 <213> Mycobacterium tuberculosis
 <400> 156
 Ala Phe Ser Leu Pro Gly Leu Thr Leu
 1 5

 <210> 157
 <211> 9
 <212> PRT
 20 <213> Mycobacterium tuberculosis
 <400> 157
 Leu Pro Gly Leu Thr Leu Pro Ser Leu
 1 5

 <210> 158
 <211> 9
 <212> PRT
 25 <213> Mycobacterium tuberculosis
 <400> 158

 <210> 159
 <211> 9
 <212> PRT
 30 <213> Mycobacterium tuberculosis
 <400> 159
 Thr Leu Pro Ser Leu Asn Ile Pro Ala
 1 5

 <210> 160
 <211> 9
 <212> PRT
 35 <213> Mycobacterium tuberculosis
 <400> 160
 Ala Thr Thr Pro Ala Asn Ile Thr Val
 1 5

 <210> 160
 <211> 9
 <212> PRT
 40 <213> Mycobacterium tuberculosis
 <400> 160
 Asn Ile Thr Val Gly Ala Phe Ser Leu
 1 5

ES 2 647 321 T3

<210> 161
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 5 <400> 161
 Ala Phe Ser Leu Pro Gly Leu Thr Leu
 1 5

 <210> 162
 <211> 9
 <212> PRT
 10 <213> Mycobacterium tuberculosis
 <400> 162
 Leu Pro Gly Leu Thr Leu Pro Ser Leu
 1 5

 <210> 163
 <211> 9
 <212> PRT
 15 <213> Mycobacterium tuberculosis
 <400> 163
 Thr Leu Pro Ser Leu Asn Ile Pro Ala
 1 5

 <210> 164
 <211> 9
 <212> PRT
 20 <213> Mycobacterium tuberculosis
 <400> 164
 Ala Thr Thr Pro Ala Asn Ile Thr Val
 1 5

 <210> 165
 <211> 9
 <212> PRT
 25 <213> Mycobacterium tuberculosis
 <400> 165
 Asn Ile Thr Val Gly Ala Phe Ser Leu
 1 5

 <210> 166
 <211> 9
 <212> PRT
 30 <213> Mycobacterium tuberculosis
 <400> 166
 Leu Pro Gly Leu Thr Leu Pro Ser Leu
 1 5

 <210> 167
 <211> 9
 <212> PRT
 35 <213> Mycobacterium tuberculosis
 <400> 167
 Ala Thr Thr Pro Ala Asn Ile Thr Val
 1 5

ES 2 647 321 T3

<210> 168
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 5 <400> 168
 Pro Ala Asn Ile Thr Val Ser Gly Phe
 1 5

 <210> 169
 <211> 9
 <212> PRT
 10 <213> Mycobacterium tuberculosis
 <400> 169
 Asn Ile Thr Val Ser Gly Phe Gln Leu
 1 5

 <210> 170
 <211> 9
 <212> PRT
 15 <213> Mycobacterium tuberculosis
 <400> 170
 Leu Pro Pro Leu Ser Ile Pro Ser Val
 1 5

 <210> 171
 <211> 9
 <212> PRT
 20 <213> Mycobacterium tuberculosis
 <400> 171
 Pro Pro Leu Ser Ile Pro Ser Val Ala
 1 5

 <210> 172
 <211> 9
 <212> PRT
 25 <213> Mycobacterium tuberculosis
 <400> 172
 Ile Pro Ser Val Ala Ile Pro Pro Val
 1 5

 <210> 173
 <211> 9
 <212> PRT
 30 <213> Mycobacterium tuberculosis
 <400> 173
 Ile Pro Pro Val Thr Val Pro Pro Ile
 1 5

 <210> 174
 <211> 9
 <212> PRT
 40 <213> Mycobacterium tuberculosis
 <400> 174
 Val Pro Pro Ile Thr Val Gly Ala Phe
 1 5

