

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 685 498**

51 Int. Cl.:

**C07K 14/35** (2006.01)  
**C12N 9/88** (2006.01)  
**C12N 9/90** (2006.01)  
**C12N 15/60** (2006.01)  
**C12N 15/61** (2006.01)  
**A61K 38/51** (2006.01)  
**A61K 38/52** (2006.01)  
**A61K 39/04** (2006.01)  
**A61K 38/00** (2006.01)  
**A61K 39/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **24.07.2009 PCT/EP2009/059585**
- 87 Fecha y número de publicación internacional: **28.01.2010 WO10010179**
- 96 Fecha de presentación y número de la solicitud europea: **24.07.2009 E 09781057 (6)**
- 97 Fecha y número de publicación de la concesión europea: **13.06.2018 EP 2315834**

54 Título: **Proteína Rv2386c de la tuberculosis, composiciones y usos de las mismas**

30 Prioridad:

**25.07.2008 US 83699 P**

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

**09.10.2018**

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:

**ELZABURU, S.L.P**

ES 2 685 498 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

Proteína Rv2386c de la tuberculosis, composiciones y usos de las mismas

**Campo de la invención**

5 La presente invención se refiere a polipéptidos y polinucleótidos para el uso en el tratamiento o prevención de la tuberculosis, en particular para el uso en el tratamiento o prevención de la tuberculosis latente, y en la prevención o el retraso de la reactivación de la tuberculosis (y también a los métodos 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 más de 2 billones de personas están infectadas con los bacilos de TB, con aproximadamente 9,2 millones de casos nuevos de TB y 1,7 millones de muertes cada año. El 10% de las personas infectadas con los bacilos de TB desarrollarán la TB activa, infectando cada persona con TB activa a un promedio de 10 a 15 personas más por año. Aunque los índices de incidencia anuales han llegado a su punto máximo a nivel mundial, todavía se está elevando el número de muertes y casos debido al crecimiento de la población (World Health Organisation (Organización Mundial de la Salud, de sus siglas en inglés) *Tuberculosis Facts* 2008).

20 *Mycobacterium tuberculosis* infecta a los individuos a través de las vías respiratorias. Los macrófagos alveolares envuelven a la bacteria, pero ésta es capaz de sobrevivir y proliferar mediante la inhibición de la fusión del fagosoma con los lisosomas ácidos. Sigue una compleja respuesta inmunológica que implica a las células T CD4+ y CD8+, dando como resultado final la formación de un granuloma. Es fundamental para el éxito de *Mycobacterium tuberculosis* como un patógeno, el hecho de que la bacteria aislada, pero no erradicada, pueda persistir durante largos períodos, dejando al individuo vulnerable al desarrollo posterior de la TB activa.

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

También se ha propuesto una hipótesis dinámica que describe el equilibrio entre la TB latente y la TB activa (Cardana P-J *Inflammation & 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 de tiempo considerable, la enfermedad activa se manifiesta más comúnmente como una inflamación aguda de los pulmones, que da como resultado cansancio, pérdida de peso, fiebre y una tos persistente. Si no se trata, normalmente da como resultado serias complicaciones y la muerte.

40 En términos generales, la tuberculosis se puede controlar empleando una terapia antibiótica prolongada, aunque tal tratamiento no es suficiente para prevenir la propagación de la enfermedad. Los individuos infectados pueden ser asintomáticos, pero contagiosos, durante algún tiempo. Además, aunque el cumplimiento con el régimen de tratamiento es crítico, es difícil monitorear el comportamiento del paciente. Algunos pacientes no llevan a cabo el plan de tratamiento, lo cual puede conducir a un tratamiento ineficaz y al desarrollo de resistencia al fármaco.

45 La TB resistente a múltiples fármacos (MDR-TB) es una forma que fracasa para responder a los medicamentos de primera línea. El 5% de todos los casos de TB son MDR-TB, presentándose cada año una estimación de 490.000 nuevos casos de MDR-TB. La TB extensamente resistente a fármacos (XDR-TB) se presenta cuando se desarrolla la resistencia a los medicamentos de segunda línea además de la MDR-TB. Se estima que anualmente se presentan 40.000 nuevos casos de MDR-TB prácticamente intratables (World Health Organisation *Tuberculosis Facts* 2008).

50 Aún cuando se lleve a cabo el tratamiento con antibiótico, la infección con *M. tuberculosis* puede no ser erradicada del individuo infectado, y puede permanecer como una infección latente que puede ser reactivada.

Para el control de la propagación de la tuberculosis, es de suma importancia una vacunación eficaz y un diagnóstico precoz exacto de la enfermedad.

El diagnóstico de la infección de TB latente se consigue normalmente utilizando la prueba cutánea de tuberculina, que implica la exposición intradérmica al derivado proteico purificado de la tuberculina (PPD). Las respuestas de

células T específicas del antígeno dan como resultado una induración cuantificable en el sitio de la inyección de 48 a 72 horas después de la inyección, lo cual indica la exposición a los antígenos micobacterianos. Sin embargo, con esta prueba, la sensibilidad y la especificidad han sido un problema, y los individuos vacunados con BCG no siempre se pueden distinguir fácilmente de los individuos infectados (esto 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 que están infectados por *M. tuberculosis* muestran una reacción al PPD por debajo de 10 mm de diámetro, mientras que se considera que las personas que tienen una reacción al PPD por encima de 10 mm de diámetro, han sido infectadas por *M. tuberculosis*. Sin embargo, esta regla no es aplicable a los individuos con inmunosupresión debido a la infección por VIH, lo cual puede dar como resultado una reacción al PPD por debajo de 10 mm de diámetro; o en los países endémicos, en donde las personas infectadas por micobacterias que no son de tuberculosis, pueden mostrar una reacción al PPD por encima de 10 mm de diámetro.

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

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

#### Breve compendio de la invención

La presente invención se refiere, en términos generales, a la identificación de Rv2386c como un antígeno de la TB (en particular un antígeno asociado con la TB latente), y a los usos relacionados en la prevención y el tratamiento de la TB, especialmente en la prevención y el tratamiento de la TB latente y en la prevención o demora de la reactivación de la TB.

La presente invención proporciona un polipéptido aislado, que comprende:

- (i) una secuencia proteica de Rv2386c;
  - (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
  - (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; para el uso en:
- (a) el tratamiento de la tuberculosis;
  - (b) el tratamiento de la tuberculosis latente;
  - (c) la prevención de la tuberculosis latente;
  - (d) la prevención de la reactivación de la tuberculosis; o
  - (e) el retraso de la reactivación de la tuberculosis.

El uso de un polipéptido que comprende:

- (i) una secuencia proteica de Rv2386c;
  - (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
  - (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- en donde la secuencia proteica se selecciona de SEQ ID NO: 1 y 6; en la fabricación de un medicamento para:
- (a) el tratamiento de la tuberculosis;
  - (b) el tratamiento de la tuberculosis latente;
  - (c) la prevención de la tuberculosis latente;

- (d) la prevención de la reactivación de la tuberculosis; o
  - (e) el retraso de la reactivación de la tuberculosis
- representa otro aspecto de la invención.

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

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c

10 en donde la secuencia proteica Rv2386c se selecciona de SEQ ID NO: 1 y 6; para el uso en:

- (a) el tratamiento de la tuberculosis latente;
- (b) la prevención de la tuberculosis latente;
- (c) la prevención de la reactivación de la tuberculosis; o
- (d) el retraso de la reactivación de la tuberculosis.

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

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;

20 en donde la secuencia proteica Rv2386c se selecciona de SEQ ID NO: 1 y 6; en la fabricación de un medicamento para:

- (a) el tratamiento de la tuberculosis latente;
- (b) la prevención de la tuberculosis latente;
- (c) la prevención de la reactivación de la tuberculosis; o
- (d) el retraso de la reactivación de la tuberculosis

representa otro aspecto de la invención.

Además, se proporciona una composición inmunogénica, que comprende:

30 (a) un polipéptido que comprende:

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;

35 en donde la secuencia de Rv2386c se selecciona de SEQ ID NO: 1 y 6; y

(b) un potenciador de la respuesta inmunitaria no específica.

La composición inmunogénica comprende de manera adecuada una combinación de Rv2386c y un segundo componente heterólogo del antígeno de la tuberculosis.

Además, se proporciona una composición que comprende:

40 (a) un polipéptido que comprende:

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma;  
o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;

5 en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; o

- (b) un polinucleótido que comprende una secuencia de ácido nucleico que codifica el polipéptido de (a); y

un componente M72.

Además, se proporciona una composición que comprende:

10 (a) un polipéptido que comprende:

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma;  
o

- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;

15 en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; y

- (b) 3D-MPL y QS21 en una formulación liposomal.

Además, se proporciona una proteína de fusión que comprende un polipéptido que comprende:

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de la identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;

20

en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; para el uso en:

- (a) el tratamiento de la tuberculosis;
- (b) el tratamiento de la tuberculosis latente;
- (c) la prevención de la tuberculosis latente;
- (d) la prevención de la reactivación de la tuberculosis; o
- (e) el retraso de la reactivación de la tuberculosis.

25

30 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 la infección por *M. tuberculosis*).

35 Un sujeto que recibe un polipéptido, polinucleótido o composición de la invención, puede haber sido previamente vacunado para tuberculosis (por ejemplo, vacunado contra la infección por *M. tuberculosis*), tal como que haya sido vacunado con *Bacillus Calmette-Guerin* (BCG). Alternativamente, un sujeto que recibe un polipéptido, polinucleótido o composición de la invención, puede no haber sido vacunado previamente para tuberculosis (por ejemplo, no vacunado contra la infección por *M. tuberculosis*), tal como que no haya sido vacunado con *Bacillus Calmette-Guerin* (BCG).

#### Descripción de las figuras

40 Figura 1: Porcentaje de células CD4 y CD8 a partir de ratones CB6F1 inmunizados que expresan citoquinas de 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 citoquinas el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD4 específica del antígeno en ratones CB6F1 inmunizados.

Figura 3: Perfil de citoquinas el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD8 específica del antígeno en ratones CB6F1 inmunizados.

Figura 4: Porcentaje de células CD4 y CD8 a partir de ratones CB6F1 inmunizados que expresan citoquinas de IFN-gamma y/o IL-2 y/o TNF-alfa el día 35 (es decir, 7 días después de la tercera inmunización).

5 Figura 5: Perfil de citoquinas el día 35 (es decir, 7 días después de la tercera inmunización) de la respuesta de CD4 específica del antígeno en ratones CB6F1 inmunizados.

Figura 6: Perfil de citoquinas el día 35 (es decir, 7 días después de la tercera inmunización) de la respuesta de CD8 específica del antígeno en ratones CB6F1 inmunizados.

10 Figura 7: Porcentaje de células CD4 y CD8 a partir de ratones C57BL/6 inmunizados que expresan citoquinas de 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 citoquinas el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD4 específica del antígeno en ratones C57BL/6 inmunizados.

Figura 9: Perfil de citoquinas el día 21 (es decir, 7 días después de la segunda inmunización) de la respuesta de CD8 específica del antígeno en ratones C57BL/6 inmunizados.

#### 15 **Descripción del listado de secuencias**

SEQ ID No: 1: secuencia polipeptídica de Rv2386c a partir de *M. tuberculosis*, cepa H37Rv.

SEQ ID No: 2: secuencia polipeptídica de Rv2386c a partir de *M. tuberculosis*, cepa H37Rv.

SEQ ID No: 3: secuencia polipeptídica de Rv2386c a partir de *M. tuberculosis*, cepa CDC1551.

SEQ ID No: 4: secuencia polipeptídica de Rv2386c a partir de *M. tuberculosis*, cepa F11.

20 SEQ ID No: 5: secuencia polipeptídica de Rv2386c a partir de *M. tuberculosis*, cepa Haarlem A.

SEQ ID No: 6: secuencia polipeptídica de Rv2386c a partir de *M. tuberculosis*, cepa C.

SEQ ID No: 7: secuencia polipeptídica de Rv2386c a partir de BCG.

SEQ ID No: 8: secuencia polipeptídica de Mtb8.4.

SEQ ID No: 9: secuencia polipeptídica de Mtb9.8.

25 SEQ ID No: 10: secuencia polipeptídica de Mtb9.9.

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.

30 SEQ ID No: 15: secuencia polipeptídica de Mtb41.

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.

35 SEQ ID No: 20: secuencia polipeptídica de MPT64.

SEQ ID No: 21: secuencia polipeptídica de Mtb32A.

SEQ ID No: 22: secuencia polipeptídica de Mtb32A maduro mutado en Ser/Ala.

SEQ ID No: 23: secuencia polipeptídica de TB10.4.

SEQ ID No: 24: secuencia polipeptídica de Mtb72f.

40 SEQ ID No: 25: secuencia polipeptídica de M72.

- SEQ ID No: 26: secuencia polipeptídica de Mtb71f.
- SEQ ID No: 27: secuencia polipeptídica de fusión M92.
- SEQ ID No: 28: secuencia polipeptídica de fusión M103.
- SEQ ID No: 29: secuencia polipeptídica de fusión M114.
- 5 SEQ ID No: 30: supuesto epítipo 1 de células CD4 humanas.
- SEQ ID No: 31: supuesto epítipo 2 de células CD4 humanas
- SEQ ID No: 32: supuesto epítipo 3 de células CD4 humanas.
- SEQ ID No: 33: supuesto epítipo 4 de células CD4 humanas
- SEQ ID No: 34: supuesto epítipo 5 de células CD4 humanas
- 10 SEQ ID No: 35: supuesto epítipo 6 de células CD4 humanas
- SEQ ID No: 36: supuesto epítipo 7 de células CD4 humanas.
- SEQ ID No: 37: supuesto epítipo 8 de células CD4 humanas.
- SEQ ID No: 38: supuesto epítipo 9 de células CD4 humanas.
- SEQ ID No: 39: supuesto epítipo 10 de células CD4 humanas.
- 15 SEQ ID No: 40: supuesto epítipo 11 de células CD4 humanas.
- SEQ ID No: 41: supuesto epítipo 12 de células CD4 humanas.
- SEQ ID No: 42: supuesto epítipo 13 de células CD4 humanas.
- SEQ ID No: 43: supuesto epítipo 14 de células CD4 humanas.
- SEQ ID No: 44: supuesto epítipo 15 de células CD4 humanas.
- 20 SEQ ID No: 45: supuesto epítipo 16 de células CD4 humanas.
- SEQ ID No: 46: supuesto epítipo 17 de células CD4 humanas.
- SEQ ID No: 47: supuesto epítipo 18 de células CD4 humanas.
- SEQ ID No: 48: supuesto epítipo 19 de células CD4 humanas.
- SEQ ID No: 49: supuesto epítipo 20 de células CD4 humanas.
- 25 SEQ ID No: 50: supuesto epítipo 21 de células CD4 humanas.
- SEQ ID No: 51: supuesto epítipo 22 de células CD4 humanas.
- SEQ ID No: 52: supuesto epítipo 23 de células CD4 humanas.
- SEQ ID No: 53: supuesto epítipo 1 de células CD8 humanas.
- SEQ ID No: 54: supuesto epítipo 2 de células CD8 humanas.
- 30 SEQ ID No: 55: supuesto epítipo 3 de células CD8 humanas.
- SEQ ID No: 56: supuesto epítipo 4 de células CD8 humanas.
- SEQ ID No: 57: supuesto epítipo 5 de células CD8 humanas.
- SEQ ID No: 58: supuesto epítipo 6 de células CD8 humanas.
- SEQ ID No: 59: supuesto epítipo 7 de células CD8 humanas.
- 35 SEQ ID No: 60: supuesto epítipo 8 de células CD8 humanas.
- SEQ ID No: 61: supuesto epítipo 9 de células CD8 humanas.
- SEQ ID No: 62: supuesto epítipo 10 de células CD8 humanas.

- SEQ ID No: 63: supuesto epítoto 11 de células CD8 humanas.
- SEQ ID No: 64: supuesto epítoto 12 de células CD8 humanas.
- SEQ ID No: 65: supuesto epítoto 13 de células CD8 humanas.
- SEQ ID No: 66: supuesto epítoto 14 de células CD8 humanas.
- 5 SEQ ID No: 67: supuesto epítoto 15 de células CD8 humanas.
- SEQ ID No: 68: supuesto epítoto 16 de células CD8 humanas.
- SEQ ID No: 69: supuesto epítoto 17 de células CD8 humanas.
- SEQ ID No: 70: supuesto epítoto 18 de células CD8 humanas.
- SEQ ID No: 71: supuesto epítoto 19 de células CD8 humanas.
- 10 SEQ ID No: 72: supuesto epítoto 20 de células CD8 humanas.
- SEQ ID No: 73: supuesto epítoto 21 de células CD8 humanas.
- SEQ ID No: 74: supuesto epítoto 22 de células CD8 humanas.
- SEQ ID No: 75: supuesto epítoto 23 de células CD8 humanas.
- SEQ ID No: 76: supuesto epítoto 24 de células CD8 humanas.
- 15 SEQ ID No: 77: supuesto epítoto 25 de células CD8 humanas.
- SEQ ID No: 78: supuesto epítoto 26 de células CD8 humanas.
- SEQ ID No: 79: supuesto epítoto 27 de células CD8 humanas
- SEQ ID No: 80: supuesto epítoto 28 de células CD8 humanas.
- SEQ ID No: 81: supuesto epítoto 29 de células CD8 humanas.
- 20 SEQ ID No: 82: supuesto epítoto 30 de células CD8 humanas.
- SEQ ID No: 83: supuesto epítoto 31 de células CD8 humanas.
- SEQ ID No: 84: supuesto epítoto 32 de células CD8 humanas.
- SEQ ID No: 85: supuesto epítoto 33 de células CD8 humanas.
- SEQ ID No: 86: supuesto epítoto 34 de células CD8 humanas.
- 25 SEQ ID No: 87: supuesto epítoto 35 de células CD8 humanas.
- SEQ ID No: 88: supuesto epítoto 36 de células CD8 humanas.
- SEQ ID No: 89: supuesto epítoto 37 de células CD8 humanas.
- SEQ ID No: 90: supuesto epítoto 38 de células CD8 humanas.
- SEQ ID No: 91: supuesto epítoto 39 de células CD8 humanas.
- 30 SEQ ID No: 92: supuesto epítoto 40 de células CD8 humanas.
- SEQ ID No: 93: supuesto epítoto 41 de células CD8 humanas.
- SEQ ID No: 94: supuesto epítoto 42 de células CD8 humanas.
- SEQ ID No: 95: supuesto epítoto 43 de células CD8 humanas.
- SEQ ID No: 96: supuesto epítoto 44 de células CD8 humano.
- 35 SEQ ID No: 97: supuesto epítoto 45 de células CD8 humanas.
- SEQ ID No: 98: supuesto epítoto 46 de células CD8 humanas.
- SEQ ID No: 99: supuesto epítoto 47 de células CD8 humanas.

- SEQ ID No: 100: supuesto epítoto 48 de células CD8 humanas.
- SEQ ID No: 101: supuesto epítoto 49 de células CD8 humanas.
- SEQ ID No: 102: supuesto epítoto 50 de células CD8 humanas.
- SEQ ID No: 103: supuesto epítoto 51 de células CD8 humanas.
- 5 SEQ ID No: 104: supuesto epítoto 52 de células CD8 humanas.
- SEQ ID No: 105: supuesto epítoto 53 de células CD8 humanas.
- SEQ ID No: 106: supuesto epítoto 54 de células CD8 humanas.
- SEQ ID No: 107: supuesto epítoto 55 de células CD8 humanas.
- SEQ ID No: 108: supuesto epítoto 56 de células CD8 humanas.
- 10 SEQ ID No: 109: supuesto epítoto 57 de células CD8 humanas.
- SEQ ID No: 110: supuesto epítoto 58 de células CD8 humanas.
- SEQ ID No: 111: supuesto epítoto 59 de células CD8 humanas.
- SEQ ID No: 112: supuesto epítoto 60 de células CD8 humanas.
- SEQ ID No: 113: supuesto epítoto 61 de células CD8 humanas.
- 15 SEQ ID No: 114: supuesto epítoto 62 de células CD8 humanas.
- SEQ ID No: 115: supuesto epítoto 63 de células CD8 humanas.
- SEQ ID No: 116: supuesto epítoto 64 de células CD8 humanas.
- SEQ ID No: 117: supuesto epítoto 65 de células CD8 humanas.
- SEQ ID No: 118: supuesto epítoto 66 de células CD8 humanas.
- 20 SEQ ID No: 119: supuesto epítoto 67 de células CD8 humanas.
- SEQ ID No: 120: supuesto epítoto 68 de células CD8 humanas.
- SEQ ID No: 121: supuesto epítoto 69 de células CD8 humanas.
- SEQ ID No: 122: supuesto epítoto 70 de células CD8 humanas.
- SEQ ID No: 123: supuesto epítoto 71 de células CD8 humanas.
- 25 SEQ ID No: 124: supuesto epítoto 72 de células CD8 humanas.
- SEQ ID No: 125: supuesto epítoto 73 de células CD8 humanas.
- SEQ ID No: 126: supuesto epítoto 74 de células CD8 humanas.
- SEQ ID No: 127: supuesto epítoto 75 de células CD8 humanas.
- SEQ ID No: 128: supuesto epítoto 76 de células CD8 humanas.
- 30 SEQ ID No: 129: supuesto epítoto 77 de células CD8 humanas.
- SEQ ID No: 130: supuesto epítoto 78 de células CD8 humanas.
- SEQ ID No: 131: supuesto epítoto 79 de células CD8 humanas.
- SEQ ID No: 132: supuesto epítoto 80 de células CD8 humanas.
- SEQ ID No: 133: supuesto epítoto 81 de células CD8 humanas.
- 35 SEQ ID No: 134: supuesto epítoto 82 de células CD8 humanas.
- SEQ ID No: 135: supuesto epítoto 83 de células CD8 humanas.
- SEQ ID No: 136: supuesto epítoto 84 de células CD8 humanas.