ES 2 647 321 T3

<210> 182
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 5 <400> 182
 Pro Ile Thr Val Gly Gln Ile Gly Val
 1 5
 <210> 183
 <211> 9
 <212> PRT
 10 <213> Mycobacterium tuberculosis
 <400> 183
 Gln Ile Gly Val Gly Gln Phe Gly Leu
 1 5
 <210> 184
 <211> 9
 <212> PRT
 15 <213> Mycobacterium tuberculosis
 <400> 184
 Gly Leu Pro Ser Ile Gly Trp Asp Val
 1 5
 <210> 185
 <211> 9
 <212> PRT
 20 <213> Mycobacterium tuberculosis
 <400> 185
 Leu Pro Ser Ile Gly Trp Asp Val Phe
 1 5
 <210> 186
 <211> 9
 <212> PRT
 25 <213> Mycobacterium tuberculosis
 <400> 186
 Asp Val Phe Leu Ser Thr Pro Arg Ile
 1 5
 <210> 187
 <211> 9
 <212> PRT
 30 <213> Mycobacterium tuberculosis
 <400> 187
 Phe Leu Ser Thr Pro Arg Ile Thr Val
 1 5
 <210> 188
 <211> 9
 <212> PRT
 40 <213> Mycobacterium tuberculosis
 <400> 188

ES 2 647 321 T3

		Thr Pro Arg Ile Thr Val Pro Ala Phe
		1 5
	<210> 189	
	<211> 9	
	<212> PRT	
5	<213> Mycobacterium tuberculosis	
	<400> 189	
		Phe Gly Ile Pro Phe Thr Leu Gln Phe
		1 5
	<210> 190	
	<211> 9	
10	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 190	
		Phe Thr Leu Gln Phe Gln Thr Asn Val
		1 5
	<210> 191	
15	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 191	
		Gln Phe Gln Thr Asn Val Pro Ala Leu
		1 5
20	<210> 192	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 192	
25		Ala Leu Gln Pro Pro Gly Gly Gly Leu
		1 5
	<210> 193	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
30	<400> 193	
		Leu Ser Thr Phe Thr Asn Gly Ala Leu
		1 5
	<210> 194	
	<211> 9	
	<212> PRT	
35	<213> Mycobacterium tuberculosis	
	<400> 194	
		Ser Thr Phe Thr Asn Gly Ala Leu Ile
		1 5
40	<210> 195	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 195	

ES 2 647 321 T3

		Thr Phe Thr Asn Gly Ala Leu Ile Phe
		1 5
5	<210> 196	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 196	
		Ala Leu Ile Phe Gly Glu Phe Asp Leu
		1 5
10	<210> 197	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 197	
		Gly Glu Phe Asp Leu Pro Gln Leu Val
		1 5
15	<210> 198	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 198	
		Leu Pro Gln Leu Val Val His Pro Tyr
		1 5
20	<210> 199	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 199	
		Gln Leu Val Val His Pro Tyr Thr Leu
		1 5
25	<210> 200	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 200	
		His Pro Tyr Thr Leu Thr Gly Pro Ile
		1 5
30	<210> 201	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 201	
		Tyr Thr Leu Thr Gly Pro Ile Val Ile
		1 5
35	<210> 202	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 202	
40	<210> 202	
	<211> 9	
	<212> PRT	
	<213> Mycobacterium tuberculosis	
	<400> 202	

ES 2 647 321 T3

		Gly	Pro	Ile	Val	Ile	Gly	Ser	Phe	Phe
		1				5				
	<210>	203								
	<211>	9								
	<212>	PRT								
5	<213>	Mycobacterium tuberculosis								
	<400>	203								
		Pro	Ile	Val	Ile	Gly	Ser	Phe	Phe	Leu
		1				5				
	<210>	204								
	<211>	9								
	<212>	PRT								
10	<213>	Mycobacterium tuberculosis								
	<400>	204								
		Ser	Phe	Phe	Leu	Pro	Ala	Phe	Asn	Ile
		1				5				
	<210>	205								
	<211>	9								
	<212>	PRT								
15	<213>	Mycobacterium tuberculosis								
	<400>	205								
		Leu	Pro	Ala	Phe	Asn	Ile	Pro	Gly	Ile
		1				5				
	<210>	206								
	<211>	9								
	<212>	PRT								
20	<213>	Mycobacterium tuberculosis								
	<400>	206								
		Ile	Pro	Gly	Ile	Asp	Val	Pro	Ala	Ile
		1				5				
	<210>	207								
	<211>	9								
	<212>	PRT								
	<213>	Mycobacterium tuberculosis								
30	<400>	207								
		Gly	Ile	Asp	Val	Pro	Ala	Ile	Asn	Val
		1				5				
	<210>	208								
	<211>	9								
	<212>	PRT								
35	<213>	Mycobacterium tuberculosis								
	<400>	208								
		Val	Pro	Ala	Ile	Asn	Val	Asp	Gly	Phe
		1				5				
	<210>	209								
	<211>	9								
	<212>	PRT								
40	<213>	Mycobacterium tuberculosis								
	<400>	209								

ES 2 647 321 T3

		Ala Ile Asn Val Asp Gly Phe Thr Leu							
		1		5					
5		<210> 210 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 210							
		Thr Leu Pro Gln Ile Thr Thr Pro Ala		1		5			
10		<210> 211 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 211							
		Leu Pro Gln Ile Thr Thr Pro Ala Ile		1		5			
15		<210> 212 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 212							
		Thr Pro Ala Ile Thr Thr Pro Glu Phe		1		5			
20		<210> 213 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 213							
		Ala Ile Thr Thr Pro Glu Phe Ala Ile		1		5			
25		<210> 214 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 214							
		Thr Pro Glu Phe Ala Ile Pro Pro Ile		1		5			
30		<210> 215 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 215							
		Ile Pro Pro Ile Gly Val Gly Gly Phe		1		5			
35		<210> 216 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 216							
40		<210> 216 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 216							

ES 2 647 321 T3

		Pro Ile Gly Val Gly Gly Phe Thr Leu
		1 5
5		<210> 217 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 217
		Leu Pro Gln Ile Thr Thr Gln Glu Ile
		1 5
10		<210> 218 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 218
		Pro Gln Ile Thr Thr Gln Glu Ile Ile
		1 5
15		<210> 219 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 219
		Glu Ile Ile Thr Pro Glu Leu Thr Ile
		1 5
20		<210> 220 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 220
		Thr Pro Glu Leu Thr Ile Asn Ser Ile
		1 5
25		<210> 221 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 221
		Glu Leu Thr Ile Asn Ser Ile Gly Val
		1 5
30		<210> 222 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 222
		Ser Ile Gly Val Gly Gly Phe Thr Leu
		1 5
35		<210> 223 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 223
40		<210> 223 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 223

ES 2 647 321 T3

		Leu Pro Gln Ile Thr Thr Pro Pro Ile
		1 5
5	<210> 224 <211> 9 <212> PRT <213> Mycobacterium tuberculosis	
	<400> 224	
		Thr Pro Pro Ile Thr Thr Pro Pro Leu
		1 5
10	<210> 225 <211> 9 <212> PRT <213> Mycobacterium tuberculosis	
	<400> 225	
		Pro Ile Thr Thr Pro Pro Leu Thr Ile
		1 5
15	<210> 226 <211> 9 <212> PRT <213> Mycobacterium tuberculosis	
	<400> 226	
		Thr Pro Pro Leu Thr Ile Asp Pro Ile
		1 5
20	<210> 227 <211> 9 <212> PRT <213> Mycobacterium tuberculosis	
	<400> 227	
		Pro Ile Asn Leu Thr Gly Phe Thr Leu
		1 5
25	<210> 228 <211> 9 <212> PRT <213> Mycobacterium tuberculosis	
	<400> 228	
		Thr Pro Pro Leu Thr Ile Glu Pro Ile
		1 5
30	<210> 229 <211> 9 <212> PRT <213> Mycobacterium tuberculosis	
	<400> 229	
		Pro Leu Thr Ile Glu Pro Ile Gly Val
		1 5
35	<210> 230 <211> 9 <212> PRT <213> Mycobacterium tuberculosis	
	<400> 230	
40	<210> 230 <211> 9 <212> PRT <213> Mycobacterium tuberculosis	
	<400> 230	