- SEQ ID No: 137: supuesto epítipo 85 de células CD8 humanas.
- SEQ ID No: 138: supuesto epítipo 86 de células CD8 humanas.
- SEQ ID No: 139: supuesto epítipo 87 de células CD8 humanas.
- SEQ ID No: 140: supuesto epítipo 88 de células CD8 humanas.
- 5 SEQ ID No: 141: supuesto epítipo 89 de células CD8 humanas.
- SEQ ID No: 142: supuesto epítipo 90 de células CD8 humanas.
- SEQ ID No: 143: supuesto epítipo 91 de células CD8 humanas.
- SEQ ID No: 144: supuesto epítipo 92 de células CD8 humanas.
- SEQ ID No: 145: supuesto epítipo 93 de células CD8 humanas.
- 10 SEQ ID No: 146: supuesto epítipo 94 de células CD8 humanas.
- SEQ ID No: 147: supuesto epítipo 95 de células CD8 humanas.
- SEQ ID No: 148: supuesto epítipo 96 de células CD8 humanas.
- SEQ ID No: 149: supuesto epítipo 97 de células CD8 humanas.
- SEQ ID No: 150: supuesto epítipo 98 de células CD8 humanas.
- 15 SEQ ID No: 151: supuesto epítipo 99 de células CD8 humanas.
- SEQ ID No: 152: supuesto epítipo 100 de células CD8 humanas.
- SEQ ID No: 153: supuesto epítipo 101 de células CD8 humanas.
- SEQ ID No: 154: supuesto epítipo 102 de células CD8 humanas.
- SEQ ID No: 155: secuencia polipeptídica de Rv1753c a partir de *M. tuberculosis*, cepa H37Rv.
- 20 SEQ ID No: 156: secuencia polipeptídica de Rv2707c a partir de *M. tuberculosis*, cepa H37Rv.

### Descripción detallada

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

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

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

Las vacunas posteriores a la exposición que se dirigen directamente hacia las bacterias persistentes latentes pueden ayudar a proteger contra la reactivación de la TB, mejorando de esta manera el control de la TB, o incluso permitiendo la eliminación de la infección. Por lo tanto, una vacuna que se dirija hacia la TB latente podría reducir significativa y económicamente los índices globales de infección por TB.

También se podrían utilizar las vacunas subunitarias basadas en antígenos de etapa tardía en combinación con

antígenos de etapa temprana, para proporcionar una vacuna de múltiples fases. Alternativamente, los antígenos de etapa tardía se podrían utilizar para complementar y mejorar la vacunación con BCG (ya sea potenciando la respuesta a la BCG, o mediante el desarrollo de cepas de BCG recombinantes avanzadas).

5 Rv2386c, también conocida como MbtI, cataliza la transformación inicial en la biosíntesis de micobactina mediante la conversión de corismato a salicilato (Zwahlen et al. *Biochemistry* 2007 46:954-964). Rv2386 se ha identificado anteriormente por estar asociado con la expresión bajo condiciones de estrés, aunque no dio como resultado una protección cuando se administró como una vacuna de ADN en un modelo de cobaya (Vipond et al. *Vaccine* 2006 24:6340-6350).

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

15 Aunque se ha demostrado que los macrófagos actúan como los efectores principales de la inmunidad a *Mycobacterium*, las células T son los inductores predominantes de tal inmunidad. La función esencial de las células T en la protección contra la tuberculosis se ilustra por el aumento en los índices de la reactivación de la TB en los individuos infectados con el virus de la inmunodeficiencia humana, debido al agotamiento asociado de las células T CD4+. Además, se ha demostrado que la transferencia adoptiva de las células T CD4+ tomadas a la altura de la respuesta inmunitaria primaria a la *M. tuberculosis* confiere protección contra la *M. tuberculosis* en los ratones deficientes en células T (Orme et al, *J. Exp. Med.* 1983 158:74-83).

20 Se ha demostrado que las células T CD4+ que reaccionan con *Mycobacterium* son potentes productoras de interferón- $\gamma$  (IFN- $\gamma$ ), el cual, a su vez, se ha demostrado que desencadena los efectos anti-micobacterianos de los macrófagos en ratones (Flynn et al. *J. Exp. Med.* 1993 178:2249-2254). Aunque está menos clara la función del IFN- $\gamma$  en los seres humanos, los estudios han demostrado que la 1,25-dihidroxi-vitamina D3, ya sea en solitario o bien en combinación con el IFN- $\gamma$  o con el factor de necrosis tumoral-alfa, activa los macrófagos humanos para inhibir la infección por *M. tuberculosis*. Además, se sabe que el IFN- $\gamma$  estimula a los macrófagos humanos para producir 1,25-dihidroxi-vitamina D3. De una 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 edición, Rom y Garay, eds., 2003), y *Harrison's Principles of Internal Medicine*, Capítulo 150, páginas 953-966 (16ª edición, Braunwald et al., eds., 2005).

30 La presente invención se refiere en términos generales a la identificación de Rv2386c como un antígeno de TB (en particular un antígeno asociado con TB latente), y a los usos relacionados con la prevención y el tratamiento de TB, especialmente en la prevención y el tratamiento de TB latente y en la prevención o demora de la reactivación de TB.

35 La invención, por lo tanto, proporciona una proteína de Rv2386c, una variante de la misma, o un fragmento inmunogénico de la misma, o un polinucleótido que codifica dicha proteína, variante o fragmento, para el uso en el tratamiento o en la prevención de TB. De una manera adecuada, el uso puede ser específicamente en la prevención y el tratamiento de TB latente (especialmente el tratamiento de TB latente). Alternativamente, el uso puede ser en la prevención o el retraso de la reactivación de TB (especialmente el retraso de la reactivación de TB, por ejemplo, por un período de meses, años, o incluso indefinidamente).

40 El término "especies de *Mycobacterium* del complejo de tuberculosis" incluye las especies que tradicionalmente se consideran como causantes de la enfermedad de la tuberculosis, así como las especies ambientales y oportunistas de *Mycobacterium* que provocan la tuberculosis y la enfermedad pulmonar en los pacientes inmunodeprimidos, tales como los pacientes con SIDA, por ejemplo, *M. tuberculosis*, *M. bovis*, o *M. africanum*, BCG, *M. avium*, *M. intracellulare*, *M. celatum*, *M. genavense*, *M. haemophilum*, *M. kansasii*, *M. simiae*, *M. vaccae*, *M. fortuitum*, y *M. scrofulaceum* (véase, por ejemplo, *Harrison's Principles of Internal Medicine*, Capítulo 150, páginas 953-966 (16ª edición, Braunwald et al., eds., 2005). La presente invención se refiere en particular a la infección con *M. tuberculosis*.

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

Los términos "infección inactiva", "infección letárgica" o "infección latente" se refieren a una infección (por ejemplo, a una infección por *M. tuberculosis*) sin los síntomas y/o lesiones evidentes de la enfermedad (de una manera adecuada, sin los síntomas manifestados de la enfermedad).

55 El término "tuberculosis primaria" se refiere a la enfermedad clínica (la manifestación de los síntomas de la enfermedad) directamente después de la infección (por ejemplo, la infección por *M. tuberculosis*). Véase, *Harrison's Principles of Internal Medicine*, Capítulo 150, páginas 953-966 (16ª edición, Braunwald et al., eds, 2005).

Los términos "tuberculosis secundaria" o "tuberculosis post-primaria" se refieren a la reactivación de una infección

letárgica, inactiva o latente (por ejemplo, una infección por *M. tuberculosis*). Véase, *Harrison's Principles of Internal Medicine*, Capítulo 150, páginas 953-966 (16ª edición, Braunwald et al., eds., 2005).

El término "reactivación de tuberculosis" se refiere a la manifestación posterior de los síntomas de la enfermedad en un individuo en quien se prueba positiva la infección (por ejemplo, en una prueba cutánea de tuberculina, de una manera adecuada en un ensayo basado en células T *in vitro*) pero que no tiene los síntomas evidentes de la enfermedad. La prueba de diagnóstico positiva indica que el individuo está infectado, sin embargo, el individuo puede o no haber manifestado anteriormente los síntomas de la enfermedad activa que se han tratado lo suficiente como para llevar a la tuberculosis hasta un estado inactivo o latente. Se reconocerá que se pueden iniciar métodos para la prevención, el retraso, o el tratamiento de la reactivación de la tuberculosis en un individuo que manifieste los síntomas activos de la enfermedad.

El término tuberculosis "resistente a fármacos" se refiere a una infección (por ejemplo, una infección por *M. tuberculosis*), en donde la cepa infecciosa no se mantiene estática ni muere por (es decir, es resistente a) uno o más agentes quimioterapéuticos denominados "de primera línea" eficaces en el tratamiento de la tuberculosis (por ejemplo, isoniazida, rifampina, etambutol, estreptomycin, y pirazinamida).

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

Un "agente quimioterapéutico" se refiere a un agente farmacológico conocido y usado en la técnica para tratar la tuberculosis (por ejemplo, la infección por *M. tuberculosis*). Ejemplos de agentes farmacológicos utilizados para tratar la tuberculosis incluyen, pero no se limitan a, amikacina, ácido aminosalicílico, capreomicina, cicloserina, etambutol, etionamida, isoniazida, kanamicina, pirazinamida, rifamicinas (es decir, rifampina, rifapentina y rifabutina), estreptomycin, ofloxacin, ciprofloxacino, claritromicina, azitromicina y fluoroquinolonas. Los agentes quimioterapéuticos "de primera línea" o "de línea frontal" utilizados para tratar la tuberculosis que no es resistente a fármacos incluyen isoniazida, rifampina, etambutol, estreptomycin, y pirazinamida. Los agentes quimioterapéuticos "de segunda línea" utilizados para tratar la tuberculosis que han demostrado resistencia a fármacos, es decir, a uno o más fármacos "de primera línea" 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.

Los términos "polipéptido", "péptido" y "proteína" se utilizan indistintamente en la presente memoria para referirse a un polímero de restos de aminoácidos. Los términos también se aplican a polímeros de aminoácidos en donde uno o más restos de aminoácidos son un mimético químico artificial de un aminoácido natural correspondiente, así como a los polímeros de aminoácidos naturales y a los polímeros de aminoácidos no naturales. De una manera adecuada, un polipéptido según la presente invención consistirá solamente en restos de aminoácidos que aparecen de manera natural, especialmente los aminoácidos codificados por el código genético.

El término "aminoácido" se refiere a los aminoácidos naturales y sintéticos, así como a los análogos de aminoácidos y miméticos de aminoácidos que funcionan de una manera similar a los aminoácidos que aparecen de manera natural. Los aminoácidos que aparecen de manera natural son aquellos codificados por el código genético, así como los aminoácidos que se modifican posteriormente, por ejemplo, hidroxiprolina,  $\gamma$ -carboxiglutamato, y O-fosfoserina. Análogos de aminoácidos se refieren a los compuestos que tienen la misma estructura química básica que un aminoácido que aparece de manera natural, es decir, un carbono  $\alpha$  que se enlaza a un hidrógeno, un grupo carboxilo, un grupo amino, y un grupo R, por ejemplo, homoserina, norleucina, sulfóxido de metionina, metionina-metil-sulfonio. Dichos análogos tienen grupos R modificados (por ejemplo, norleucina) o cadenas principales de péptidos modificados, pero retienen la misma estructura química básica que un aminoácido que aparece de manera natural. Miméticos de aminoácidos se refieren a los compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que funciona de una manera similar a un aminoácido que aparece de manera natural. De una manera adecuada, un aminoácido es un aminoácido que se aparece de manera natural o un análogo de aminoácido, especialmente un aminoácido que se presenta naturalmente, y en particular los aminoácidos codificados por el código genético.

"Ácido nucleico" se refiere a desoxirribonucleótidos o ribonucleótidos y polímeros de los mismos, ya sea en forma monocatenaria o bicatenaria. El término abarca los ácidos nucleicos que contienen análogos de nucleótidos conocidos o restos o enlaces de cadena principal modificados, que son sintéticos, de origen natural, y no naturales, que tienen propiedades de enlace similares a los del ácido nucleico de referencia, y los cuales se metabolizan de una manera similar a los nucleótidos de referencia. Ejemplos de estos análogos incluyen, sin limitación, fosforotioatos, fosforamidatos, metil-fosfonatos, metil-fosfonatos quirales, ribonucleótidos de 2-O-metilo, ácidos nucleicos peptídicos (PNAs). Adecuadamente, el término "ácido nucleico" se refiere a los desoxirribonucleótidos o ribonucleótidos que aparecen de manera natural, y a los polímeros de los mismos.

A menos que se indique de otra manera, una secuencia de ácido nucleico particular también abarca implícitamente variantes de la misma modificadas de forma conservadora (por ejemplo, sustituciones de codones degenerados), y secuencias complementarias, así como la secuencia explícitamente indicada (de una manera adecuada, se refiere a

la secuencia explícitamente indicada). De una manera específica, las sustituciones de codones degenerados se puede lograr mediante la generación de secuencias en donde la tercera posición de uno o más codones seleccionados (o todos) está sustituida con restos de base mixta y/o de desoxi-inosina (Batzer et al., *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka et al., *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini et al., *Mol. Cell. Probes* 8:91-98 (1994)). El término "ácido nucleico" se utiliza de una manera intercambiable con gen, ADNc, ARNm, oligonucleótido, y polinucleótido.

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

El término "secuencia proteica de Rv2386c", como se emplea en la presente memoria, significa la secuencia polipeptídica proporcionada en SEQ ID No: 1 o un homólogo de la misma a partir de una especie de *Mycobacterium* del complejo de tuberculosis, por ejemplo, una especie tal como *M. tuberculosis*, *M. bovis*, o *M. africanum*, o una especie de *Mycobacterium* que es ambiental u oportunista, y que causa infecciones oportunistas, tales como infecciones pulmonares en los 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, páginas 953-966, 16ª edición, Braunwald et al., eds, 2005).

Para asegurar un alto índice de eficacia entre los huéspedes vacunados, los componentes de una vacuna deben estar bien conservados entre las cepas de significado clínico. De una manera adecuada, la proteína Rv2386c se deriva a partir de *M. tuberculosis* H37Rv (es decir, la secuencia polipeptídica proporcionada en SEQ ID No: 1) o un homólogo de la misma a partir de otra cepa de *M. tuberculosis* (tal como CDC1551, F11, Haarlem cepas A y C). Las cepas de *M. tuberculosis* que están asociadas con la resistencia a fármacos (por ejemplo, MDR o especialmente XDR), son una base particularmente valiosa para la secuencia proteica de Rv2386c. Las cepas de interés incluyen:

CDC1551- cepa transmisible y virulenta

Familia Haarlem (tal como Haarlem A) - Cepas resistentes a fármacos encontradas en poblaciones humanas saturadas. Se han encontrado miembros de la familia Haarlem de las cepas de *M. tuberculosis* en muchas partes del mundo. El primer representante de la familia se descubrió en Haarlem, Holanda.

KZN4207- Aislado sensible a fármacos a partir de pacientes de KwaZulu-Natal, Sudáfrica.

KZN1435- Aislado resistente a múltiples fármacos (MDR) a partir de pacientes de KwaZulu-Natal, Sudáfrica.

KZN605- Aislado extensamente resistente a fármacos (XDR) a partir de pacientes de KwaZulu-Natal, Sudáfrica.

C - Altamente transmitida en la Ciudad de Nueva York. En un estudio, se encontró que esta cepa es más común entre los usuarios de fármacos inyectables y resistente a los intermediarios de nitrógeno reactivo (Friedman et al., *J. Infect. Dis.* 1997 176(2):478-84).

94\_M4241A - Aislado en San Francisco en 1994 a partir de un paciente nacido en China. Esta cepa se analizó anteriormente mediante el análisis de delección genómica (Gagneux et al., *PNAS* 2006 103(8):2869-2873).

02\_1987- Aislado en San Francisco en 2002 a partir de un paciente nacido en Corea del Sur. Esta cepa se analizó anteriormente mediante el análisis de delección genómica (Gagneux et al., *PNAS* 2006 103(8):2869-2873).

T92 - Aislado en San Francisco en 1999 a partir de un paciente nacido en Las Filipinas. Esta cepa se publicó en Hirsh et al., *PNAS* 2004 101:4871-4876).

T85 - Aislado en San Francisco en 1998 a partir de un paciente nacido en China. Esta cepa se publicó en Hirsh et al., *PNAS* 2004 101:4871-4876).

EAS054 - Aislado en San Francisco en 1993 a partir de un paciente nacido en India. Esta cepa se analizó anteriormente mediante el análisis de delección genómica (Gagneux et al., *PNAS* 2006 103(8):2869-2873).

Gagneux et al., *PNAS* 2006 103(8):2869-2873 y Herbert et al. *Infect. Immun.* 2007 75(12):5798-5805 proporcionan valiosos antecedentes sobre el rango de cepas de *M. tuberculosis* que se sabe que existen.

De una manera más adecuada, la proteína Rv2386c se selecciona a partir de las secuencias polipeptídicas proporcionadas en SEQ ID Nos: 1 y 3-7, en particular en SEQ ID Nos: 1 y 3-6, tal como SEQ ID NO: 1.

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

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma;  
o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c.

5 Los polinucleótidos comprenderán adecuadamente (tal como consistirán en) una variante de SEQ ID NO: 2 o un fragmento de SEQ ID NO: 2 que codifica un fragmento inmunogénico de una proteína Rv2386c.

### Combinaciones

10 Los polipéptidos relacionados con Rv2386c de la presente invención pueden comprender además otros componentes diseñados para mejorar su inmunogenicidad o para mejorar estos antígenos en otros aspectos. Por ejemplo, se puede facilitar un mejor aislamiento de los antígenos de polipéptido a través de la adición de un tramo de restos de histidina (comúnmente conocido como etiqueta de histidina) hacia un extremo del antígeno.

15 El término "etiqueta de histidina" se refiere a una cadena de restos de histidina, típicamente de seis restos, que se insertan dentro de la secuencia de referencia. Para minimizar la alteración de la actividad asociada con la secuencia de referencia, normalmente se inserta una etiqueta de histidina en el N terminal, usualmente inmediatamente después del resto de metionina de inicio, o bien en el C terminal. Usualmente son heterólogos para la secuencia nativa, pero se incorporan debido a que facilitan el aislamiento mejorando la unión de la proteína a las resinas de cromatografía de afinidad de metal inmovilizado (IMAC). Generalmente, la presencia o ausencia de una etiqueta de histidina no es significativa desde el punto de vista de provocar una respuesta inmunitaria deseable contra la proteína de referencia. Sin embargo, para evitar el riesgo de una reacción adversa contra la propia etiqueta de histidina, se considera mejor minimizar la longitud de la etiqueta de histidina, por ejemplo, hasta cuatro o menos restos, en particular hasta dos restos (o excluir el uso de una etiqueta de histidina totalmente).

20 Para mejorar la magnitud y/o el alcance de la respuesta inmunitaria provocada, las composiciones, polipéptidos y ácidos nucleicos para emplear en la invención pueden comprender múltiples copias del antígeno y/o polipéptidos heterólogos adicionales (o polinucleótidos que los codifiquen) a partir de las especies de *Mycobacterium* (en particular *M. tuberculosis*).