ES 2 647 321 T3

		Ile Glu Pro Ile Gly Val Gly Gly Phe
		1 5
5	<210> 231 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 231	
		Pro Pro Leu Thr Val Pro Gly Ile His
		1 5
10	<210> 232 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 232	
		Pro Leu Thr Val Pro Gly Ile His Leu
		1 5
15	<210> 233 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 233	
		Gly Ile His Leu Pro Ser Thr Thr Ile
		1 5
20	<210> 234 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 234	
		His Leu Pro Ser Thr Thr Ile Gly Ala
		1 5
25	<210> 235 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 235	
		Leu Pro Ser Thr Thr Ile Gly Ala Phe
		1 5
30	<210> 236 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 236	
		Phe Ala Ile Pro Gly Gly Pro Gly Tyr
		1 5
35	<210> 237 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 237	
40	<210> 237 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 237	

ES 2 647 321 T3

		Ser Thr Ala Pro Ser Ser Gly Phe Phe
		1 5
5		<210> 238 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 238
		Trp Phe Asn Thr Asn Pro Ala Gly Leu
		1 5
10		<210> 239 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 239
		Asn Phe Gly Gly Leu Ser Ser Gly Phe
		1 5
15		<210> 240 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 240
		Gly Leu Ser Ser Gly Phe Ser Asn Leu
		1 5
20		<210> 241 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 241
		Asn Leu Gly Ser Gly Val Ser Gly Phe
		1 5
25		<210> 242 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 242
		Phe Ala Asn Arg Gly Ile Leu Pro Phe
		1 5
30		<210> 243 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 243
		Asn Arg Gly Ile Leu Pro Phe Ser Val
		1 5
35		<210> 244 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 244
40		<210> 244 <211> 9 <212> PRT <213> Mycobacterium tuberculosis <400> 244

ES 2 647 321 T3

Ile Leu Pro Phe Ser Val Ala Ser Val
 1 5

5 <210> 245
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 245

Leu Pro Phe Ser Val Ala Ser Val Val
 1 5

10 <210> 246
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 246

Ser Val Ala Ser Val Val Ser Gly Phe
 1 5

15 <210> 247
 <211> 9
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 247

Gly Phe Ala Asn Ile Gly Thr Asn Leu
 1 5

20 <210> 248
 <211> 450
 <212> PRT
 <213> Mycobacterium tuberculosis
 <400> 248

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
 25 35 40 45

ES 2 647 321 T3

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

ES 2 647 321 T3

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> 249

<211> 324

<212> PRT

5 <213> Mycobacterium tuberculosis

<400> 249

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

REIVINDICACIONES

1. Un polipéptido aislado, que comprende:
 - (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 - (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
 - (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 para su uso en el tratamiento o prevención de la tuberculosis.
2. Un polipéptido para su uso de acuerdo con la reivindicación 1, que comprende una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7.
3. Un polipéptido para su uso de acuerdo con la reivindicación 1, que comprende una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7.
4. Un polipéptido para su uso de acuerdo con la reivindicación 1, que comprende un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7.
5. Un polipéptido para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 4, para su uso en el tratamiento de la tuberculosis.
6. Un polipéptido para su uso de acuerdo con la reivindicación 5, para su uso en el tratamiento de la tuberculosis latente.
7. Un polipéptido para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 4, para su uso en la prevención de la tuberculosis.
8. Un polipéptido para su uso de acuerdo con la reivindicación 7, para su uso en la prevención de la tuberculosis latente.
9. Un polipéptido para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 4, para su uso en la prevención de la reactivación de la tuberculosis.
10. Un polipéptido para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 4, para su uso en el retraso de la reactivación de la tuberculosis.
11. Un polipéptido para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 10, en el que la tuberculosis se asocia con una infección por *Mycobacterium tuberculosis*.
12. Un polinucleótido aislado que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende:
 - (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 - (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
 - (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 para su uso en el tratamiento o prevención de la tuberculosis.
13. Un polinucleótido para su uso de acuerdo con la reivindicación 12, que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una secuencia de proteína Rv1753c de SEQ ID No: 1 o 3-7.
14. Un polinucleótido para su uso de acuerdo con la reivindicación 12, que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una variante de la secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7.
15. Un polinucleótido para su uso de acuerdo con la reivindicación 12, que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7.
16. Un polinucleótidos para su uso de acuerdo con una cualquiera de las reivindicaciones 12 a 15, para su uso en el tratamiento de la tuberculosis.
17. Un polinucleótido para su uso de acuerdo con la reivindicación 16, para su uso en el tratamiento de la tuberculosis latente.
18. Un polinucleótidos para su uso de acuerdo con una cualquiera de las reivindicaciones 12 a 15, para su uso en la prevención de la tuberculosis.