Un experto en este campo reconocerá que, cuando se utilizan varios componentes en combinación, se puede variar la presentación precisa. Por ejemplo, se podría presentar un componente de Rv2386c y una copia adicional del antígeno o un componente de antígeno heterólogo adicional:

- (1) como dos componentes polipeptídicos individuales;
- 30 (2) como una proteína de fusión que comprende ambos componentes polipeptídicos;
- (3) como un polipéptido y un componente de polinucleótido;
- (4) como dos componentes de polinucleótido individuales;
- (5) como un único polinucleótido que codifica dos componentes polipeptídicos individuales; o
- 35 (6) como un único polinucleótido que codifica una proteína de fusión que comprende ambos componentes polipeptídicos.

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

45 El término "heterólogo", cuando se utiliza con referencia a partes 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 se produce normalmente de una manera recombinante, teniendo dos o más secuencias a partir de genes no relacionados dispuestos para formar un nuevo ácido nucleico funcional, por ejemplo, un promotor a partir de una fuente, y una región codificante a partir de otra fuente. De una manera similar, una proteína heteróloga indica que la proteína comprende dos o más subsecuencias que no se encuentran en la misma relación entre sí en la naturaleza (por ejemplo, una proteína de fusión).

50 "Polipéptido de fusión" o "proteína de fusión" se refiere a una proteína que tiene al menos dos polipéptidos heterólogos (por ejemplo, al menos dos polipéptidos de *Mycobacterium* sp.) unidos covalentemente, ya sea directamente o bien a través de un enlazador de aminoácidos. Los polipéptidos que forman la proteína de fusión

- típicamente enlazan el C-terminal al N-terminal, aunque también pueden enlazar el C-terminal al C-terminal, el N-terminal al N-terminal, o el N-terminal al C-terminal. Los polipéptidos de la proteína de fusión pueden estar en cualquier orden. Este término también se refiere a las variantes modificadas conservadoramente, variantes polimórficas, alelos, mutantes, fragmentos inmunogénicos, y homólogos interespecies de los antígenos que forman la proteína de fusión. Los antígenos de *Mycobacterium tuberculosis* se describen en Coler et al., *Nature* 393:537 (1998), que describe el genoma entero de *Mycobacterium tuberculosis*. Los antígenos a partir de otras especies de *Mycobacterium* que corresponden a antígenos de *M. tuberculosis* se pueden identificar, por ejemplo, utilizando algoritmos de comparación de secuencias, como se describen en la presente memoria, u otros métodos conocidos por los expertos en la técnica, por ejemplo, ensayos de hibridación y ensayos de unión a anticuerpos.
- El término “fusionado” se refiere al enlace covalente entre dos polipéptidos en una proteína de fusión. Los polipéptidos típicamente se unen a través de un enlace peptídico, ya sea directamente entre sí o bien 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.
- Ejemplos de antígenos de *M. tuberculosis* que se pueden combinar con Rv2386c incluyen uno o más de (por ejemplo, de 1 a 5, tal como de 1 a 3, en particular 1) los siguientes (tales como uno o más de (i) a (xii)):
- (i) Mtb8.4 (también conocido como DPV y Rv1174c), cuya secuencia polipeptídica se describe en SEQ ID No: 102 de WO97/09428 (ADNc en SEQ ID No: 101), y en Coler et al, *Journal of Immunology* 1998 161:2356-2364. Es de un interés particular la secuencia de Mtb8.4 madura, que está ausente del péptido señal líder (es decir, los restos aminoácidos 15-96 de SEQ ID No: 102 de WO97/09428). La secuencia del polipéptido de longitud completa de Mtb8.4 se muestra en SEQ ID No: 8;
- (ii) Mtb9.8 (también conocido como MSL y Rv0287), cuya secuencia polipeptídica se describe en SEQ ID No: 109 de WO98/53075 (los fragmentos de MSL se describen en SEQ ID NOs: 110-124 de WO98/53075, siendo de un interés particular SEQ ID NOs: 119 y 120), y también en Coler et al, *Vaccine* 2009 27:223-233 (en particular los fragmentos reactivos mostrados en la Figura 2 del mismo). La secuencia polipeptídica de longitud completa para Mtb9.8 se muestra en SEQ ID No: 9;
- (iii) Mtb9.9 (también conocido como Mtb9.9A, MTI, MTI-A y Rv1793), cuya secuencia polipeptídica se describe en SEQ ID No: 19 WO98/53075 y en Alderson et al, *Journal of Experimental Medicine* 2000 7:551-559 (los fragmentos de MTI se describen en SEQ ID NOs: 17 y 51-66 de WO98/53075, siendo de un interés particular SEQ ID NOs: 17, 51, 52, 53, 56 y 62-65). Se describen varias variantes polipeptídicas de MTI en SEQ ID NOs: 21, 23, 25, 27, 29 y 31 de WO98/53075 y en Alderson et al, *Journal of Experimental Medicine* 2000 7:551-559. La secuencia polipeptídica de longitud completa para Mtb9.9 se muestra en SEQ ID No: 10;
- (iv) Ra12 (también conocido como el antígeno C-terminal Mtb32A), cuya secuencia polipeptídica se describe en SEQ ID No: 10 de WO01/98460 y en Skeiky et al, *Journal of Immunology* 2004 172:7618-7682. La secuencia polipeptídica de longitud completa para Ra12 se muestra en SEQ ID No: 11;
- (v) Ra35 (también conocido como el antígeno N-terminal Mtb32A), cuya secuencia polipeptídica se describe en SEQ ID No: 8 de WO01/98460 y en Skeiky et al, *Journal of Immunology* 2004 172:7618-7682. La secuencia polipeptídica de longitud completa para Ra35 se muestra en SEQ ID No: 12;
- (vi) TbH9 (también conocido como Mtb39, Mtb39A, TbH9FL y Rv1196), cuya secuencia polipeptídica se describe en SEQ ID No: 107 de WO97/09428, y también en Dillon et al, *Infection and Immunity* 1999 67(6):2941-2950, y en Skeiky et al, *Journal of Immunology* 2004 172:7618-7682. La secuencia polipeptídica de longitud completa para TbH9 se muestra en SEQ ID No: 13;
- (vii) Mtb40 (también conocido como HTCC1 y Rv3616c), cuya secuencia polipeptídica se describe en SEQ ID No: 138 de WO98/53075 (ADNc en SEQ ID No: 137). La secuencia polipeptídica de longitud completa para Mtb40 se muestra en SEQ ID No: 14;
- (viii) Mtb41 (también conocido como MTCC2 y Rv0915c), cuya secuencia polipeptídica se describe en SEQ ID No: 142 de WO98/53075 (ADNc en SEQ ID No: 140), y en Skeiky et al, *Journal of Immunology* 2000 165:7140-7149. La secuencia polipeptídica de longitud completa para Mtb41 se muestra en SEQ ID No: 15;
- (ix) ESAT-6 (también conocido como esxA y Rv3875), cuya secuencia polipeptídica se describe en SEQ ID No: 103 de WO97/09428 (ADNc en SEQ ID No: 104), y en Sorensen et al, *Infection and Immunity* 1995 63(5):1710-1717. La secuencia polipeptídica de longitud completa para ESAT-6 se muestra en SEQ ID No: 16;
- (x) Antígenos complejos Ag85 (por ejemplo, Ag85A, también conocido como fbpA y Rv3804c; o Ag85B, también conocido como fbpB y Rv1886c), los cuales se discuten, por ejemplo, en Content et al, *Infection and Immunity* 1991 59:3205-3212 y en Huygen et al, *Nature Medicine* 1996 2(8):893-898. La secuencia polipeptídica de longitud completa para Ag85A se muestra en SEQ ID No: 17 (la proteína madura de 43-338 restos, es decir, que carece del péptido señal, es de particular interés). La secuencia polipeptídica de longitud completa para Ag85B se muestra en SEQ ID No: 18 (la proteína madura de 41-325 restos, es decir, que carece del péptido señal, es de

particular interés);

- (xi) Alfa-cristalina (también conocido como hspX y Rv2031c), que se describe en Verbon et al, *Journal of Bacteriology* 1992 174:1352-1359, y en Friscia et al, *Clinical and Experimental Immunology* 1995 102:53-57 (son de particular interés los fragmentos correspondientes 71-91, 21-40, 91-110 y 111-130 restos). La secuencia polipeptídica de longitud completa para alfa-cristalina se muestra en SEQ ID No: 19;
- (xii) Mpt64 (también conocido como Rv1980c), que se describe en Roche et al, *Scandinavian Journal of Immunology* 1996 43:662-670. La secuencia polipeptídica de longitud completa para MPT64 se muestra en SEQ ID No: 20 (la proteína madura de 24-228 restos, es decir, que carece del péptido señal, es de particular interés);
- (xiii) Mtb32A, cuya secuencia polipeptídica se describe en SEQ ID No: 2 (longitud completa), y 8-330 restos de SEQ ID NO: 4 (madura) de WO01/98460, especialmente las variantes que tienen al menos una de las tríadas catalíticas mutadas (por ejemplo, el resto catalítico de serina, el cual, por ejemplo, puede mutar a alanina). La secuencia polipeptídica de longitud completa para Mtb32A se muestra en SEQ ID No: 21. La forma madura de Mtb32A que tiene una mutación de Ser/Ala se muestra en SEQ ID No: 22;
- (xiv) TB10.4, la secuencia polipeptídica de longitud completa para TB10.4 se muestra en SEQ ID No: 23
- (xv) Rv1753c, la secuencia polipeptídica de longitud completa para Rv1753c a partir de *Mycobacterium tuberculosis* H37Rv se muestra en SEQ ID No: 155; y/o
- (xvi) Rv2707c, la secuencia polipeptídica de longitud completa para Rv2707c a partir de *Mycobacterium tuberculosis* H37Rv se muestra en SEQ ID No: 156.
- o combinaciones de los mismos, tales como (por ejemplo, combinaciones tales como (a) a (g)):
- (a) una combinación de los componentes Ra12, TbH9 y Ra35, por ejemplo, en la forma de una proteína de fusión, tal como Mtb72f. La secuencia polipeptídica de Mtb72f se describe en SEQ ID No: 6 de WO2006/117240 (ADNc en SEQ ID No: 5), y en Skeiky et al, *Journal of Immunology* 2004 172:7618-7682 (en donde incorpora una etiqueta de histidina opcional para ayudar a la purificación, cuando se utiliza en la presente invención de una manera adecuada, Mtb72f está ausente de los restos opcionales de histidina). La secuencia polipeptídica para Mtb72f se muestra en SEQ ID No: 24;
- (b) una combinación de los componentes Ra12, TbH9 y Ra35 mutado con Ser/Ala (es decir, en donde el resto catalítico de serina ha sido reemplazado con alanina), por ejemplo en forma de una proteína de fusión, tal como M72. La secuencia polipeptídica de M72 se describe en SEQ ID No: 4 de WO2006/117240 (ADNc en SEQ ID No: 3), en 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 en M72 está ausente la doble histidina opcional (es decir, 4-725 restos de SEQ ID No: 4 de WO2006/117240 son de particular interés). La secuencia polipeptídica para M72 se muestra en SEQ ID No: 25;
- (c) una combinación de 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 SEQ ID No: 16 de WO99/051748 (ADNc en SEQ ID No: 15), en donde incorpora una etiqueta de histidina opcional para ayudar a la purificación, cuando se utiliza en la presente invención de una manera adecuada, Mtb71f corresponde a 9-710 restos aminoácidos de SEQ ID NO: 16 de WO99/051748. La secuencia polipeptídica para Mtb71f se muestra en SEQ ID No: 26;
- (d) una combinación de Mtb72f o M72 (de una manera adecuada sin restos opcionales de histidina para ayudar a la expresión) con Mtb9.8 y Mtb9.9, por ejemplo, en una proteína de fusión. La secuencia polipeptídica para una fusión de M72-Mtb9.9-Mtb9.8 se muestra en SEQ ID No: 27 (fusión de M92); cuando se utiliza en la presente invención, la fusión de M72-Mtb9.9-Mtb9.8 puede incorporar opcionalmente una doble histidina después del resto de metionina de inicio para ayudar a la fabricación;
- (e) una combinación de Mtb72f o M72 (de una manera adecuada sin restos opcionales de histidina 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 SEQ ID No: 18 de WO03/070187 (ADNc en SEQ ID No: 10), en donde incorpora una etiqueta de histidina opcional para ayudar a la purificación, cuando se utiliza en la presente invención, de una manera adecuada, Mtb103f corresponde a 8-1016 restos aminoácidos de SEQ ID NO: 18 de WO03/070187. También es de particular interés M103, es decir, Mtb103f que incorpora una mutación de Ser/Ala en el componente Ra35, cuando se utiliza en la presente invención, de una manera adecuada, M103 corresponde a 8-1016 restos aminoácidos de SEQ ID NO: 18 de WO03/070187, en donde el resto Ser en la posición 710 ha sido reemplazado por Ala. La secuencia polipeptídica para M103 se muestra en SEQ ID No: 28, cuando se utiliza en la presente invención, la fusión de M72-Mtb9.9-Mtb9.8 puede incorporar opcionalmente una doble histidina después del resto de metionina de inicio para ayudar a la fabricación;
- (f) una combinación de Mtb72f o M72 (de una manera adecuada sin restos opcionales de histidina 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 SEQ ID No: 16 de WO03/070187 (ADNc en SEQ ID No: 9), en donde incorpora una etiqueta de histidina opcional para ayudar a la purificación, cuando se utiliza en la presente invención, de una manera adecuada, Mtb114f corresponde a los restos aminoácidos 8-1154 de SEQ ID NO: 16 de WO03/070187. También es de particular interés M114, es decir, Mtb114f que incorpora una mutación de Ser/Ala en el componente Ra35; cuando se utiliza en la presente invención, de una manera adecuada, M114 corresponde a 8-1154 restos aminoácidos de SEQ ID NO: 16 a partir de WO03/070187- También es de particular interés M114, es decir, Mtb114f que incorpora una mutación de Ser/Ala en el componente Ra35, cuando se utiliza en la presente invención de manera adecuada M114 corresponde a 8-1154 restos aminoácidos de SEQ ID NO: 16 a partir de WO03/070187 en donde el resto Ser en posición 710 ha sido reemplazado por Ala. La secuencia polipeptídica para M114 se muestra en SEQ ID No: 29, cuando se utiliza en la presente invención, la fusión de M72-Mtb9.9-Mtb9.8 puede incorporar opcionalmente una doble histidina después del resto de metionina de inicio para ayudar a la fabricación;

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

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

Las combinaciones de un componente Rv2386c y un componente Mtb40 son de un interés particular. Obviamente, tales combinaciones podrían contener opcionalmente otros componentes de antígeno adicionales (por ejemplo, un componente de M72).

Otra combinación de interés comprende un componente Rv2386c y un componente M72.

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

Otras combinaciones de interés incluyen aquellas que comprenden un componente Rv2386c y un componente Rv2707c.

Una combinación adicional de interés comprende un componente Rv2386c y un componente alfa-cristalina.

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

Cada una de las secuencias de antígenos individuales anteriores se describen también en Cole et al, *Nature* 1998 393:537-544 y en 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/](http://www.sanger.ac.uk/Projects/M_tuberculosis/)), y en otros lugares.

Muchos de los antígenos anteriores se describen también en las Solicitudes de Patente U.S. Números 08/523.435, 08/523.436, 08/658.800, 08/659.683, 08/818.111, 08/818.112, 08/942.341, 08/942.578, 08/858.998, 08/859.381, 09/056.556, 09/072.596, 09/072.967, 09/073.009, 09/073.010, 09/223.040, 09/287.849, en las Solicitudes de Patente PCT PCT/US98/10407, PCT/US98/10514, PCT/US99/03265, PCT/US99/03268, PCT/US99/07717, WO97/09428 y WO97/09429, WO98/16645, WO98/16646, cada una de las cuales se incorpora en la presente memoria como referencia.

Las composiciones, polipéptidos, y ácidos nucleicos para el uso en la invención también pueden comprender polipéptidos adicionales a partir de otras fuentes. Por ejemplo, las composiciones y las proteínas de fusión para el uso de la invención pueden incluir polipéptidos o ácidos nucleicos que codifican polipéptidos, en donde el polipéptido mejora la expresión del antígeno, por ejemplo, NS1, una proteína del virus de influenza (véanse, por ejemplo, WO99/40188 y WO93/04175). Los ácidos nucleicos para el uso en la invención se pueden diseñar basándose en la preferencia de codones en una especie de elección, por ejemplo, en seres humanos (en el caso de la expresión *in vivo*) o en una bacteria particular (en el caso de la producción de polipéptidos).

El componente Rv2386c también se puede administrar con uno o más agentes quimioterapéuticos eficaces contra la tuberculosis (por ejemplo, contra la infección por *M. tuberculosis*). Ejemplos de tales agentes quimioterapéuticos incluyen, pero no se limitan a, amikacina, ácido aminosalicílico, capreomicina, cicloserina, etambutol, etionamida, isoniazida, kanamicina, pirazinamida, rifamicinas (es decir, rifampina, rifapentina y rifabutina), estreptomina, ofloxacino, ciprofloxacino, claritromicina, azitromicina y fluoroquinolonas. Tal quimioterapia se determina por el juicio del médico tratante utilizando combinaciones de fármacos preferidos. Agentes quimioterapéuticos de "primera línea" utilizados para tratar la tuberculosis (por ejemplo, la infección por *M. tuberculosis*) que no son resistentes a fármacos, incluyen isoniazida, rifampina, etambutol, estreptomina y pirazinamida. Agentes quimioterapéuticos de "segunda línea" utilizados para tratar la tuberculosis (por ejemplo, la infección por *M. tuberculosis*) que han

demostrado resistencia a uno o más fármacos de “primera línea”, incluyen ofloxacino, ciprofloxacino, etionamida, ácido aminosalicílico, cicloserina, amikacina, kanamicina y capreomicina.

5 Los agentes quimioterapéuticos convencionales se administran en general durante un período relativamente largo (aproximadamente 9 meses). La combinación de los agentes quimioterapéuticos convencionales con la administración de un componente Rv2386c según la presente invención, puede permitir que se reduzca el período del tratamiento quimioterapéutico (por ejemplo, a 8 meses, 7 meses, 6 meses, 5 meses, 4 meses, 3 meses o menos) sin una disminución en la eficacia.

10 Es de un interés particular el uso de un componente Rv2386c en conjunto con el *Bacillus Calmette-Guerin* (BCG). Por ejemplo, en forma de un BCG modificado que exprese de una manera recombinante Rv2386c (o una variante o un fragmento del mismo, como se describe en la presente memoria). Alternativamente, el componente Rv2386c se puede utilizar para mejorar la respuesta de un sujeto a la vacunación con BCG, ya sea mediante administración conjunta o potenciando una vacunación previa con BCG. Cuando se utiliza para mejorar la respuesta de un sujeto a la vacunación con BCG, el componente Rv2386c se puede proporcionar obviamente en forma de un polipéptido o de un polinucleótido (opcionalmente junto con componentes antigénicos adicionales, como se describe anteriormente).

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

20 Los polipéptidos, polinucleótidos y composiciones para el uso en la presente invención usualmente se administrarán 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**

25 Los epítomos de células T son tramos contiguos cortos de aminoácidos que son reconocidos por células T (por ejemplo, células T CD4+ o CD8+). La identificación de epítomos de células T se puede lograr a través de experimentos de mapeo de epítomos que son bien conocidos por el experto en la técnica (véase, por ejemplo, Paul, *Fundamental Immunology*, 3ª edición, 243-247 (1993); Beißbarth et al, *Bioinformatics* 2005 21(Suplemento 1):i29-i37).

30 Alternativamente, los epítomos se pueden predecir empleando los planteamientos discutidos en los Ejemplos.

Como resultado de la involucración crucial de la respuesta de células T en la tuberculosis, es fácilmente evidente que los fragmentos del polipéptido Rv2386c de longitud completa que contienen al menos un epítomo de células T serán contribuir a la inmunoprotección. Tales fragmentos son referidos en la presente memoria como fragmentos inmunogénicos.

35 Los fragmentos inmunogénicos para el uso según la presente invención típicamente comprenderán al menos 9 aminoácidos contiguos a partir de la secuencia polipeptídica de longitud completa (por ejemplo, de al menos 10), tal 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 200 aminoácidos contiguos). De una manera adecuada, los fragmentos inmunogénicos serán al menos el 20%, tal como al menos 50%, al menos 70%, o al menos 80% de la longitud de la secuencia polipeptídica de longitud completa.

40 Se entenderá que en una población diversa, tal como en los seres humanos, diferentes tipos de HLA significan que los epítomos específicos pueden no ser reconocidos por todos los miembros de la población. En consecuencia, para maximizar el nivel de reconocimiento y la escala de respuesta inmunitaria a un polipéptido, generalmente es deseable que un fragmento inmunogénico contenga una pluralidad de epítomos a partir de la secuencia de longitud completa (de una manera adecuada, todos los epítomos).