19. Un polinucleótido para su uso de acuerdo con la reivindicación 18, para su uso en la prevención de la tuberculosis latente.
20. Un polinucleótidos para su uso de acuerdo con una cualquiera de las reivindicaciones 12 a 15, para su uso en la prevención de la reactivación de la tuberculosis.
- 5 21. Un polinucleótidos para su uso de acuerdo con una cualquiera de las reivindicaciones 12 a 15, para su uso en el retraso de la reactivación de la tuberculosis.
22. Un polinucleótidos para su uso de acuerdo con una cualquiera de las reivindicaciones 12 a 15, en el que la tuberculosis se asocia con una infección por *Mycobacterium tuberculosis*.
- 10 23. Una composición que comprende un componente Rv1753c y un componente M72, en la que dicho componente Rv1753c es un polipéptido que comprende:
- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 - (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
 - (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- 15 y dicho componente M72 es:
- (i) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos el 90 % con la SEQ ID No: 25; o
 - (ii) un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos el 90 % con la SEQ ID No: 25;
- 20 y en el que dicho componente M72 se proporciona en forma de un polipéptido, el componente Rv1753c y un componente M72 se pueden proporcionar como dos componentes polipeptídicos individuales o como una proteína de fusión que comprende ambos componentes polipeptídicos.
24. Una composición que comprende un componente Rv1753c y un componente M72, en la que dicho componente Rv1753c es un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que
- 25 comprende:
- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 - (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
 - (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- 30 y dicho componente M72 es:
- (i) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos el 90 % con la SEQ ID No: 25; o
 - (ii) un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos el 90 % con la SEQ ID No: 25;
- 35 y en la que dicho componente M72 se proporciona en forma de un polinucleótido, el componente Rv1753c y un componente M72 se pueden proporcionar como dos componentes polinucleotídicos individuales, como un solo polinucleótido que codifica dos componentes polipeptídicos individuales o como un solo polinucleótido que codifica una proteína de fusión que comprende ambos componentes polipeptídicos.
25. Una composición que comprende:
- 40 (a) un polipéptido que comprende:
- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 - (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
 - (iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
- 45 y
- (b) monofosforil lípido A 3-de-O-acilado y QS21 en una formulación de liposoma.
26. Uso de un polipéptido que comprende:
- (i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;
 - (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o
- 50

(iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7,

o un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende:

(i) una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7;

5 (ii) una variante de una secuencia de la proteína Rv1753c que tiene una identidad de al menos el 90 % con la SEQ ID No: 1 o 3-7; o

(iii) un fragmento inmunogénico de una secuencia de la proteína Rv1753c de SEQ ID No: 1 o 3-7,

en la fabricación de un medicamento para el tratamiento o la prevención de la tuberculosis.