45 Los fragmentos particulares de la proteína Rv2386c que pueden ser útiles incluyen aquellos que contienen al menos un epítomo CD4+, de una manera adecuada al menos dos epítomos CD4+, y especialmente todos los epítomos CD4+ (tales como los epítomos descritos en los Ejemplos y en SEQ ID NOs: 30-52, en particular aquellos asociados con una pluralidad de alelos HLA, por ejemplo, aquellos asociados con 2, 3, 4, 5 o más alelos).

50 Otros fragmentos de la proteína Rv2386c que pueden ser útiles incluyen aquellos que contiene al menos un epítomo CD8, de una manera adecuada al menos dos epítomos CD8, y especialmente todos los epítomos CD8 (tales como los epítomos descritos en los Ejemplos y en SEQ ID NOs: 53-154, en particular aquellos asociados con una pluralidad de alelos HLA, por ejemplo, aquellos asociados con 2, 3, 4, 5 o más alelos).

55 Cuando se utiliza un fragmento individual del polipéptido de longitud completa, se considera que tal fragmento es inmunogénico cuando provoca una respuesta que es al menos el 20%, de una manera adecuada al menos 50%, y

especialmente al menos 75% (tal como al menos 90%) de la actividad de la secuencia de referencia en un ensayo de re-estimulación *in vitro* de células mononucleares de sangre periférica (PBMC) o de sangre entera con antígenos específicos (por ejemplo, una re-estimulación durante un período 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 por medio de linfoproliferación, producción de citoquinas en el sobrenadante de cultivo (medido mediante ELISA, CBA, etc.) o caracterización de respuestas de células T y B mediante tinción intra y extracelular (por ejemplo, utilizando anticuerpos específicos para inmuno-marcadores, tales como CD3, CD4, CD8, IL2, TNFa, IFNg, CD40L, CD69, etc.), seguido por el análisis con un citómetro de flujo. De una manera adecuada, se considera que un fragmento es inmunogénico cuando provoca una respuesta que es al menos 20%, de una manera adecuada al menos 50%, y especialmente al menos 75% (tal como al menos 90%) de la actividad de la secuencia de referencia en un ensayo de proliferación de células T y/o de producción de IFN-gamma.

En algunas circunstancias se puede utilizar una pluralidad de fragmentos del polipéptido de longitud completa (que pueden superponerse o no, y pueden o no cubrir la totalidad de la secuencia de longitud completa) para obtener una respuesta biológica equivalente a la secuencia de longitud completa en sí misma. Por ejemplo, al menos dos fragmentos inmunogénicos (tal como tres, cuatro, o cinco) como se describe anteriormente, que en combinación proporcionen al menos 50%, de una manera adecuada al menos 75%, y especialmente al menos 90% de la actividad de la secuencia de referencia en un ensayo de re-estimulación *in vitro* de PBMC o de sangre entera (por ejemplo, un ensayo de proliferación de células T y/o de producción de IFN-gamma).

### Variantes

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

Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier proteína dada. Por ejemplo, los codones GCA, GCC, GCG y GCU codifican todos el aminoácido alanina. Por tanto, en cada posición en la que un codón especifica a alanina, el codón se puede alterar hasta cualquiera de los codones correspondientes descritos, sin alterar el polipéptido codificado. Tales variaciones de ácidos nucleicos conducen a variantes "silenciosas" o "degeneradas", que son una especie de variaciones modificadas conservadoramente. Cada secuencia de ácido nucleico de la presente memoria que codifica un polipéptido también describe cada variación silenciosa posible del ácido nucleico. Un experto reconocerá que cada codón en un ácido nucleico (excepto AUG, el cual es ordinariamente el único codón para metionina, y TGG, el cual es ordinariamente el único codón para triptófano) se puede modificar para proporcionar una molécula idéntica funcionalmente. Por consiguiente, cada variación silenciosa de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita.

Un polinucleótido para el uso en la invención puede contener un número de variaciones silenciosas (por ejemplo, se pueden alterar 1-50, tal como 1-25, en particular 1-5, y especialmente 1 codón(es)) cuando se compara con la secuencia de referencia. Un polinucleótido para el uso en la invención puede contener un número de variaciones conservadoras no silenciosas (por ejemplo, se pueden alterar 1-50, tal como 1-25, en particular 1-5, y especialmente 1 codón(es)) cuando se compara con la secuencia de referencia. Las variaciones no silenciosas son aquellas que dan como resultado un cambio en la secuencia de aminoácidos codificada (ya sea a través de la sustitución, delección, o adición de restos de aminoácidos). Los expertos en la técnica reconocerán que una secuencia de polinucleótido particular puede contener variaciones conservadoras tanto silenciosas como no silenciosas.

Con respecto a las variantes de una secuencia proteica, la persona experta reconocerá que las sustituciones, delecciones, o adiciones individuales al polipéptido, que altera, agrega, o elimina un único aminoácido o un pequeño porcentaje de aminoácidos es una "variante modificada conservadoramente", en donde la alteración o alteraciones dan como resultado la sustitución de un aminoácido por un aminoácido similar funcionalmente, o la sustitución/delección /adición de restos que no impactan sustancialmente la función biológica de la variante.

Las tablas de sustitución conservadora, que proporcionan aminoácidos similares funcionalmente, son bien conocidas en la técnica. Tales variantes modificadas conservadoramente son en adición a, y no excluyen, variantes polimórficas, homólogos inter-especies, y alelos de la invención.

Un polipéptido para el uso en la invención puede contener un número de sustituciones conservadoras (por ejemplo, se pueden alterar 1-50, tal como 1-25, en particular 1-10, y especialmente 1 resto (s) de aminoácidos) cuando se compara con la secuencia de referencia. En general, tales sustituciones conservadoras caerán dentro de una de las agrupaciones de aminoácidos especificadas más adelante, aunque en algunas circunstancias son posibles otras sustituciones sin afectar sustancialmente las propiedades inmunogénicas del antígeno. Los siguientes ocho grupos contienen cada uno aminoácidos que son típicamente sustituciones conservadoras entre sí:

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

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

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

15 Las variantes de proteína también pueden incluir aquellas en donde se insertan aminoácidos adicionales comparándose con la secuencia de referencia, por ejemplo, tales inserciones pueden presentarse en 1-10 localizaciones (tal como 1-5 localizaciones, de una manera adecuada 1 ó 2 localizaciones, en particular 1 localización) y, por ejemplo, pueden implicar la adición de 50 aminoácidos o menos en cada localización (tal como 20 o menos, en particular 10 o menos, especialmente 5 o menos). De una manera adecuada, estas inserciones no aparecen 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 un tramo corto de restos de histidina (por ejemplo, 2-6 restos) para ayudar a la expresión y/o purificación del antígeno en cuestión.

20 Las variantes de proteína incluyen aquellas en donde se han suprimido aminoácidos, comparándose con la secuencia de referencia, por ejemplo, estas deleciones pueden presentarse en 1-10 localizaciones (tal como 1-5 localizaciones, de una manera adecuada en 1 ó 2 localizaciones, en particular en 1 localización) y pueden, por ejemplo, implicar la supresión de 50 aminoácidos o menos en cada localización (tal como 20 o menos, en particular 10 o menos, especialmente 5 o menos). De una manera adecuada, estas supresiones no aparecen en la región de un epítipo y, por tanto, no tienen un impacto significativo sobre las propiedades inmunogénicas del antígeno.

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

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

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

35 Los términos "idéntico" o porcentaje de "identidad", en el contexto de dos o más ácidos nucleicos o secuencias polipeptídicas, se refieren a dos o más secuencias o sub-secuencias que son iguales o tienen un porcentaje especificado de restos de aminoácidos o nucleótidos que son iguales (es decir, 90%, 95%, 98% o 99% de identidad sobre una región especificada), al compararse y alinearse para una máxima correspondencia en una ventana de comparación, o una región designada medida utilizando uno de los siguientes algoritmos de comparación de secuencias, o mediante alineación manual e inspección visual. Entonces se dice que tales secuencias son "sustancialmente idénticas". Esta definición también se refiere al complemento de una secuencia de prueba. Opcionalmente, la identidad existe sobre una región que tiene al menos aproximadamente de 25 a aproximadamente 50 aminoácidos o nucleótidos de longitud, u opcionalmente sobre una región que tiene 75-100 aminoácidos o nucleótidos de longitud. De una manera adecuada, la comparación se lleva a cabo sobre una ventana correspondiente a toda la longitud de la secuencia de referencia.

45 Para la comparación de secuencias, típicamente una secuencia actúa como una secuencia de referencia, con la cual se comparan las secuencias de prueba. Cuando se utiliza un algoritmo de comparación de secuencias, las secuencias de prueba y de referencia se introducen en una computadora, se designan las coordenadas de la sub-secuencia, si es necesario, y se designan los parámetros del programa del algoritmo de secuencias. Se pueden utilizar parámetros del programa predeterminados, o se pueden designar parámetros alternativos. Entonces el algoritmo de comparación de secuencias calcula el porcentaje de identidad de secuencias para las secuencias de prueba en relación con la secuencia de referencia, basándose en los parámetros del programa.

50 Una "ventana de comparación", como se emplea en la presente memoria, hace referencia a un segmento en el que se puede comparar una secuencia con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias estén alineadas óptimamente. Los métodos de alineación de secuencias para comparación son bien conocidos en la técnica. 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 el método de búsqueda de similitudes de Pearson y Lipman, *Proc. Nat'l. Acad. Sci. EE.UU.* 85:2444 (1988), mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante alineación manual e inspección visual (véase, por ejemplo, *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., suplemento 1995)).

Un ejemplo de un algoritmo útil es el PILEUP. PILEUP crea una alineación de secuencia múltiple a partir de un grupo de secuencias relacionadas, utilizando alineaciones progresivas por pares, para mostrar la relación y el porcentaje de identidad de secuencias. También traza un árbol o dendograma que muestra las relaciones de agrupación utilizadas para crear la alineación. PILEUP utiliza una simplificación del método de alineación progresiva de Feng y Doolittle, *J. Mol. Evol.* 35:351-360 (1987). El método empleado es similar al método descrito por Higgins y Sharp, *CABIOS* 5:151-153 (1989). El programa puede alinear hasta 300 secuencias, cada una de una longitud máxima de 5.000 nucleótidos o aminoácidos. El procedimiento de alineación múltiple empieza con la alineación por pares de las dos secuencias más similares, produciendo un grupo de dos secuencias alineadas. Este grupo se alinea entonces con la siguiente secuencia más relacionada o grupo de secuencias alineadas. Dos grupos de secuencias se alinean mediante una extensión simple de la alineación por pares de dos secuencias individuales. La alineación final se logra mediante una serie de alineaciones progresivas por pares. El programa se ejecuta mediante la designación de secuencias específicas y sus coordenadas de aminoácidos o nucleótidos para las regiones de comparación de secuencias, y mediante la designación de los parámetros del programa. Utilizando PILEUP, se compara una secuencia de referencia con otras secuencias de prueba para determinar la relación del porcentaje de identidad de secuencias, utilizando los siguientes parámetros: peso de hueco predeterminado (3,00), peso de longitud de hueco predeterminado (0,10), y huecos de extremo ponderados. PILEUP se puede obtener del paquete del programa informático de análisis de secuencias GCG, por ejemplo, versión 7.0 (Devereaux *et al.*, *Nuc. Acids Res.* 12:387-395 (1984)).

Otro ejemplo de algoritmo que es adecuado para determinar el porcentaje de identidad de secuencias y de similitud de secuencias, son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.*, *Nuc. Acids Res.* 25:3389-3402 (1977), y Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990), respectivamente. El programa informático para llevar a cabo los análisis BLAST está disponible públicamente a través del National Center for Biotechnology Information (sitio web en [www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)). Este algoritmo implica identificar primero los pares de secuencias de alta puntuación (HSPs), mediante la identificación de palabras cortas de una longitud *W* en la secuencia requerida, que concuerdan o satisfacen alguna puntuación umbral de valor positivo *T* al alinearse con una palabra de la misma longitud en una secuencia de base de datos. *T* se refiere como el umbral de puntuación de palabra vecina (Altschul *et al.*, *supra*). Estos aciertos de palabra de vecindad iniciales actúan como semillas para iniciar las búsquedas para encontrar HSPs más largos que los contienen. Los aciertos de palabra se extienden en ambas direcciones a lo largo de cada secuencia hasta donde se pueda aumentar la puntuación de alineación acumulativa. Las puntuaciones acumuladas se calculan utilizando, para las secuencias de nucleótidos, los parámetros *M* (puntaje de recompensa para un par de restos emparejados; siempre >0) y *N* (puntaje de penalización para los restos mal emparejados; siempre <0). Para las secuencias de aminoácidos, se utiliza una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabra en cada dirección se detienen cuando: la puntuación de alineación acumulativa disminuye en la cantidad *X* desde su máximo valor alcanzado; la puntuación acumulativa llega a cero o menos, debido a la acumulación de una o más alineaciones de restos de puntuación negativa; o se alcanza el final de cualquier secuencia. Los parámetros *W*, *T*, y *X* del algoritmo BLAST determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) utiliza como valores por defecto una longitud de palabra (*W*) de 11, una expectativa (*E*) o 10, *M*=5, *N*=-4, y una comparación de ambas cadenas. Para las secuencias de aminoácidos, el programa BLASTP utiliza por defecto una longitud de palabra de 3, y una expectativa (*E*) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, *Proc. Natl. Acad. Sci. EE.UU.* 89:10915 (1989)), alineaciones (*B*) de 50, expectativa (*E*) de 10, *M*=5, *N*=-4, y una comparación de ambas cadenas.

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

La presente divulgación también se extiende a los polinucleótidos que comprenden una primera secuencia de nucleótidos que se hibrida selectivamente bajo condiciones moderadamente rigurosas (tal como bajo condiciones altamente rigurosas) al complemento de una segunda secuencia de nucleótidos que codifica un polipéptido que comprende:

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c.

5 La frase “condiciones de hibridación altamente rigurosas” se refiere a las condiciones bajo las cuales una sonda se hibridará a su subsecuencia diana, típicamente en una mezcla compleja de ácido nucleico, pero no a otras secuencias. Las condiciones altamente rigurosas dependen de la secuencia y serán diferentes en circunstancias diferentes. Las secuencias más largas se hibridan específicamente a temperaturas más altas. Se encuentra una extensa guía sobre la hibridación de ácidos nucleicos en Tijssen, *Techniques in Biochemistry and Molecular Biology -- Hybridization with Nucleic Probes*, “Overview of principles of hybridisation and the strategy of nucleic acid assays”  
 10 (1993). Generalmente, las condiciones altamente rigurosas se seleccionan para que sean aproximadamente 5-10°C más bajas que el punto de fusión térmico ( $T_m$ ) para la secuencia específica a un pH de fuerza iónica definido. La  $T_m$  es la temperatura (bajo una fuerza iónica, pH y concentración de ácido nucleico definidos) a la que 50% de las sondas complementarias a la diana hibridan a la secuencia diana en equilibrio (como las secuencias diana están presentes en exceso, a  $T_m$ , el 50% de las sondas están ocupadas en equilibrio). Las condiciones altamente  
 15 rigurosas serán aquellas en donde la concentración de sal sea menor de aproximadamente 1,0 M de ion sodio, típicamente una concentración de aproximadamente 0,01 a 1,0 M de ion sodio (u otras sales) a un pH de 7,0 a 8,3, y la temperatura es al menos aproximadamente 30°C para las sondas cortas (por ejemplo, 10 a 50 nucleótidos), y al menos aproximadamente 60°C para las sondas largas (por ejemplo, mayores de 50 nucleótidos). Las condiciones altamente rigurosas también se pueden lograr con la adición de agentes desestabilizantes, tales como formamida.  
 20 Para la hibridación selectiva o específica, una señal positiva es una hibridación de al menos dos veces el fondo, opcionalmente 10 veces la hibridación de fondo.

Ejemplos de condiciones de hibridación altamente rigurosas pueden ser las 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 lavado en 0,2x SSC, y 0,1% de SDS a 65°C.

25 Los ácidos nucleicos que no se 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 utilizando la máxima degeneración de codones permitida por el código genético. En tales casos, los ácidos nucleicos típicamente hibridan bajo condiciones de hibridación moderadamente rigurosas.

30 Ejemplos de “condiciones de hibridación moderadamente rigurosas” incluyen una hibridación en un tampón de formamida al 40%, NaCl 1 M, SDS al 1%, a 37°C, y un lavado en 1X SSC a 45°C. Una hibridación positiva es al menos dos veces el fondo. Los expertos comunes en la técnica reconocerán fácilmente que se pueden emplear condiciones alternativas de hibridación y lavado para proporcionar condiciones de rigurosidad similar.

35 La frase “hibrida selectivamente (o específicamente) a”, se refiere al enlace, duplex, o hibridación de una molécula solamente a una secuencia de nucleótidos particular, bajo condiciones de hibridación rigurosas, cuando esa secuencia está presente en una mezcla compleja (por ejemplo, ADN o ARN celular total o de biblioteca).

40 En cualquier caso, las variantes de una secuencia polipeptídica tendrán esencialmente la misma actividad que la secuencia de referencia (en el caso de los polinucleótidos, las secuencias de polinucleótidos variantes codificarán un polipéptido que tiene esencialmente la misma actividad que la secuencia de referencia). Esencialmente la misma actividad significa una actividad de al menos 50%, de una manera adecuada al menos 75%, y especialmente al menos 90% de la secuencia de referencia en un ensayo de re-estimulación *in vitro* de PBMC o de sangre entera con antígenos específicos (por ejemplo, re-estimulación durante un período 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 por medio de linfoproliferación, producción de citoquinas en el sobrenadante de cultivo (medidas mediante ELISA, CBA, etc.), o  
 45 caracterización de respuestas de células T y B mediante tinción intra y extracelular (por ejemplo, utilizando anticuerpos específicos para inmuno-marcadores, tales como CD3, CD4, CD8, IL2, TNFa, IFNg, CD40L, CD69, etc.), seguido por análisis con un citómetro de flujo. De una manera adecuada, esencialmente la misma actividad significa una actividad de al menos 50%, de una manera adecuada de al menos 75%, y especialmente de al menos 90% de la secuencia de referencia en un ensayo de proliferación de células T y/o de producción de IFN-gamma.

### Composiciones de polinucleótidos

50 Como se emplea en la presente memoria, 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, pero que está sustancialmente aislado del ADN genómico total de la especie a partir de la que se obtiene el polinucleótido, o está libre de él.

55 Como será entendido por los expertos en la técnica, los polinucleótidos para el uso en esta invención pueden incluir secuencias genómicas, secuencias extra-genómicas y codificadas por plásmidos, y segmentos genéticos modificados más pequeños que expresen, o que se puedan adaptar para expresar, proteínas, polipéptidos, péptidos,

y similares. Tales segmentos se pueden aislar naturalmente, o pueden ser modificados sintéticamente por la mano del hombre.

5 “Aislado”, como se emplea en la presente memoria, significa que un polinucleótido está sustancialmente lejos de otras secuencias de codificación, y que el polinucleótido no contiene grandes porciones de ADN codificante no relacionado, tales como los fragmentos cromosómicos grandes u otros genes funcionales o regiones codificantes de polipéptidos. Un ácido nucleico aislado se separa de otros marcos de lectura abierta que flanquean el gen y que codifican proteínas diferentes del gen. Por supuesto, esto se refiere al segmento de ADN como se aisló originalmente, y no excluye a los genes o a las regiones codificantes agregadas posteriormente al segmento por la mano del hombre.

10 Como será reconocido por el experto, los polinucleótidos pueden ser monocatenarios (codificante o anti-sentido) o bicatenarios, y pueden ser moléculas de ADN (genómico, ADNc, o sintético) o de ARN. Las moléculas de ARN incluyen moléculas de HnARN, que contienen intrones y corresponden a una molécula de ADN de una manera de uno a uno, y moléculas de ARNm, que no contienen intrones. Pueden estar presentes, pero no es necesario, secuencias codificantes o no codificantes adicionales dentro de un polinucleótido para el uso en la presente invención, y un polinucleótido puede, pero no es necesario, estar enlazado a otras moléculas y/o materiales de soporte.

15 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 tal secuencia. Las variantes de polinucleótidos pueden contener una o más sustituciones, adiciones, deleciones, y/o inserciones, como se describe adicionalmente más adelante, preferiblemente de manera que no se disminuya la inmunogenicidad del polipéptido codificado, en relación a la proteína de referencia. El efecto sobre la inmunogenicidad del polipéptido codificado se puede evaluar generalmente como se describe en la presente memoria.