Figura 1

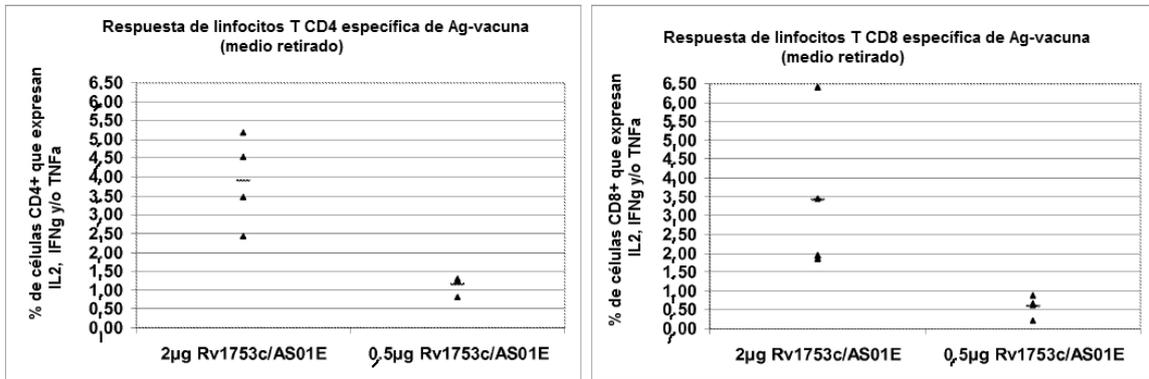


Figura 2

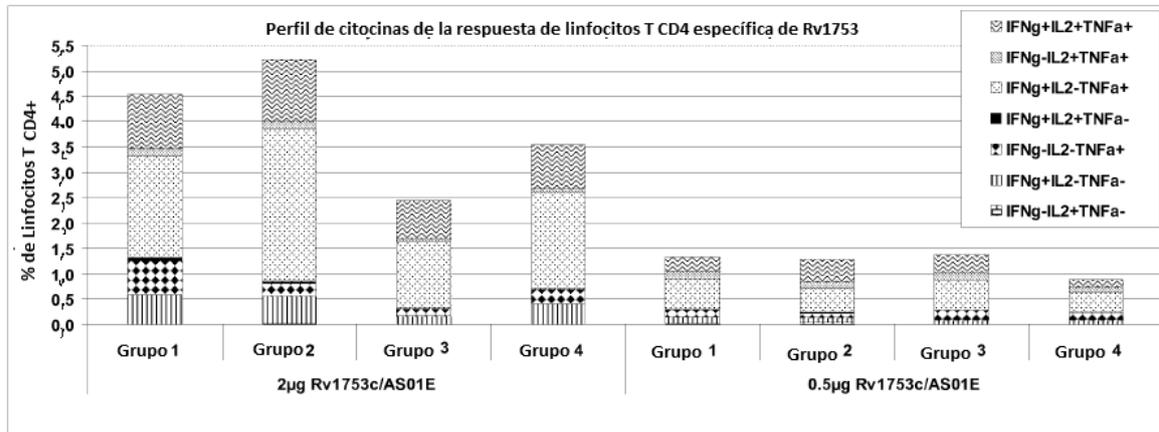


Figura 3

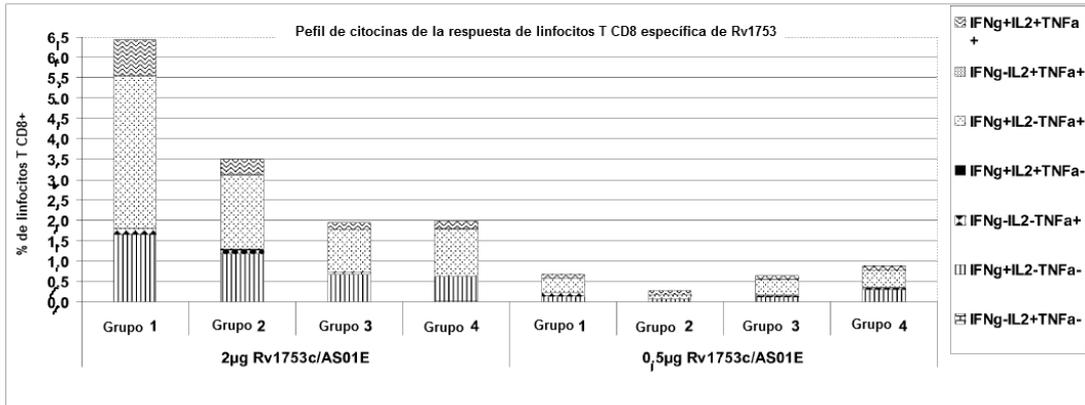