20 La presente invención proporciona polinucleótidos y polipéptidos aislados que comprenden diferentes longitudes de tramos contiguos de secuencia idéntica o complementaria a una o más de las secuencias descritas en la presente memoria. Por ejemplo, esta invención proporciona polinucleótidos que comprenden al menos aproximadamente 30, 40, 50, 75, 100, 150, 200, 300, 400, 500, o 1.000 o más nucleótidos contiguos de la secuencia de referencia descrita en la presente memoria, 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 enteros a través de 200-500; 500-1.000, y similares.

25 Por otra parte, será apreciado por los expertos ordinarios en la técnica que, como resultado de la degeneración del código genético, hay muchas secuencias de nucleótidos que codifican un polipéptido como se describe en la presente memoria. Algunos de estos polinucleótidos tienen una identidad relativamente baja con la secuencia de nucleótidos de cualquier gen nativo. No obstante, la presente invención contempla específicamente los polinucleótidos que varían debido a las diferencias en el uso de codones, por ejemplo, los polinucleótidos que se optimizan para la selección de codones humanos y/o de primates. Además, los alelos de los genes que comprenden las secuencias de polinucleótidos proporcionadas en la presente memoria están dentro del alcance de la presente invención. Los alelos son genes endógenos que se alteran como resultado de una o más mutaciones, tales como deleciones, adiciones, y/o sustituciones de nucleótidos. El ARNm y la proteína resultantes pueden, pero no es necesario, tener una estructura o función alterada. Los alelos se pueden identificar empleando técnicas convencionales (tales como hibridación, amplificación, y/o comparación de secuencias de bases de datos).

### Identificación y caracterización de nucleótidos

35 Los polinucleótidos se pueden identificar, preparar, y/o manipular empleando cualquiera de una variedad de técnicas bien establecidas. Por ejemplo, un polinucleótido se puede identificar, como se describe con mayor detalle a continuación, mediante el cribado de una micromatriz de ADNcs. Tales cribados se pueden llevar a cabo, por ejemplo, utilizando una micromatriz de Synteni (Palo Alto, CA) según las instrucciones del fabricante (y esencialmente como se describe por Schena *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 93:10614-10619 (1996), y Heller *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 94:2150-2155 (1997)). Alternativamente, los polinucleótidos se pueden amplificar a partir del ADNc preparado a partir de células que expresan las proteínas descritas en la presente memoria, tales como las células de *M. tuberculosis*. Tales polinucleótidos se pueden amplificar mediante reacción en cadena de polimerasa (PCR). Para este planteamiento, se pueden diseñar cebadores específicos de la secuencia, basándose en las secuencias proporcionadas en la presente memoria, y se pueden comprar o sintetizar.

40 Se puede utilizar una porción amplificada de un polinucleótido para aislar un gen de longitud completa a partir de una biblioteca adecuada (por ejemplo, una biblioteca de ADNc de *M. tuberculosis*), empleando técnicas bien conocidas. Dentro de tales técnicas, se criba una biblioteca (ADNc o genómico) utilizando una o más sondas de polinucleótido o cebadores adecuados para la amplificación. Preferiblemente, una biblioteca se selecciona por tamaños para incluir moléculas más grandes. También se pueden preferir las bibliotecas con cebado aleatorio para identificar las regiones 5' y secuencia arriba de los genes. Se prefieren las bibliotecas genómicas para obtener

intrones y extender las secuencias 5'.

Para las técnicas de hibridación, una secuencia parcial se puede marcar (por ejemplo, mediante traducción de cortes o marcado final con <sup>32</sup>P) empleando técnicas bien conocidas. Entonces se rastrea en general una biblioteca bacteriana o bacteriófaga mediante filtros de hibridación que contienen colonias bacterianas desnaturalizadas (o pastos que contienen placas de fagos) con la sonda marcada (véase Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (2000)). Las colonias o placas que hibridan se seleccionan y se expanden, y se aísla el ADN para un análisis adicional. Los clones de ADNc se pueden analizar para determinar la cantidad de secuencia adicional, por ejemplo, mediante PCR, utilizando un cebador a partir de la secuencia parcial y un cebador a partir del vector. Se pueden generar mapas de restricción y secuencias parciales para identificar uno o más clones superpuestos. Entonces se puede determinar la secuencia completa empleando técnicas convencionales, que pueden implicar la generación de una serie de clones de delección. Luego se pueden ensamblar las secuencias superpuestas resultantes en una sola secuencia contigua. Se puede generar una molécula de ADNc de longitud completa ligando fragmentos adecuados, empleando técnicas bien conocidas.

Alternativamente, existen numerosas técnicas de amplificación para obtener una secuencia de codificación de longitud completa a partir de una secuencia de ADNc parcial. Dentro de estas técnicas, la amplificación se lleva a cabo generalmente mediante PCR. Se puede emplear cualquiera de una variedad de kits comercialmente disponibles para llevar a cabo la etapa de amplificación. Se pueden diseñar cebadores utilizando, por ejemplo, un programa informático bien conocido en la técnica. Los cebadores son preferiblemente de 22-30 nucleótidos de longitud, tienen un contenido de GC de al menos el 50%, y se reasocian a la secuencia diana a temperaturas de aproximadamente 68°C a 72°C. La región amplificada se puede secuenciar como se describe anteriormente, y superponerse secuencias ensambladas en una secuencia contigua.

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

En ciertos casos, es posible obtener una secuencia de ADNc de longitud completa mediante el análisis de secuencias proporcionadas en una base de datos de etiquetas de secuencia expresadas (EST), tal como la disponible en GenBank. Las búsquedas de las ESTs solapadas se pueden llevar a cabo generalmente empleando programas bien conocidos (por ejemplo, búsquedas de NCBI BLAST), y tales ESTs se pueden utilizar para generar una secuencia de longitud completa contigua. También se pueden obtener secuencias de ADN de longitud completa mediante el análisis de fragmentos genómicos.

### Expresión de polinucleótidos en células hospedadoras

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

Como será entendido por los expertos en la técnica, puede ser conveniente, en algunos casos, producir secuencias de nucleótidos que codifican polipéptidos que posean codones que no aparecen de manera natural. Por ejemplo, se pueden seleccionar codones preferidos por un huésped procarionota o eucariota particular para aumentar la velocidad de expresión de la proteína, o para producir un transcrito de ARN recombinante que tiene propiedades deseables, tal como una vida media que sea más larga que la de un transcrito generado a partir de la secuencia que aparece de manera natural.

Por otra parte, las secuencias de polinucleótidos se pueden diseñar empleando métodos generalmente conocidos en la técnica, para alterar las secuencias de codificación de polipéptidos por una variedad de razones, que incluyen, pero no se limitan a, alteraciones que modifican la clonación, procesamiento, y/o expresión del producto genético. Por ejemplo, para diseñar secuencias de nucleótidos se puede emplear la transposición de ADN mediante

fragmentación aleatoria y reensamble de PCR de fragmentos de genes y oligonucleótidos sintéticos. Además, se puede emplear mutagénesis dirigida al sitio para insertar nuevos sitios de restricción, alterar patrones de glicosilación, cambiar la preferencia de codones, producir variantes de empalme, o introducir mutaciones, etc.

5 Las secuencias de ácidos nucleicos naturales, modificadas, o recombinantes se pueden ligar a una secuencia heteróloga para codificar una proteína de fusión. Por ejemplo, para cribar las bibliotecas de péptidos en busca de inhibidores de la actividad del polipéptido, puede ser útil codificar una proteína quimérica que pueda reconocerse por un anticuerpo disponible comercialmente. También se puede modificar genéticamente una proteína de fusión para que contenga un sitio de disociación localizado entre la secuencia de codificación del polipéptido y la secuencia proteica heteróloga, de manera que el polipéptido se pueda disociar y purificar a parte de la fracción heteróloga.

10 Se pueden sintetizar secuencias que codifican un polipéptido deseado, del todo o en parte, empleando métodos químicos bien conocidos en la técnica (véase Caruthers, M. H. *et al.*, *Nucl. Acids Res. Symp. Ser.* páginas 215-223 (1980), Horn *et al.*, *Nucl. Acids Res. Symp. Ser.* páginas 225-232 (1980)).

15 Alternativamente, la proteína en sí misma se puede producir empleando métodos químicos para sintetizar la secuencia de aminoácidos de un polipéptido, o una porción de la misma. Por ejemplo, se puede llevar a cabo síntesis de péptidos empleando diferentes técnicas en fase sólida (Roberge *et al.*, *Science* 269:202-204 (1995)), y se puede lograr la síntesis automatizada, por ejemplo, empleando el Sintetizador de Péptidos ABI 431A (Perkin Elmer, Palo Alto, CA).

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

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

35 Se pueden utilizar una variedad de vectores de expresión/sistemas huésped para contener y expresar las secuencias de polinucleótidos. Éstos incluyen, pero no se limitan a, microorganismos, tales como bacterias transformadas con bacteriófagos recombinantes, plásmidos o vectores de expresión de ADN cósmido; levadura transformada con vectores de expresión de levadura; sistemas de células de insecto infectadas con vectores de expresión de virus (por ejemplo, baculovirus); sistemas de células vegetales transformados con vectores de expresión de virus (por ejemplo, virus del mosaico de la coliflor, CaMV; virus del mosaico del tabaco; TMV), o con vectores de expresión bacterianos (por ejemplo, plásmidos Ti o pBR322); o sistemas de células animales.

40 Los "elementos de control" o "secuencias reguladoras" presentes en un vector de expresión, son las regiones no traducidas de los potenciadores del vector, promotores, regiones no traducidas 5' y 3' que interactúan con las proteínas celulares del huésped para llevar a cabo la transcripción y la traducción. Tales elementos pueden variar en su fuerza y especificidad. Dependiendo del sistema de vector y del huésped utilizado, se puede utilizar cualquier número de elementos de transcripción y de traducción adecuados, incluyendo promotores constitutivos e inducibles. Por ejemplo, cuando se clona en sistemas bacterianos, se pueden utilizar promotores inducibles, tales como el promotor lacZ híbrido del fagémido PBLUESCRIPT (Stratagene, La Jolla, Calif.), o el plásmido PSFORT1 (Gibco BRL, Gaithersburg, MD), y similares. En sistemas celulares de mamífero, generalmente se prefieren los promotores a partir de genes de mamífero o a partir de virus de mamífero. Si es necesario generar una línea celular que contenga múltiples copias de la secuencia que codifica un polipéptido, se pueden utilizar convenientemente vectores basados en SV40 o EBV, con un marcador seleccionable apropiado.

55 En sistemas bacterianos, se puede seleccionar un número de vectores de expresión dependiendo del uso pretendido para el polipéptido expresado. Por ejemplo, cuando se necesitan grandes cantidades, por ejemplo, para la inducción de anticuerpos, se pueden utilizar vectores que dirijan una expresión de alto nivel de proteínas de fusión que se purifican fácilmente. Tales vectores incluyen, pero no se limitan a, los vectores de clonación y expresión en *E. coli* multifuncionales, tales como BLUESCRIPT (Stratagene), en los que la secuencia que codifica el polipéptido de interés se puede ligar en el vector en el marco con secuencias para el Met amino-terminal y los siguientes 7 restos de  $\beta$ -galactosidasa, de tal manera que se produce una proteína híbrida; vectores pIN (Van Heeke y Schuster, *J. Biol. Chem.* 264:5503-5509 (1989)); y similares. También se pueden utilizar vectores pGEX (Promega, Madison, Wis.) para expresar polipéptidos extraños como proteínas de fusión con glutatión S-transferasa (GST). En general, tales

proteínas de fusión son solubles y se pueden purificar fácilmente a partir de células lisadas mediante adsorción a perlas de glutatión-agarosa, seguido por elución en presencia de glutatión libre. Las proteínas preparadas en tales sistemas se pueden diseñar para incluir sitios de disociación de proteasa de heparina, trombina, o factor XA, de manera que el polipéptido clonado de interés pueda liberarse a voluntad del resto de GST

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

- 10 En los casos en los que se utilizan vectores de expresión vegetales, la expresión de secuencias que codifican los polipéptidos puede ser activada mediante cualquiera de varios promotores. Por ejemplo, se pueden utilizar promotores virales, tales como los promotores 35S y 19S de CaMV en solitario o en combinación con la secuencia líder omega a partir de TMV (Takamatsu, *EMBO J.* 6:307-311 (1987)). Alternativamente, se pueden utilizar promotores vegetales, tales como la subunidad pequeña de RUBISCO, o promotores de choque por calor (Coruzzi *et al.*, *EMBO J.* 3:1671-1680 (1984); Broglie *et al.*, *Science* 224:838-843 (1984); y Winter *et al.*, *Results Probl. Cell Differ.* 17:85-105 (1991)). Estas construcciones se pueden introducir en células vegetales mediante la transformación directa del ADN, o mediante transfección mediada por patógeno. Estas técnicas se describen en un número de revisiones generalmente disponibles (véase, por ejemplo, Hobbs en *McGraw Hill Yearbook of Science and Technology*, páginas 191-196 (1992)).

- 20 También se puede utilizar un sistema de insecto para expresar un polipéptido de interés. Por ejemplo, en uno de tales sistemas, se utiliza el virus de polihedrosis nuclear de *Autographa californica* (AcNPV) como un vector para expresar genes extraños en células de *Spodoptera frugiperda* o en *Trichoplusia larvae*. Las secuencias que codifican el polipéptido se pueden clonar en la región no esencial del virus, tal como el gen de polihedrina, y colocarse bajo el control del promotor de polihedrina. La inserción con éxito de la secuencia que codifica el polipéptido dejará inactivo el gen de polihedrina y producirá el virus recombinante que carece de la proteína de la cubierta. Entonces los virus recombinantes se pueden utilizar para infectar, por ejemplo, células de *S. frugiperda* o *Trichoplusia larvae*, en donde se puede expresar el polipéptido de interés (Engelhard *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 91:3224-3227 (1994)).

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

- 30 También se pueden utilizar señales de inicio específicas para lograr una traducción más eficiente de secuencias que codifican un polipéptido de interés. Tales señales incluyen el codón de inicio ATG y las secuencias adyacentes. En los casos en los que se inserten secuencias que codifican el polipéptido, su codón de inicio, y las secuencias corriente arriba, en el vector de expresión apropiado, pueden no necesitarse señales de control de transcripción o traducción adicionales. Sin embargo, en los casos en los que solamente se inserta la secuencia de codificación, o una porción de la misma, se deben proporcionar señales de control de traducción exógenas, incluyendo el codón de inicio ATG. Además, el codón de inicio debe estar en el marco de lectura correcto para asegurar la traducción del inserto entero. Los elementos de traducción exógenos y los codones de inicio pueden ser de diferentes orígenes, tanto naturales como sintéticos. La eficacia de la expresión se puede mejorar mediante la inclusión de potenciadores que son apropiados para el sistema celular particular que se utiliza, tales como los descritos en la bibliografía (Scharf *et al.*, *Results Probl. Cell Differ.* 20:125-162 (1994)).

- 35 Además, se puede seleccionar una cepa de células huésped por su capacidad para modular la expresión de las secuencias insertadas, o para procesar la proteína expresada en la forma deseada. Tales modificaciones del polipéptido incluyen, pero no se limitan a, acetilación, carboxilación, glicosilación, fosforilación, lipidación, y acilación. También se puede emplear el procesamiento post-traduccionales que disocia una forma "prepro" de la proteína, para facilitar la inserción, el plegamiento, y/o la función correctos. Se pueden seleccionar diferentes células huésped, tales como CHO, HeLa, MDCK, HEK293, y W138, que tienen una maquinaria celular específica y mecanismos característicos para tales actividades post-traduccionales, para asegurar la modificación y el procesamiento correctos de la proteína extraña.

- 40 Para la producción de alto rendimiento a largo plazo de proteínas recombinantes, generalmente se prefiere la expresión estable. Por ejemplo, las líneas celulares que expresan establemente un polinucleótido de interés, se pueden transformar utilizando vectores de expresión que pueden contener orígenes virales de replicación y/o elementos de expresión endógenos, y un gen marcador seleccionable, sobre el mismo vector o sobre un vector

separado. Después de la introducción del vector, se puede permitir que las células crezcan durante 1-2 días en un medio enriquecido antes de cambiarse al medio selectivo. El propósito del marcador seleccionable es conferir resistencia a la selección, y su presencia permite el crecimiento y la recuperación de las células que expresan con éxito las secuencias introducidas. Los clones resistentes de las células transformadas establemente se pueden proliferar empleando técnicas de cultivo de tejido apropiadas para el tipo de célula.

Se puede utilizar varios sistemas de selección para recuperar las líneas celulares transformadas. Éstas incluyen, pero no se limitan a, timidina quinasa del virus del herpes simple (Wigler *et al.*, *Cell* 11:223-32 (1977)), y adenina fosforribosiltransferasa (Lowy *et al.*, *Cell* 22:817-23 (1990)), genes que se pueden emplear en células tk.sup. o aprt.sup., respectivamente. También, se puede utilizar la resistencia a antimetabolitos, a antibióticos, o a herbicidas como base para la selección; por ejemplo, dhfr que confiere resistencia al metotrexato (Wigler *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 77:3567-70 (1980)); npt, que confiere resistencia a los aminoglicósidos, neomicina y G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol.* 150:1-14 (1981)); y als o pat, que confieren resistencia al clorsulfurón y a la fosfotricina acetiltransferasa, respectivamente (Murry, *supra*). Se han descrito genes seleccionables adicionales, por ejemplo, trpB, que permite que las células utilicen indol en lugar de triptófano, o hisD, que permite que las células utilicen histinol en lugar de histidina (Hartman y Mulligan, *Proc. Natl. Acad. Sci. EE.UU.* 85:8047-51 (1988)). Recientemente, el uso de marcadores visibles ha ganado popularidad con marcadores tales como antocianinas,  $\beta$ -glucuronidasa y su sustrato GUS, y luciferasa y su sustrato luciferina, siendo ampliamente utilizados no sólo para identificar transformantes, sino también para cuantificar la cantidad de expresión de proteína transitoria o estable atribuible a un sistema de vector específico (Rhodes *et al.*, *Methods Mol. Biol.* 55:121-131 (1995)).

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

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

En la técnica se conocen una variedad de protocolos para detectar y medir la expresión de productos codificados por los polinucleótidos, utilizando bien anticuerpos policlonales o bien anticuerpos monoclonales específicos para el producto. Ejemplos incluyen el ensayo inmunoabsorbente ligado a enzimas (ELISA), radioinmunoensayo (RIA), y clasificación celular activada por fluorescencia (FACS). Para algunas aplicaciones, se puede preferir un inmunoensayo basado en monoclonal, de dos sitios, que utiliza anticuerpos monoclonales que reaccionan con dos epítopos no interferentes sobre un polipéptido dado, pero también se puede emplear un ensayo de enlace competitivo. Estos y otros ensayos se describen, entre otros lugares, en Hampton *et al.*, *Serological Methods, a Laboratory Manual* (1990), y en Maddox *et al.*, *J. Exp. Med.* 158:1211-1216 (1983).

Los expertos en la técnica conocen una amplia variedad de etiquetas y técnicas de conjugación, y se pueden utilizar en diferentes ensayos de ácidos nucleicos y aminoácidos. Los medios para producir hibridación marcada o sondas PCR para detectar secuencias relacionadas con los polinucleótidos incluyen oligomarcado, traducción de cortes, marcado final, o amplificación PCR utilizando un nucleótido marcado. Alternativamente, las secuencias, o cualquiera de las porciones de las mismas, se pueden clonar en un vector para la producción de una sonda de ARNm. Tales vectores se conocen en la técnica, están disponibles comercialmente, y se pueden utilizar para sintetizar sondas de ARN *in vitro* mediante la adición de una ARN polimerasa apropiada, tal como T7, T3, o SP6, y nucleótidos marcados. Estos procedimientos se pueden realizar empleando una variedad de kits disponibles comercialmente. Las moléculas o etiquetas indicadoras adecuadas que se pueden utilizar incluyen radionucleidos, enzimas, agentes fluorescentes, quimioluminescentes, o cromogénicos, así como sustratos, cofactores, inhibidores, partículas magnéticas, y similares.