Figura 4

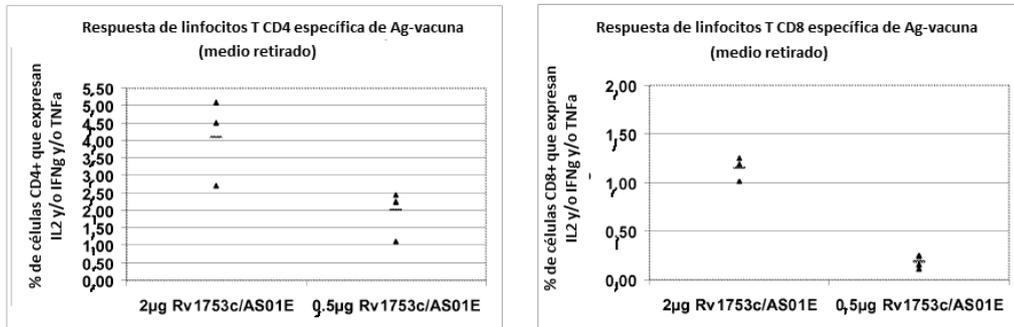


Figura 5

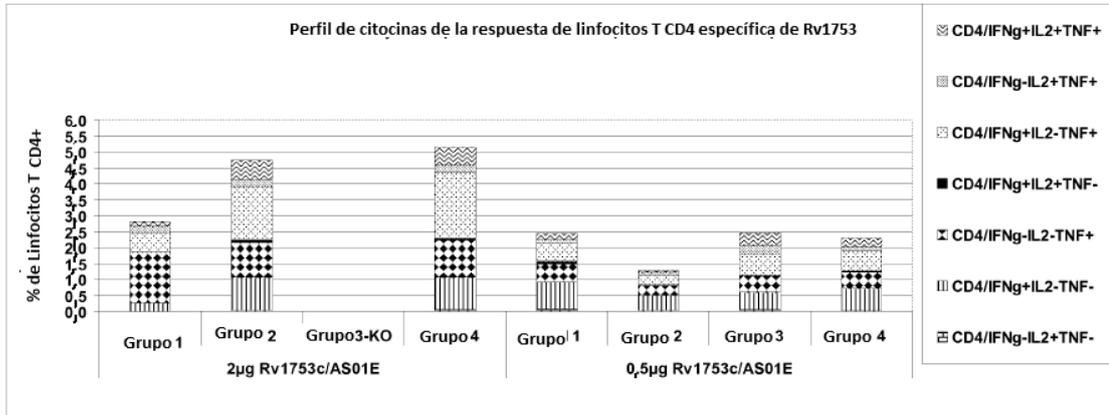


Figura 6

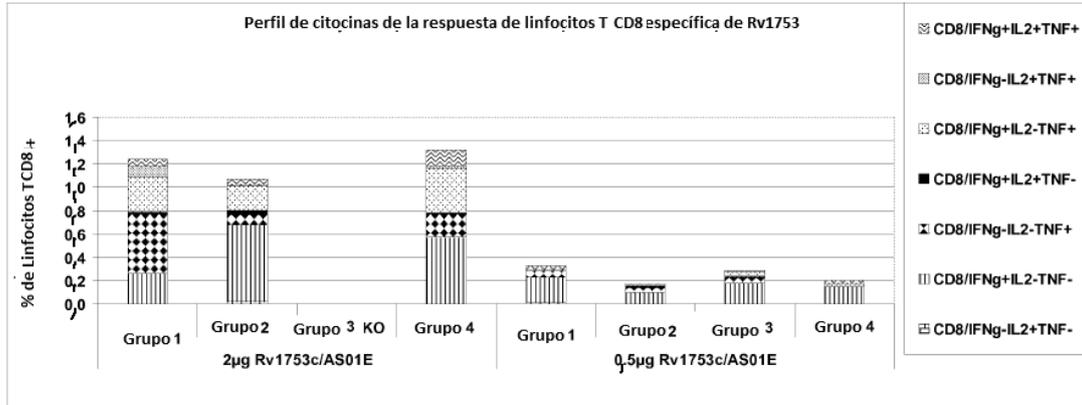


Figura 7