Las células huésped transformadas con una secuencia de polinucleótido de interés se pueden cultivar bajo condiciones adecuadas para la expresión y recuperación de la proteína a partir del cultivo celular. La proteína producida por una célula recombinante se puede secretar o contener intracelularmente dependiendo de la secuencia y/o del vector utilizado. Como será entendido por los expertos en la técnica, se pueden diseñar vectores de expresión que contengan los polinucleótidos para contener secuencias de señales que dirigen la secreción del polipéptido codificado a través de una membrana celular procariota o eucariota. Se pueden utilizar otras construcciones recombinantes para unir secuencias que codifican un polipéptido de interés a la secuencia de nucleótidos que codifica un dominio de polipéptido que facilitará la purificación de las proteínas solubles. Tales dominios que facilitan la purificación incluyen, pero no se limitan a, péptidos quelantes de metales, tales como módulos de histidina-triptófano, que permiten la purificación sobre metales inmovilizados, dominios de proteína A que permiten la purificación sobre inmunoglobulina inmovilizada, y el dominio utilizado en el sistema de purificación

de extensión/afinidad FLAGS (Immunex Corp., Seattle, Wash.). La inclusión de secuencias enlazadoras disociables, tales como aquellas específicas para el factor XA o la enterocinasa (Invitrogen, San Diego, Calif.) entre el dominio de purificación y el polipéptido codificado, se pueden utilizar para facilitar la purificación. Uno de tales 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 precedentes a una tiorredoxina o a un sitio de escisión de enterocinasa. Los restos de histidina facilitan la purificación en la IMIAC (cromatografía por afinidad de iones metálicos inmovilizados), como se describe en Porath *et al.*, *Prot. Exp. Purif.* 3:263-281 (1992), mientras que el sitio de escisión de enterocinasa proporciona un medio para purificar el polipéptido deseado a partir de la proteína de fusión. Una discusión de los vectores que contienen proteínas de fusión se proporciona en Kroll *et al.*, *DNA Cell Biol.* 12:441-453 (1993).

### Técnicas de entrega de polinucleótidos *in vivo*

En realizaciones adicionales, se introducen construcciones genéticas que comprenden uno o más de los polinucleótidos de la invención en células *in vivo*. Esto se puede lograr empleando cualquiera de una variedad de planteamientos bien conocidos, varios de los cuales se ilustran a continuación con propósitos de ilustración.

#### 15 1. Adenovirus

Uno de los métodos preferidos para el suministro *in vivo* de una o más secuencias de ácidos nucleicos implica el uso de un vector de expresión de adenovirus. "Vector de expresión de adenovirus" pretende incluir aquellas 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 en sentido o anti-sentido. Por supuesto, en el contexto de una construcción anti-sentido, la expresión no requiere que se sintetice el producto genético.

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

El adenovirus es particularmente adecuado para utilizarse como un vector de transferencia génica, debido a su genoma de tamaño mediano, a su facilidad de manipulación, alta titulación, amplio intervalo de células diana, y a su alta infectividad. Ambos extremos del genoma viral contienen repeticiones invertidas de 100-200 pares de bases (ITRs), que son elementos *cis* necesarios para la replicación y el empaquetamiento del ADN viral. Las regiones temprana (E) y tardía (L) del genoma contienen diferentes unidades de transcripción que se dividen mediante el inicio de la replicación del ADN viral. La región E1 (E1A y E1B) codifica las proteínas responsables de la regulación de la transcripción del genoma viral y unos cuantos genes celulares. La expresión de la región E2 (E2A y E2B) da como resultado la síntesis de las proteínas para la replicación del ADN viral. Estas proteínas están implicadas en la replicación del ADN, la expresión tardía del gen, y la desactivación de la célula huésped (Renan, 1990). Los productos de los genes tardíos, que incluyen la mayoría de las proteínas de la cápside viral, se expresan solamente después del procesamiento significativo de un único transcrito primario emitido por el principal promotor tardío (MLP). El MLP (localizado a 16,8 m.u.) es particularmente eficaz durante la fase tardía de la infección, y todo el ARNm emitido a partir de este promotor posee una secuencia líder tripartita-5' (TPL) que los convierte en el ARNm preferido para la traducción.

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

La generación y propagación de los vectores de adenovirus actuales, que son deficientes en la replicación, dependen de una línea celular auxiliar única, designada como 293, que se transformó a partir de células de riñón embrionario humano, mediante fragmentos de ADN Ad5, y que expresa constitutivamente proteínas E1 (Graham *et al.*, 1977). Debido a 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 ya sea en la región E1, en la región D3, o en ambas regiones (Graham y Prevec, 1991). En la naturaleza, el adenovirus puede empaquetar aproximadamente el 105% del genoma de tipo salvaje (Ghosh-Choudhury *et al.*, 1987), proporcionando la capacidad para aproximadamente 2 kB extra de ADN. Combinado con las aproximadamente 5,5 kB de ADN que son reemplazables en las regiones E1 y E3, la máxima capacidad del vector de adenovirus actual es inferior a 7,5 kB, o aproximadamente el 15% de la longitud total del vector. Más del 80% del genoma viral del adenovirus permanece en la cadena principal del vector, y es la fuente de la citotoxicidad transmitida por el vector. Además, la deficiencia en la

replicación del virus con E1 suprimido es incompleta. Por ejemplo, se ha observado una fuga de la expresión génica viral con los vectores actualmente disponibles a altas multiplicidades de infección (MOI) (Mulligan, 1993).

Las líneas celulares auxiliares se pueden derivar a partir de células humanas, tales como células de riñón embrionarias humanas, células musculares, células hematopoyéticas, u otras células mesenquimales o epiteliales embrionarias humanas. Alternativamente, las células auxiliares se pueden derivar a partir de células de otras especies de mamífero que sean permisivas para el adenovirus humano. Estas células incluyen, por ejemplo, las células Vero u otras células mesenquimales o epiteliales embrionarias de mono. Como se mencionó anteriormente, la línea celular auxiliar actualmente preferida es la 293.

Racher *et al.* (1995) han descrito mejores métodos para cultivar las células 293 y propagar el adenovirus. En un formato, se cultivan agregados celulares naturales inoculando células individuales en matraces giratorios siliconizados de 1 litro (Techne, Cambridge, Reino Unido) que contienen 100-200 ml del medio. Después de agitar a 40 rpm, se estima la viabilidad celular con azul de tripano. En otro formato, se emplean microportadores Fibra-Cel (Bibby Sterlin, Stone, Reino Unido) (5 g/l) como sigue. Se agrega un inóculo celular, resuspendido en 5 ml del medio, al vehículo (50 ml) en un matraz Erlenmeyer de 250 ml, y se deja estacionario, con agitación ocasional, durante 1 a 4 horas. Luego se reemplaza el medio con 50 ml de medio fresco, y se inicia la agitación. Para la producción del virus, se permite que las células crezcan hasta aproximadamente una confluencia del 80%, después de cuyo tiempo, se reemplaza el medio (hasta el 25% del volumen final), y se agrega el adenovirus a una MOI de 0,05. Los cultivos se dejan estacionarios durante la noche, seguido de lo cual, se aumenta el volumen hasta el 100%, y se comienza la agitación durante otras 72 horas.

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

Como se mencionó anteriormente, el vector típico según la presente invención es defectuoso en 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 a partir de la cual se hayan eliminado las secuencias que codifican E1. Sin embargo, la posición de inserción de la construcción dentro de las secuencias de adenovirus no es crítica para la invención. El polinucleótido que codifica el gen de interés también se puede insertar en lugar de la región E3 suprimida, en los vectores de reemplazo de E3, como se describe por Karlsson *et al.* (1986), o en la región E4, donde una línea celular auxiliar o un virus auxiliar complementa al defecto de E4.

El adenovirus es fácil de cultivar y manipular, y exhibe un amplio rango de huéspedes *in vitro* e *in vivo*. Este grupo de virus se puede obtener en altas titulaciones, por ejemplo, de  $10^9$ - $10^{11}$  unidades formadoras de placas por ml, y son altamente infecciosos. El ciclo de vida del adenovirus no requiere de la integración en el genoma de la célula huésped. Los genes extraños suministrados por los vectores de adenovirus son episomales, y por lo tanto, tienen una baja genotoxicidad para las células huésped. No se han reportado efectos secundarios en estudios de vacunación con adenovirus de tipo salvaje (Couch *et al.*, 1963; Top *et al.*, 1971), demostrando su seguridad y su potencial terapéutico como vectores de transferencia genética *in vivo*.

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

Los vectores de adenovirus se pueden originar a partir de adenovirus humanos. Alternativamente, se pueden originar a partir de adenovirus de otras especies, por ejemplo de chimpancé, que pueden tener la ventaja de que los vectores virales no se neutralizan por los anticuerpos contra los adenovirus humanos que circulan en muchos sujetos humanos (véase, por ejemplo, Tatsis N *et al.* *Gene Therapy* 2006 13:421-429).

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

## 2. Retrovirus

Los retrovirus son un grupo de virus de ARN monocatenarios caracterizados por la capacidad para convertir su ARN en ADN bicatenario en células infectadas mediante un proceso de transcripción inversa (Coffin, 1990). Entonces el

ADN resultante se integra establemente en los cromosomas celulares como un provirus, y dirige la síntesis de proteínas virales. La integración da como resultado la retención de las secuencias genéticas virales en la célula receptora y sus descendientes. El genoma retroviral contiene tres genes, gag, pol, y env, que codifican las proteínas de la cápside, la enzima polimerasa, y los componentes de envoltura, respectivamente. Una secuencia que se encuentra corriente arriba desde el gen gag, contiene una señal para empaquetar el genoma en viriones. Dos secuencias de repetición terminal larga (LTR) están presentes en los extremos 5' y 3' del genoma viral. Éstas contienen fuertes secuencias promotoras y potenciadoras, y también se requieren para la integración en el genoma de la célula huésped (Coffin, 1990).

Para construir un vector retroviral, se inserta un ácido nucleico que codifica una o más secuencias de oligonucleótidos o de polinucleótidos de interés, en el genoma viral, en el lugar de ciertas secuencias virales, para producir un virus que es defectuoso en la replicación. Para producir viriones, se construye una línea celular de empaquetamiento que contiene los genes gag, pol, y env, pero sin los componentes LTR y de empaquetamiento (Mann *et al.*, 1983). Cuando se introduce un plásmido recombinante que contiene un ADNc, junto con la LTR retroviral y secuencias de empaquetamiento 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 luego se secretan en los medios de cultivo (Nicolas y Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). Entonces se recolecta el medio que contiene los retrovirus recombinantes, opcionalmente se concentra, y se utiliza para la transferencia genética. Los vectores retrovirales son capaces de infectar a una amplia variedad de tipos de células. Sin embargo, la integración y la expresión estable requieren de la división de las células huésped (Paskind *et al.*, 1975).

Recientemente se desarrolló un planteamiento novedoso diseñado para permitir la dirección específica de los vectores de retrovirus, basado en la modificación química de un retrovirus mediante la adición química de restos de lactosa a la envoltura viral. Esta modificación podría permitir la infección específica de los hepatocitos por medio de receptores sialoglicoproteicos.

Se diseñó un planteamiento diferente para dirigir los retrovirus recombinantes, en el que se utilizaron anticuerpos biotinilados contra una proteína de envoltura retroviral y contra un receptor celular específico. Los anticuerpos se acoplaron mediante los componentes de biotina, utilizando estreptavidina (Roux *et al.*, 1989). Utilizando anticuerpos contra antígenos del complejo mayor de histocompatibilidad de clase I y clase II, demostraron la infección de una variedad de células humanas que portaban estos antígenos superficiales con un virus ecotrópico *in vitro* (Roux *et al.*, 1989).

### 3. Virus adeno-asociados

AAV (Ridgeway, 1988; Hermonat y Muzycska, 1984) es un parvovirus, descubierto como una contaminación de los suministros adenovirales. Es un virus ubicuo (los anticuerpos están presentes en el 85% de la población humana de EE.UU.) que no se ha vinculado 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 AAV-2 es el mejor caracterizado. AAV tiene un ADN lineal monocatenario que se encapsula en las proteínas de la cápside VP1, VP2, y VP3 para formar un virión icosaédrico de 20 a 24 nm de diámetro (Muzyczka y McLaughlin, 1988).

El ADN de AAV es de aproximadamente 4,7 kilobases de longitud. Contiene dos marcos de lectura abierta, y está flanqueado por dos ITRs. Existen dos genes principales en el genoma AAV: *rep* y *cap*. El gen *rep* codifica las proteínas responsables de las replications virales, mientras que *cap* codifica la proteína de la cápside VP1-3. Cada ITR forma una estructura de horquilla en forma de T. Estas repeticiones terminales son los únicos componentes *cis* esenciales del AAV para la integración cromosómica. Por lo tanto, el AAV se puede utilizar como un vector con todas las secuencias de codificación viral eliminadas y reemplazadas por el casete de genes para el suministro. Se han identificado tres promotores virales, y se denominaron como p5, p19, y p40, según su posición en el mapa. La transcripción de p5 y p19 da como resultado la producción de proteínas rep, y la transcripción de p40 produce las proteínas de la cápside (Hermonat y Muzyczka, 1984).

Existen varios factores que hicieron que los investigadores estudiaran la posibilidad de utilizar rAAV como un vector de expresión. Uno es que los requerimientos para suministrar un gen para que se integre en el cromosoma huésped son sorprendentemente pocos. Es necesario tener los ITRs de 145 pb, que son solamente el 6% del genoma de AAV. Esto deja espacio en el vector para ensamblar una inserción de ADN de 4,5 kb. Aunque esta capacidad portadora puede impedir que el AAV suministre genes grandes, es ampliamente adecuada para suministrar construcciones antisentido.

AAV también es una buena elección de vehículos de suministro debido a su seguridad. Hay un mecanismo de rescate relativamente complicado: no sólo se requiere el adenovirus de tipo salvaje sino también los genes AAV para movilizar rAAV. De la misma manera, AAV no es patogénico y no está asociado con ninguna enfermedad. La eliminación de secuencias de codificación viral minimiza las reacciones inmunes a la expresión del gen viral, y por lo tanto, rAAV no provoca una respuesta inflamatoria.

#### 4. Otros vectores virales como construcciones de expresión

Se pueden emplear otros vectores virales como construcciones de expresión en la presente invención para el suministro de secuencias de oligonucleótidos o polinucleótidos a una célula huésped. Se pueden emplear vectores derivados a partir de virus, tales como virus vaccinia (Ridgeway, 1988; Coupar *et al.*, 1988), lentivirus, virus de la polio, y virus del herpes. Se puede esperar que también sean útiles otros vectores derivados de poxvirus, tales como vectores derivados de la viruela aviar. Éstos ofrecen varias características atractivas para diferentes células de mamífero (Friedmann, 1989; Ridgeway, 1988; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

Con el reciente reconocimiento de virus de hepatitis B defectuosos, se obtuvo una nueva perspectiva sobre la relación estructura-función de diferentes secuencias virales. Los estudios *in vitro* mostraron que el virus podría retener la capacidad para el empaquetamiento dependiente del auxiliar y la transcripción inversa, a pesar de la supresión de hasta el 80% de su genoma (Horwich *et al.*, 1990). Esto sugirió que grandes porciones del genoma se podrían reemplazar con material genético extraño. El hepatotropismo y la persistencia (integración) fueron propiedades particularmente atractivas para la transferencia del gen dirigida al hígado. Chang *et al.* (1991) introdujeron el gen de cloranfenicol acetiltransferasa (CAT) en el genoma del virus de hepatitis B de pato en lugar de las secuencias de codificación de polimerasa, superficiales, y pre-superficiales. Se co-transfectó con el virus de tipo salvaje en una línea celular de hepatoma aviar. Se utilizaron medios de cultivo que contienen altas titulaciones del virus recombinante para infectar los hepatocitos de patito primario. La expresión estable del gen CAT se detectó durante al menos 24 días después de la transfección (Chang *et al.*, 1991).

Los vectores "virales" adicionales incluyen partículas tipo virus (VLPs) y fagos.

#### 5. Vectores no virales

Para efectuar la expresión de las secuencias de oligonucleótidos o polinucleótidos, se debe suministrar la construcción de expresión a una célula. Este suministro se puede llevar a cabo *in vitro*, como en los procedimientos de laboratorio para transformar líneas celulares, o *in vivo* o *ex vivo*, como en el tratamiento de ciertos estados de enfermedad. Como se describe anteriormente, un mecanismo preferido para el suministro es mediante infección viral donde la construcción de expresión se encapsula en una partícula viral infecciosa.

Una vez que se ha suministrado la construcción de expresión a la célula, el ácido nucleico que codifica las secuencias de oligonucleótidos o polinucleótidos deseadas se puede colocar y expresar en diferentes sitios. El ácido nucleico que codifica la construcción se puede integrar establemente en el genoma de la célula. Esta integración puede estar en la localización y orientación específicas mediante recombinación homóloga (reemplazo de genes), o se puede integrar en una localización aleatoria no específica (aumento de genes). El ácido nucleico se puede mantener establemente en la célula como un segmento episomal separado de ADN. Tales segmentos de ácidos nucleicos o "episomas" codifican secuencias suficientes para permitir el mantenimiento y la replicación, independientemente de, o en sincronización con, el ciclo de la célula huésped. Cómo se suministra la construcción de expresión a una célula, y en dónde permanece el ácido nucleico en la célula, depende del tipo de construcción de expresión empleada.

La construcción de expresión que comprende una o más secuencias de oligonucleótidos o polinucleótidos puede consistir simplemente en ADN o plásmidos recombinantes desnudos. La transferencia de la construcción se puede llevar a cabo, por ejemplo, mediante cualquier método que permeabilice física o químicamente la membrana celular. Esto es particularmente aplicable para la transferencia *in vitro*, pero se puede aplicar también al uso *in vivo*. Dubensky *et al.* (1984) inyectaron con éxito ADN de poliomavirus en forma de precipitados de fosfato de calcio en hígado y bazo de ratones adultos y recién nacidos, demostrando la replicación viral activa y una infección aguda. Benvenisty y Reshef (1986) también demostraron que la inyección intraperitoneal directa de plásmidos precipitados con fosfato de calcio, da como resultado la expresión de los genes transfectados. Se prevé que el ADN que codifica un gen de interés también se puede transferir de una manera similar *in vivo*, y expresar el producto genético.

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

Los órganos seleccionados, que incluyen hígado, piel, y tejido muscular de ratas y ratones, se han bombardeado *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). Esto puede requerir de la exposición quirúrgica del tejido o de las células, para eliminar cualquier tejido intermedio entre la pistola y el órgano diana, es decir, el tratamiento *ex vivo*. Nuevamente, mediante este método puede suministrarse ADN que codifica un gen particular, y aún así incorporarse.

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

### Composiciones polipeptídicas

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

En general, un polipéptido para el uso de la invención será un polipéptido aislado (es decir, separado de aquellos componentes con los que normalmente se puede encontrar en la naturaleza).

5 Por ejemplo, una proteína aparece de manera natural se aísla si se separa de algunos o de todos los materiales co-existentes en el sistema natural. Preferiblemente, tales polipéptidos son al menos aproximadamente 90% puros, más preferiblemente al menos aproximadamente 95% puros, y lo más preferiblemente al menos aproximadamente 99% puros. Se considera que un polinucleótido está aislado si, por ejemplo, se clona en un vector que no forma parte del medio ambiente natural.

10 Los polipéptidos se pueden preparar empleando cualquiera de una variedad de técnicas bien conocidas. Los polipéptidos recombinantes codificados por secuencias de ADN, como se describen anteriormente, se pueden preparar fácilmente a partir de secuencias de ADN empleando cualquiera de una variedad de vectores de expresión conocidos por los expertos ordinarios en la técnica. La expresión se puede lograr en cualquier célula huésped apropiada que se ha transformado o transfectado con un vector de expresión que contiene una molécula de ADN  
15 que codifica un polipéptido recombinante. Las células huésped adecuadas incluyen procariotas, levaduras, y células eucariotas superiores, tales como células de mamífero y células vegetales. Preferiblemente, las células huésped empleadas son *E. coli*, levadura, o una línea celular de mamífero, tales como COS o CHO. Los sobrenadantes de sistemas de huésped/vector adecuados que secreten la proteína o el polipéptido recombinante en el medio de cultivo, se pueden concentrar primero utilizando un filtro disponible comercialmente. Después de la concentración, el  
20 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 emplear uno o más etapas de HPLC en fase inversa para purificar adicionalmente un polipéptido recombinante.

Los polipéptidos para emplearse en la invención, fragmentos inmunogénicos de los mismos, y otras variantes que tienen menos de aproximadamente 100 aminoácidos, y en general menos de aproximadamente 50 aminoácidos,  
25 también se pueden generar por medios sintéticos, empleando técnicas bien conocidas por los expertos ordinarios en la técnica. Por ejemplo, tales polipéptidos se pueden sintetizar empleando cualquiera de las técnicas en fase sólida disponibles comercialmente, tal como el método de síntesis en fase sólida de Merrifield, donde los aminoácidos se agregan en secuencia a una cadena de aminoácidos creciente. Véase Merrifield, *J. Am. Chem. Soc.* 85:2149-2146 (1963). El equipo para la síntesis automatizada de polipéptidos está disponible comercialmente con proveedores  
30 tales como Perkin Elmer/Applied BioSystems Division (Foster City, CA), y se puede operar según 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 la presente memoria, o que comprende al menos un polipéptido como se describe en la presente memoria y una secuencia no relacionada, ejemplos de tales proteínas incluyen las  
35 proteínas de tétanos, tuberculosis, y hepatitis (véase, por ejemplo, Stoute *et al.*, *New Engl. J. Med.* 336:86-91 (1997)). Un compañero de fusión, por ejemplo, puede asistir para proporcionar epítomos T auxiliares (un compañero de fusión inmunológico), preferiblemente epítomos T auxiliares reconocidos por los seres humanos, o puede asistir en la expresión de la proteína (un potenciador de expresión) en rendimientos más altos que la proteína recombinante nativa. Ciertos compañeros de fusión preferidos son compañeros de fusión tanto inmunológicos como  
40 potenciadores de la expresión. Se pueden seleccionar otros compañeros de fusión para aumentar la solubilidad de la proteína, o para hacer posible que la proteína se dirija hacia los compartimientos intracelulares deseados. Además, los compañeros de fusión incluyen etiquetas de afinidad, que facilitan la purificación de la proteína.

Las proteínas de fusión se pueden preparar en general empleando técnicas convencionales, incluyendo la conjugación química. Preferiblemente, una proteína de fusión se expresa como una proteína recombinante, que  
45 permite la producción de mayores niveles, en relación con una proteína no fusionada, en un sistema de expresión. En resumen, las secuencias de ADN que codifican componentes polipeptídicos se pueden ensamblar por separado, y se pueden ligar en un vector de expresión apropiado. El extremo 3' de la secuencia de ADN que codifica un componente polipeptídico se liga, con o sin un enlazador peptídico, al extremo 5' de una secuencia de ADN que codifica el segundo componente polipeptídico, de tal manera que los marcos de lectura de las secuencias estén en fase. Esto permite la traducción a una sola proteína de fusión que mantiene la actividad biológica de ambos  
50 polipéptidos componentes.

Se puede emplear una secuencia enlazadora peptídica para separar el primer y el segundo componente polipeptídico en una distancia suficiente para asegurar que cada polipéptido se pliegue en sus estructuras secundaria y terciaria. Tal secuencia enlazadora peptídica se incorpora en la proteína de fusión empleando técnicas  
55 convencionales bien conocidas en la técnica. Las secuencias enlazadoras peptídicas adecuadas se pueden seleccionar basándose en los siguientes factores: (1) su capacidad para adoptar una conformación extendida flexible; (2) su incapacidad para adoptar una estructura secundaria que podría interactuar con epítomos funcionales en el primer y segundo polipéptido; y (3) la falta de restos hidrofóbicos o cargados que podrían reaccionar con los epítomos funcionales polipeptídicos. Las secuencias enlazadoras peptídicas preferidas contienen restos Gly, Asn, y

Ser. También se pueden utilizar en la secuencia enlazadora otros aminoácidos casi neutros, tales como Thr y Ala. Las secuencias de aminoácidos que se pueden emplear útilmente como enlazadoras incluyen aquellas que se describen en Maratea *et al.*, *Gene* 40:39-46 (1985); Murphy *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 83:8258-8262 (1986); la Patente U.S. N° 4.935.233, y la Patente U.S. N° 4.751.180. La secuencia enlazadora puede ser en general de 1 a aproximadamente 50 aminoácidos de longitud. Las secuencias enlazadoras no se requieren cuando el primer y el segundo polipéptido tienen regiones de aminoácidos N-terminales no esenciales que se pueden utilizar para separar los dominios funcionales y para prevenir la interferencia estérica.

En realizaciones preferidas, un compañero de fusión inmunológico se deriva a partir de la proteína D, una proteína superficial de la bacteria gram-negativa *Haemophilus influenza* B (WO 91/18926). Preferiblemente, un derivado de proteína D comprende aproximadamente el primer tercio de la proteína (por ejemplo, los primeros 100-110 aminoácidos N-terminales), y un derivado de proteína D puede lipídarse. Dentro de ciertas realizaciones preferidas, se incluyen los primeros 109 restos de un compañero de fusión de lipoproteína D en el extremo N para proporcionar al polipéptido epítomos de células T exógenos adicionales, y para aumentar el nivel de expresión en *E. coli* (funcionando, por tanto, como un potenciador de expresión). La cola del lípido asegura la presentación óptima del antígeno a las células presentadoras de antígeno. Otros compañeros de fusión incluyen la proteína no estructural del virus de influenza, NS1 (hemaglutinina). Típicamente, se utilizan los 81 aminoácidos N-terminales, aunque se pueden utilizar diferentes fragmentos que incluyan a los epítomos T auxiliares.

En otra realización, el compañero de fusión inmunológico es una proteína conocida como LYTA, o una porción de la misma (preferiblemente una porción C-terminal). LYTA se deriva de *Streptococcus pneumoniae*, que sintetiza una N-acetil-L-alanina amidasa conocida como amidasa LYTA (codificada por el gen *LytA*; *Gene* 43:265-292 (1986)). LYTA es una autolisina que degrada específicamente ciertos enlaces en la estructura base del peptidoglicano. El dominio C-terminal de la proteína LYTA es responsable de la afinidad con la colina, o con algunos análogos de colina, tales como DEAE. Esta propiedad se ha explotado para el desarrollo de plásmidos que expresan C-LYTA de *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 el amino terminal (véase *Biotechnology* 10:795-798 (1992)). Dentro de una realización preferida, se puede incorporar una porción repetida de LYTA en una proteína de fusión. Una porción repetida se encuentra en la región C-terminal que comienza en el resto 178. Una porción de repetición particularmente preferida incorpora los restos 188-305.

### Composiciones farmacéuticas

Las composiciones de polipéptidos, y/o anticuerpos para el uso en la invención descritas en la presente memoria, se formularán en disoluciones aceptables farmacéuticamente o aceptables fisiológicamente, para la administración a una célula o a un animal, ya sea solas, o bien en combinación con una o más de otras modalidades de terapia.

También se entenderá que, si se desea, el segmento de ácido nucleico (por ejemplo, ARN o ADN) que expresa un polipéptido como se describe en la presente memoria, se puede administrar también en combinación con otros agentes, tales como, por ejemplo, otras proteínas o polipéptidos, o varios agentes activos farmacéuticamente, incluyendo agentes quimioterapéuticos eficaces contra una infección por *M. tuberculosis*. De hecho, prácticamente no hay límite para otros componentes que también se pueden incluir, dado que los agentes adicionales no provocan un efecto adverso significativo al contacto con las células diana o los tejidos del huésped. Por tanto, las composiciones se pueden suministrar junto con otros agentes diferentes, según se requiera en la instancia particular. Tales composiciones se pueden purificar a partir de células huésped o de otras fuentes biológicas, o alternativamente, se pueden sintetizar químicamente como se describe en la presente memoria. Asimismo, tales composiciones pueden comprender además composiciones de ARN o ADN sustituidas o derivadas.

La formulación de los excipientes y disoluciones de vehículo aceptables farmacéuticamente es bien conocida por los expertos en la técnica, así como el desarrollo de regímenes de dosificación y tratamiento adecuados para utilizar las composiciones particulares descritas en la presente memoria en una variedad de regímenes de tratamiento, incluyendo, por ejemplo, la administración y formulación oral, parenteral, intravenosa, intranasal, e intramuscular. Otras vías de administración incluyen a través de las superficies de mucosas.

Típicamente, las formulaciones que comprenden una cantidad eficaz terapéuticamente suministran de aproximadamente 0,1 ug a aproximadamente 1000 ug del polipéptido por administración, más típicamente de aproximadamente 2,5 ug a aproximadamente 100 ug del polipéptido por administración.

Naturalmente, la cantidad del compuesto o compuestos activos en cada composición útil terapéuticamente se puede preparar de tal manera que se obtenga una dosificación adecuada en cualquier dosis unitaria dada del compuesto. Factores tales como la solubilidad, biodisponibilidad, vida media biológica, vía de administración, vida útil del producto, así como otras consideraciones farmacológicas, serán contemplados por un experto en la técnica de preparación de tales formulaciones farmacéuticas, y como tal, pueden ser deseables una variedad de dosificaciones y regímenes de tratamiento.

1. Suministro oral

En ciertas aplicaciones, las composiciones farmacéuticas descritas en la presente memoria se pueden suministrar

mediante administración oral a un animal. Como tales, estas composiciones se pueden formular con un diluyente inerte o con un vehículo comestible asimilable, o se pueden encerrar en cápsulas de gelatina de cubierta dura o blanda, o se pueden comprimir en comprimidos, o se pueden incorporar directamente con el alimento de la dieta.

5 Los compuestos activos se pueden incorporar incluso con excipientes, y se pueden utilizar en forma de comprimidos ingeribles, comprimidos bucales, píldoras, cápsulas, elixires, suspensiones, jarabes, obleas, y similares (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; Patente U.S. N° 5.641.515; Patente U.S. N° 5.580.579, y Patente U.S. N° 5.792.451, cada una específicamente incorporada en la presente memoria como referencia en su totalidad). Los comprimidos, píldoras, pastillas, cápsulas, y similares, también pueden contener lo siguiente: un aglutinante, como 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 se puede añadir un agente edulcorante, tal como sacarosa, lactosa, o sacarina, o un agente saborizante, tal como menta, aceite de gaulteria, o saborizante de cereza. Cuando la forma unitaria de dosificación es una cápsula, puede contener, además de los materiales del tipo anterior, un vehículo líquido. Se pueden presentar diferentes materiales como recubrimientos o para modificar de otra manera la forma física de la unidad de dosificación. Por ejemplo, los comprimidos, pastillas, o cápsulas se pueden recubrir con goma laca, azúcar, o ambos. Un jarabe de elixir puede contener el compuesto activo, sacarosa como agente edulcorante, metil- y propil-parabenos como conservadores, un colorante y un saborizante, tal como saborizante de cereza o de naranja. Por supuesto, cualquier material utilizado en la preparación de cualquier forma unitaria de dosificación debe ser puro farmacéuticamente y sustancialmente no tóxico en las cantidades empleadas. Además, los compuestos activos se pueden incorporar en preparaciones y formulaciones de liberación sostenida.

Para la administración oral, las composiciones de la presente invención se pueden incorporar alternativamente con uno o más excipientes en la forma de un enjuague bucal, dentífrico, comprimido bucal, pulverización oral, o formulación sublingual administrada oralmente. Por ejemplo, un enjuague bucal se puede preparar incorporando el ingrediente activo en la cantidad requerida en un disolvente apropiado, tal como una disolución de borato de sodio (Solución de Dobell). Alternativamente, el ingrediente activo se puede incorporar en una disolución oral, tal como una que contiene borato de sodio, glicerina, y bicarbonato de potasio, o se puede dispersar en un dentífrico, o se puede agregar en una cantidad eficaz terapéuticamente a una composición que puede incluir agua, aglutinantes, abrasivos, agentes saborizantes, agentes espumantes, y humectantes. Alternativamente, las composiciones se pueden configurar en forma de comprimido o disolución que se puede colocar debajo de la lengua o disolverse de otra manera en la boca.

## 2. Suministro inyectable

En ciertas circunstancias, será deseable suministrar las composiciones farmacéuticas descritas en la presente memoria, parenteralmente, intravenosamente, intramuscularmente, intradérmicamente, o incluso intraperitonealmente, como se describe en la Patente U.S. N° 5.543.158; en la Patente U.S. N° 5.641.515, y en la Patente U.S. N° 5.399.363 (cada una incorporada específicamente en la presente memoria como referencia en su totalidad). Las disoluciones de los compuestos activos como base libre o como sales aceptables farmacológicamente, se pueden preparar en agua adecuadamente mezclada con un tensoactivo, tal como hidroxipropilcelulosa. Las dispersiones también se pueden preparar en glicerol, polietilenglicoles líquidos, y mezclas de los mismos, y en aceites. Bajo condiciones ordinarias de almacenamiento y uso, estas preparaciones contienen un conservador para prevenir el crecimiento de microorganismos.

Las formas farmacéuticas adecuadas para uso inyectable incluyen disoluciones o dispersiones acuosas estériles, y polvos estériles para la preparación extemporánea de disoluciones o dispersiones inyectables estériles (Patente U.S. N° 5.466.468, incorporada específicamente en la presente memoria como referencia en su totalidad). En todos los casos, la forma debe ser estéril, y debe ser fluida en la medida en que exista facilidad de inyección. Debe ser estable bajo las condiciones de fabricación y almacenamiento, y se debe conservar contra la acción contaminante de los microorganismos, tales como bacterias y hongos. El vehículo puede ser un disolvente o un 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. Se puede mantener la fluidez apropiada, por ejemplo, mediante el uso de un recubrimiento, tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de la dispersión, y mediante el uso de tensoactivos. La prevención de la acción de los microorganismos se puede facilitar mediante varios agentes antibacterianos y antifúngicos, por ejemplo parabenos, clorobutanol, fenol, ácido sórbico, timerosal, y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares o cloruro de sodio. Se puede conseguir una absorción prolongada de las composiciones inyectables mediante el uso en las composiciones de agentes que retrasan la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Para la administración parenteral en una disolución acuosa, por ejemplo, la disolución debe tamponarse adecuadamente si es necesario, y el diluyente líquido primero se vuelve isotónico con suficiente solución salina o glucosa. Estas disoluciones acuosas particulares son especialmente adecuadas para la administración intravenosa, intramuscular, subcutánea, e intraperitoneal. En relación con esto, expertos en la técnica conocerán un medio acuoso estéril que se puede emplear a la luz de la presente divulgación. Por ejemplo, una dosificación se puede disolver en 1 ml de una disolución isotónica de NaCl, y agregarse a 1000 ml de fluido de hipodermoclasia, o

inyectarse en el sitio de infusión propuesto (véase, por ejemplo, *Remington's Pharmaceutical Sciences*, 15ª Edición, páginas 1035-1038 y 1570-1580). Necesariamente se presentará alguna variación en la dosificación, dependiendo de la condición del sujeto que se está tratando. La persona responsable de la administración, en cualquier caso, determinará la dosis apropiada para el sujeto individual. Por otra parte, para la administración a seres humanos, las preparaciones deben satisfacer con la esterilidad, pirogenicidad, y los estándares generales de seguridad y pureza según lo requieren los estándares de la Oficina de Biológicos de la FDA.

Las disoluciones inyectables estériles se preparan mediante la incorporación de los compuestos activos en la cantidad requerida en el disolvente apropiado con varios de los otros ingredientes enumerados anteriormente, según se requieran, seguido de esterilización por filtración. Generalmente, las dispersiones se preparan mediante la incorporación de varios ingredientes activos esterilizados en un vehículo estéril que contiene el medio de dispersión básico, y los otros ingredientes requeridos de los enumerados anteriormente. En el caso de los polvos estériles para la preparación de disoluciones inyectables estériles, los métodos de preparación preferidos son las técnicas de secado al vacío y secado por congelación, que proporcionan un polvo del ingrediente activo más cualquier ingrediente deseado adicional a partir de una disolución filtrada previamente esterilizada del mismo.

Las composiciones descritas en la presente memoria se pueden formular en una forma neutra o de sal. Las sales aceptables farmacéuticamente, incluyen las sales de adición de ácido (formadas con los grupos amino libres de la proteína), y las cuales se forman con ácidos inorgánicos, tales como, por ejemplo, ácidos clorhídrico o fosfórico, o con ácidos orgánicos, tales como ácido acético, oxálico, tartárico, mandélico, y similares. Las sales formadas con los grupos carboxilo libres también pueden derivarse a partir de bases inorgánicas, tales como, por ejemplo, hidróxidos de sodio, potasio, amonio, calcio, o férrico, y bases orgánicas tales como isopropilamina, trimetilamina, histidina, procaína, y similares. Después de la formulación, las disoluciones se administrarán de una manera compatible con la formulación de dosificación y en una cantidad tal que sea eficaz terapéuticamente. Las formulaciones se administran fácilmente en una variedad de formas de dosificación, tales como disoluciones inyectables, cápsulas de liberación de fármacos, y similares.

Como se emplea en la presente memoria, "vehículo" incluye cualquiera y todos los disolventes, medios de dispersión, vehículos, recubrimientos, diluyentes, agentes antibacterianos y antifúngicos, agentes isotónicos y de retardo de la absorción, tampones, disoluciones de vehículo, suspensiones, coloides, y similares. El uso de tales medios y agentes para sustancias activas farmacéuticas es bien conocido en la técnica. Excepto en la medida en que cualquier medio o agente convencional sea incompatible con el ingrediente activo, se contempla su uso en las composiciones terapéuticas. También se pueden incorporar ingredientes activos complementarios en las composiciones.

La frase "aceptable farmacéuticamente" se refiere a las entidades moleculares y composiciones que no produzcan una reacción alérgica o adversa similar cuando se administran a un ser humano. La preparación de una composición acuosa que contiene una proteína como ingrediente activo se entiende bien en la técnica. Típicamente, tales composiciones se preparan como inyectables, ya sea como disoluciones o suspensiones líquidas; también se pueden preparar formas sólidas adecuadas para disolución en, o suspensión en, un líquido antes de la inyección. La preparación también se puede emulsionar.

### 3. Suministro nasal y bucal

En ciertas realizaciones, las composiciones farmacéuticas se pueden suministrar mediante pulverizaciones intranasales, pulverizaciones bucales, inhalación, y/u otros vehículos de suministro en aerosol. Los métodos para suministrar genes, ácidos nucleicos, y composiciones peptídicas directamente a los pulmones, por ejemplo, a través de aerosoles nasales y bucales, se han descrito, por ejemplo, en la Patente U.S. Nº 5.756.353 y en la Patente U.S. Nº 5.804.212 (cada una incorporada específicamente en la presente memoria como referencia en su totalidad). Asimismo, el suministro de fármacos utilizando resinas de micropartículas intranasales (Takenaga *et al.*, 1998) y compuestos de lisofosfatidil-glicerol (Patente U.S. Nº 5.725.871, incorporada específicamente en la presente memoria como referencia en su totalidad), también son bien conocidos en las técnicas farmacéuticas. De la misma manera, el suministro transmucosal de fármacos en forma de una matriz de soporte de politetrafluoroetileno se describe en la Patente U.S. Nº 5.780.045 (incorporada específicamente en la presente memoria como referencia en su totalidad).

### 4. Suministro mediado por liposomas, nanocápsulas, y micropartículas

En ciertas realizaciones, los inventores contemplan el uso de liposomas, nanocápsulas, micropartículas, microesferas, partículas de lípido, 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 suministrarse bien encapsuladas en una partícula de lípido, un liposoma, una vesícula, una nanoesfera, o bien en una nanopartícula, o similares.

Dichas formulaciones se pueden preferir para la introducción de formulaciones aceptables farmacéuticamente de los ácidos nucleicos o construcciones descritas en la presente memoria. La formación y uso de liposomas son conocidos en general por los expertos en la técnica (véase, por ejemplo, Couvreur *et al.*, 1977; Couvreur, 1988;

Lasic, 1998; que describe el uso de liposomas y nanocápsulas en la terapia antibiótica dirigida a infecciones y enfermedades bacterianas intracelulares). Recientemente, se desarrollaron liposomas con una mejor estabilidad en suero y mejores tiempos medios de circulación (Gabizon y Papahadjopoulos, 1988; Allen y Choun, 1987; Patente U.S. N° 5.741.516, incorporada específicamente en la presente memoria como referencia en su totalidad). Además,  
 5 se han revisado varios métodos de preparaciones de liposomas y tipo liposomas como vehículos de fármacos potenciales (Takakura, 1998; Chandran *et al.*, 1997; Margalit, 1995; Patente U.S 5.567.434; Patente U.S. 5.552.157; Patente U.S. 5.565.213; Patente U.S. 5.738.868 y Patente U.S. 5.795.587, cada una incorporada específicamente en la presente memoria como referencia en su totalidad).