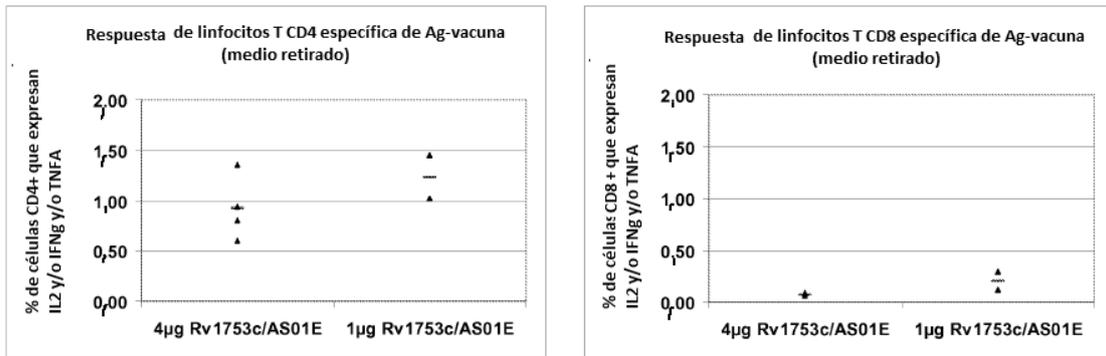


Figura 8

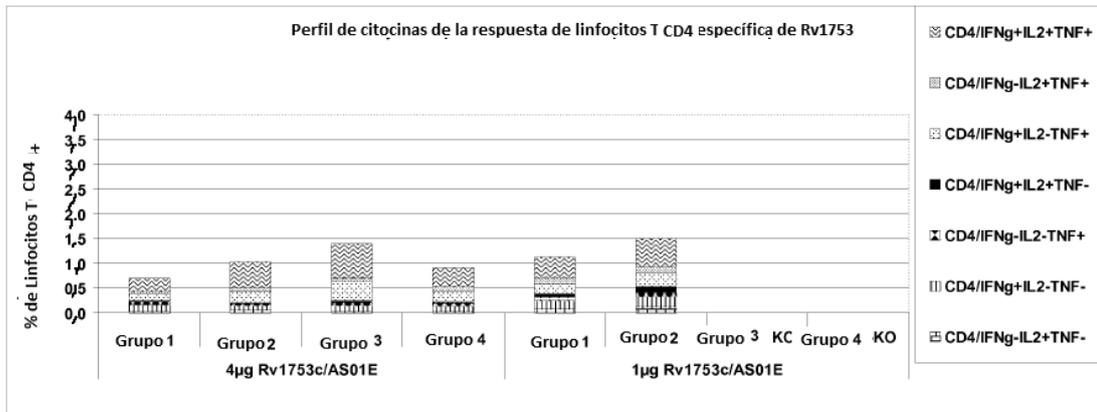


Figura 9

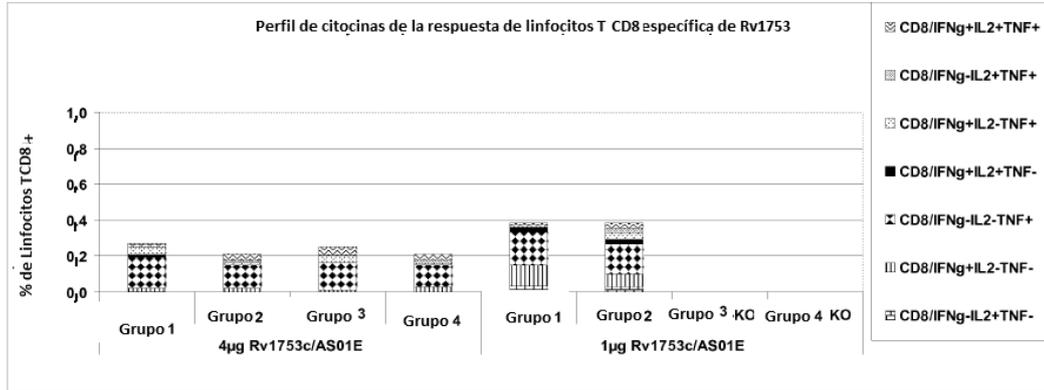


Figura 10