Los liposomas se han utilizado con éxito con varios tipos de células que normalmente son resistentes a la transfección mediante otros procedimientos, incluyendo suspensiones de células T, cultivos de hepatocitos primarios, y células PC 12 (Renneisen *et al.*, 1990; Muller *et al.*, 1990). Además, los liposomas están libres de las limitaciones de la longitud del ADN que son típicas de los sistemas de suministro basados en virus. Los liposomas se han utilizado eficazmente para introducir genes, fármacos (Heath y Martin, 1986; Heath *et al.*, 1986; Balazsovits *et al.*, 1989; Fresta y Puglisi, 1996), agentes radioterapéuticos (Pikul *et al.*, 1987), enzimas (Imaizumi *et al.*, 1990a; Imaizumi *et al.*, 1990b), virus (Faller y Baltimore, 1984), factores de transcripción y efectores alostéricos (Nicolau y Gersonde, 1979), en una variedad de líneas celulares y animales cultivados. Además, se han completado varios estudios clínicos de éxito que examinan la eficacia del suministro de fármacos mediado por liposomas (Lopez-Berestein *et al.*, 1985a; 1985b; Coune, 1988; Sculier *et al.*, 1988). Además, varios estudios sugieren que el uso de liposomas no se asocia con respuestas autoinmunes, toxicidad, o localización gonadal después del suministro sistémico (Mori y Fukatsu, 1992).  
 10  
 15  
 20

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

Los liposomas tienen un parecido con las membranas celulares, y se contemplan para utilizarse en relación con la presente invención como vehículos para las composiciones peptídicas. Son ampliamente adecuados, ya pueden quedar atrapadas tanto las sustancias solubles en agua como las solubles en lípido, es decir, en los espacios acuosos y dentro la propia bicapa, respectivamente. Es posible que los liposomas que lleven fármacos se puedan emplear incluso para el suministro de agentes activos específicos del sitio, mediante la modificación selectiva de la formulación liposomal.  
 30

Además de las enseñanzas de Couvreur *et al.* (1977; 1988), se puede utilizar la siguiente información para generar formulaciones liposomales. Los fosfolípidos pueden formar una variedad de estructuras diferentes de los liposomas cuando se dispersan en agua, dependiendo de la proporción molar de lípido a agua. En bajas proporciones, 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 una baja permeabilidad a las sustancias iónicas y polares, pero a temperaturas elevadas, sufren una transición de fases que altera marcadamente su permeabilidad. La transición de fases implica un cambio desde una estructura ordenada estrechamente empaquetada, conocida como el estado de gel, hasta una estructura menos ordenada, menos empaquetada, conocida como el estado fluido. Esto ocurre a una temperatura de transición de fase característica, y da como resultado un aumento en la permeabilidad a iones, azúcares, y fármacos.  
 35  
 40

Además de la temperatura, la exposición a las 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 esta manera cambios en la permeabilidad. El colesterol inhibe esta penetración de proteínas, aparentemente al empaquetar los fosfolípidos con más fuerza. Se contempla que las formaciones de liposomas más útiles para el suministro de antibióticos e inhibidores, contendrán colesterol.  
 45

La capacidad para atrapar solutos varía entre diferentes tipos de liposomas. Por ejemplo, las MLVs son moderadamente eficientes para atrapar solutos, pero las SUVs son extremadamente ineficientes. Las SUVs ofrecen la ventaja de la homogeneidad y reproducibilidad en la distribución de tamaños, sin embargo, las grandes vesículas unilamelares (LUVs) ofrecen un compromiso entre el tamaño y la eficiencia de atrape. Éstas se preparan mediante evaporación de éter, y son de tres a cuatro veces más eficientes para atrapar el soluto que las MLVs.  
 50

Además de las características de los liposomas, un determinante importante para atrapar compuestos es el de las propiedades fisicoquímicas del compuesto en sí. Los compuestos polares se atrapan 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 la permeación o cuando se rompe la bicapa, pero los compuestos no polares permanecen afiliados a la bicapa, a menos que se altere por la temperatura o la exposición a las lipoproteínas. Ambos tipos muestran máximos índices de flujo a la temperatura de transición de fase.  
 55

Los liposomas interactúan con las células por medio de cuatro mecanismos diferentes: endocitosis por células fagocíticas del sistema reticuloendotelial, tales como macrófagos y neutrófilos; adsorción a la superficie celular, ya

5 sea mediante fuerzas hidrofóbicas o electrostáticas débiles no específicas, o bien mediante interacciones específicas con los componentes de la superficie celular; fusión con la membrana de la célula plasmática mediante la inserción de la bicapa lipídica del liposoma en la membrana plasmática, con la liberación simultánea de contenidos liposomales en el citoplasma; y mediante transferencia de lípidos liposomales a las membranas celulares o subcelulares, o viceversa, sin ninguna asociación de los contenidos liposomales. Con frecuencia es difícil determinar qué mecanismo está operativo, y más de uno puede operar al mismo tiempo.

10 El destino y la disposición de los liposomas inyectados intravenosamente dependen de sus propiedades físicas, tal como su tamaño, fluidez, y carga superficial. Pueden persistir en los tejidos durante horas o días, dependiendo de su composición, y las vidas medias en sangre están en el intervalo de minutos a varias horas. Los liposomas más grandes, tales como las MLVs y LUVs, se absorben rápidamente por las células fagocíticas del sistema reticuloendotelial, pero la fisiología del sistema circulatorio restringe la salida de tales especies grandes en la mayoría de los sitios. Pueden salir solamente en los lugares en donde existan grandes aberturas o poros en el endotelio capilar, tal como los sinusoides del hígado o el bazo. Por tanto, estos órganos son el sitio predominante de absorción. Por otra parte, las SUVs muestran una distribución tisular más amplia, pero todavía son altamente secuestradas en el hígado y el bazo. En general, este comportamiento *in vivo* limita la dirección potencial de los liposomas sólo a los órganos y tejidos accesibles por su gran tamaño. Éstos incluyen la sangre, el hígado, el bazo, la médula ósea, y los órganos linfoides.

15 El direccionamiento en general no es una limitación en términos de la presente invención. Sin embargo, si se desea una dirección específica, existen métodos disponibles para que esto se lleve a cabo. Se pueden utilizar anticuerpos para unirse a la superficie del liposoma, y para dirigir el anticuerpo y sus contenidos del fármaco hacia los receptores antigénicos específicos localizados sobre una superficie de tipo celular particular. También se pueden utilizar determinantes de carbohidratos (componentes de superficie celular de glicoproteína o glicolípido que tiene un papel en el reconocimiento, interacción, y adhesión de célula-célula) como sitios de reconocimiento, debido a que tienen el potencial para dirigir los liposomas hacia tipos particulares de células. En su mayor parte, se contempla que se use la inyección intravenosa de preparaciones liposomales, pero también son concebibles otras vías de administración.

20 Alternativamente, la invención proporciona formulaciones de nanocápsulas aceptables farmacéuticamente de las composiciones de la presente invención. Las nanocápsulas generalmente pueden atrapar los compuestos de una manera estable y reproducible (Henry-Michelland *et al.*, 1987; Quintanar-Guerrero *et al.*, 1998; Douglas *et al.*, 1987). Para evitar los efectos secundarios debidos a la sobrecarga polimérica intracelular, tales partículas ultrafinas (de un tamaño alrededor de 0,1  $\mu\text{m}$ ) deben diseñarse utilizando polímeros capaces de degradarse *in vivo*. Las nanopartículas de polialquil-cianoacrilato biodegradables que satisfacen estos requerimientos se contemplan para su uso en la presente invención. Tales partículas se pueden fabricar fácilmente, como se describe (Couvreur *et al.*, 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.*, 1998; Pinto-Alphandry *et al.*, 1995, y la Patente U.S. N° 5.145.684, específicamente incorporados en la presente memoria como referencia en su totalidad).

35 También se pueden utilizar parches cutáneos, para el suministro transcutáneo.

### Composiciones inmunogénicas

40 En ciertas realizaciones preferidas de la presente invención, se proporcionan composiciones inmunogénicas. Las composiciones inmunogénicas comprenderán generalmente uno o más polipéptidos, tales como aquellos que se discuten anteriormente, en combinación con un inmunoestimulante. Un inmunoestimulante puede ser cualquier sustancia que mejore o potencie una respuesta inmunitaria (mediada por anticuerpo y/o células) a un antígeno exógeno. Ejemplos de inmunoestimulantes incluyen adyuvantes, microesferas biodegradables (por ejemplo, galactida poliláctica) y liposomas (en los que se incorpora el compuesto; véase, por ejemplo, Fullerton, Patente U.S. N° 4.235.877).

45 La preparación de composiciones inmunogénicas se describe en general, por ejemplo, en Powell y Newman, eds., *Vaccine Design* (planteamiento de la subunidad y adyuvante) (1995). Las composiciones farmacéuticas y las composiciones inmunogénicas dentro del alcance de la presente invención también pueden contener otros compuestos, los cuales pueden ser biológicamente activos o inactivos. Por ejemplo, pueden presentarse una o más porciones inmunogénicas de otros antígenos de *M. tuberculosis*, ya sea incorporados en un polipéptido de fusión, o bien como un compuesto separado, dentro de la composición farmacéutica o inmunogénica.

50 Las composiciones inmunogénicas ilustrativas pueden contener un polinucleótido (por ejemplo, ADN) que codifica uno o más de los polipéptidos como se describe anteriormente, de manera que el polipéptido se genera *in situ* (provocando de esta manera una respuesta inmunitaria). Como se indicó anteriormente, el ADN puede estar presente dentro de cualquier variedad de sistemas de suministro conocidos por los expertos ordinarios en la técnica, incluyendo sistemas de expresión de ácido nucleico, bacterias, y sistemas de expresión viral. En la técnica se conocen numerosas técnicas de suministro de genes, tales como las descritas por Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198 (1998), y las referencias citadas en el mismo. Los sistemas de expresión de ácido nucleico apropiados contienen las secuencias de ADN necesarias para la expresión en el paciente (tal como un promotor y una señal de terminación adecuadas). Los sistemas de suministro bacteriano implican la administración de una célula huésped bacteriana (por ejemplo, una cepa de *Mycobacterium*, *Bacillus* o *Lactobacillus*, incluyendo

*Bacillus Calmette-Guerin* o *Lactococcus lactis*), que expresan el polipéptido (por ejemplo, sobre su superficie celular, o que secreta el polipéptido) (véase, por ejemplo, Ferreira *et al.*, *An Acad Bras Cienc* (2005) 77:113-124; y Raha *et al.*, *Appl Microbiol Biotechnol* (2005) PubMedID 15635459). En una realización preferida, el ADN se puede introducir utilizando un sistema de expresión viral (por ejemplo, vaccinia u otro virus de la viruela, retrovirus, o adenovirus), que puede implicar el uso de un virus no patológico (defectuoso), competente de replicación. Los sistemas adecuados se describen, por ejemplo, en Fisher-Hoch *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 86:317-321 (1989); Flexner y *et al.*, *Ann. N. Y. Acad. Sci.* 569: 86-103 (1989); Flexner *et al.*, *Vaccine* 8:17-21 (1990); Patentes U.S. N<sup>os</sup> 4.603.112, 4.769.330, y 5.017.487; WO 89/01973; Patente U.S. N<sup>o</sup> 4.777.127; Patente GB 2.200.651; Patente EP 0.345.242; WO 91/02805; Berkner, *Biotechniques* 6:616-627 (1988); Rosenfeld *et al.*, *Science* 252:431-434 (1991); Kolls *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 91:215-219 (1994); Kass-Eisler *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 90:11498-11502 (1993); Guzman *et al.*, *Circulation* 88:2838-2848 (1993); y Guzman *et al.*, *Cir. Res.* 73:1202-1207 (1993). Las técnicas para incorporar ADN en estos sistemas de expresión son bien conocidas por los expertos ordinarios en la técnica. El ADN también puede estar "desnudo", como se describe, por ejemplo, en Ulmer *et al.*, *Science* 259:1745-1749 (1993), y revisado por Cohen, *Science* 259:1691-1692 (1993). La absorción del ADN desnudo se puede aumentar recubriendo el ADN en perlas biodegradables, que se transportan eficientemente dentro de las células. Será evidente que una composición inmunogénica puede comprender tanto un componente de polinucleótido como de polipéptido. Tal composición inmunogénica puede proporcionar una mejor respuesta inmunitaria.

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

Aunque se puede emplear cualquier vehículo adecuado conocido por los expertos ordinarios en la técnica en las composiciones inmunogénicas de esta invención, el tipo de vehículo variará dependiendo del modo de administración. Las composiciones de la presente invención se pueden formular para cualquier modo de administración apropiado, incluyendo, por ejemplo, la administración tópica, oral, nasal, intravenosa, intracraneal, intraperitoneal, subcutánea, o intramuscular. Para la administración parenteral, tal como inyección subcutánea, el vehículo comprende preferiblemente agua, solución salina, alcohol, una grasa, una cera, o un tampón. Para la administración oral, se puede emplear cualquiera de los vehículos anteriores, o un vehículo sólido, tal como manitol, lactosa, almidón, estearato de magnesio, sacarina de sodio, talco, celulosa, glucosa, sacarosa, y carbonato de magnesio. También se pueden emplear microesferas biodegradables (por ejemplo, polilactato poliglicolato) como vehículos para las composiciones farmacéuticas de esta invención. Las microesferas biodegradables adecuadas se describen, por ejemplo, en las Patentes U.S. N<sup>os</sup> 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 y proteínas descritos en la Patente U.S. N<sup>o</sup> 5.928.647, que son capaces de inducir respuestas de linfocitos T citotóxicos de clase I restringida en un huésped.

Tales composiciones también pueden comprender tampones (por ejemplo, solución salina tamponada neutra o solución salina tamponada con fosfato), carbohidratos (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 de un receptor, agentes de suspensión, agentes espesantes, y/o conservadores. Alternativamente, las composiciones de la presente invención se pueden formular como un liofilizado. Los compuestos también se pueden encapsular dentro de liposomas empleando tecnología bien conocida.

Se puede emplear cualquiera de una variedad de inmunoestimulantes en las composiciones inmunogénicas de esta invención. Por ejemplo, se puede incluir un adyuvante. La mayoría de los adyuvantes contienen una sustancia diseñada para proteger al antígeno del catabolismo rápido, tal como hidróxido de aluminio o aceite mineral, y un estimulante de las respuestas inmunes, tal como lípido A, especies de *Bordetella pertussis* o *Mycobacterium*, o proteínas derivadas de *Mycobacterium*. Por ejemplo, se puede utilizar *M. vaccae* ("pVac") desglícolipidada, deslipidada. Los adyuvantes adecuados están disponibles comercialmente, por ejemplo, como Adyuvante Incompleto y Adyuvante Completo de Freund (Difco Laboratories, Detroit, MI); Adyuvante Merck 65 (Merck and Company, Inc., Rahway, NJ); AS01B, AS02A, AS15, AS-2, y derivados de los mismos (GlaxoSmithKline, Filadelfia, PA); CWS (esqueleto de la pared celular de un bacilo de tuberculosis), TDM (dicorinomicolato de trehalosa), Leif (factor de inicio de elongación de Leishmania), sales de aluminio, tales como gel de hidróxido de aluminio (alumbre) o fosfato de aluminio; sales de calcio, hierro, o zinc; una suspensión insoluble de tirosina acilada; azúcares acilados; polisacáridos catiónicamente o aniónicamente derivados; polifosfocenos; microesferas biodegradables; monofosforil lípido A (MPL®), y quil A (por ejemplo, QS-21). También se pueden utilizar como adyuvantes citoquinas, tales como GM-CSF o interleucina-2, -7, ó -12.

Un adyuvante se refiere a los componentes de 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 del tipo Th-2. Las citoquinas tipo Th1 (por ejemplo, IFN- $\gamma$ , IL-2, e IL-12) tienden a favorecer la inducción de la respuesta inmunitaria mediada por células a un antígeno

administrado, mientras que las citoquinas tipo Th-2 (por ejemplo, IL-4, IL-5, IL-6, IL-10) tienden a favorecer la inducción de respuestas inmunitarias humerales. Los adyuvantes capaces de tener una estimulación preferencial de una respuesta inmunitaria mediada por células Th-1 se describen en WO 94/00153 y WO 95/17209.

5 Dentro de las composiciones inmunogénicas proporcionadas en la presente memoria, la composición adyuvante se diseña preferiblemente para inducir una respuesta inmunitaria predominantemente del tipo Th1. Después de la aplicación de una composición inmunogénica como se proporciona en la presente memoria, un paciente soportará típicamente una respuesta inmunitaria que incluye respuestas tipo Th1 y tipo Th2. Dentro de una realización preferida, en la que una respuesta es predominantemente del tipo Th1, el nivel de citoquinas tipo Th1 aumentará en mayor medida que el nivel de citoquinas tipo Th2. Los niveles de estas citoquinas se pueden evaluar fácilmente empleando los ensayos convencionales. Para una revisión de las familias de citoquinas, véase Janeway *et al.*, *Immunobiology*, 5ª Edición, 2001.

15 Las composiciones de Rv2386c usualmente comprenden uno o más adyuvantes, por ejemplo, AS01B (monofosforil lípido A 3-des-O-acilado (3D-MPL®), y QS21 en una formulación de liposomas; véase la Publicación de Patente U.S. N° 2003/0143240); AS02A (3D-MPL® y QS21, y una emulsión de aceite en agua; véase, Bojang *et al.*, *Lancet* (2001) 358:1927); ENHANZYN® (Detox); 3D-MPL®; saponinas, incluyendo Quil A y sus componentes, por ejemplo, QS21 y miméticos de saponina; CWS (esqueleto de la pared celular de un bacilo de tubérculo); TDM (dicorinomicolato de trehalosa); 4-fosfatos de aminoalquil glucosaminida (AGPs); oligonucleopéptidos inmunoestimulantes, por ejemplo, CPG; Leif (factor de inicio de elongación de Leishmania); y derivados de los mismos. En una realización preferida, se administra un polipéptido de Rv2386c con uno o más adyuvantes seleccionados del grupo que consiste en 3D-MPL® y QS21 en una formulación de liposomas, por ejemplo, AS01B y 20 3D-MPL® y QS21, y una emulsión de aceite en agua (por ejemplo, AS02A). Los sistemas adyuvantes AS01B y AS02A se describen adicionalmente en Pichyangkul, *et al.*, *Vaccine* (2004) 22:3831-40.

25 Cuando se suministra el antígeno Rv2386c como un ácido nucleico, se puede suministrar, por ejemplo, en un vector viral (es decir, un vector de adenovirus), o en una célula huésped de bacteria mutante (es decir, una célula huésped mutante, avirulenta, de *Mycobacterium*, *Lactobacillus* o *Bacillus*, incluyendo *Bacillus Calmette-Guerin* (BCG), y *Lactococcus lactis*).

30 Los adyuvantes preferidos para utilizarse en provocar una respuesta predominantemente tipo Th1 incluyen, por ejemplo, una combinación de monofosforil lípido A (MPL®), preferiblemente monofosforil lípido A 3-O-desacilado (3D-MPL®), opcionalmente con una sal de aluminio (véase, por ejemplo, Ribi *et al.*, 1986, *Immunology and Immunopharmacology of Bacterial Endotoxins*, Plenum Publ. Corp., NY, páginas 407-419; Patente GB 2122204B; Patente GB 2220211; y Patente US N° 4.912.094). Una forma preferida de 3D-MPL® está en la forma de una emulsión que tiene un tamaño de partículas pequeño menor de 0,2 µm de diámetro, y su método de elaboración se describe en WO 94/21292. Se han descrito formulaciones acuosas que comprenden monofosforil lípido A y un tensoactivo en WO 98/43670. Ejemplos de adyuvantes preferidos incluyen AS01B (MPL® y QS21 en una 35 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, disponible en GlaxoSmithKline. Los adyuvantes MPL® están disponibles en GlaxoSmithKline (véanse las Patentes US N°s 4.436.727; 4.877.611; 4.866.034 y 4.912.094).

40 Los oligonucleótidos que contienen CpG (en los que el dinucleótido CpG no está metilado) también inducen una respuesta predominantemente Th1. CpG es una abreviatura para los motivos de los dinucleótidos de citosina-guanosina presentes en el ADN. Tales oligonucleótidos son bien conocidos y se describen, por ejemplo, en WO 96/02555, WO 99/33488, y en las Patentes U.S. N°s 6.008.200 y 5.856.462. También se describen secuencias de ADN inmunoestimulantes, por ejemplo, por Sato *et al.*, *Science* 273:352 (1996). CpG cuando se formula en composiciones inmunogénicas, se administra generalmente en solución libre junto con el antígeno libre 45 (WO96/02555; McCluskie y Davis, *supra*), o covalentemente conjugado con un antígeno (WO 98/16247), o se formula con un vehículo, tal como hidróxido de aluminio ((antígeno superficial de hepatitis), Davis *et al.*, *supra*; Brazolot-Millan *et al.*, *Proc. Natl. Acad. Sci. EE.UU.*, 1998, 95(26), 15553-8). CpG se conoce en la técnica como un adyuvante que se puede administrar tanto mediante la vía sistémica como por vía mucosal (WO 96/02555, Patente EP 468520, Davis *et al.*, *J. Immunol.*, 1998, 160(2):870-876; McCluskie y Davis, *J. Immunol.*, 1998, 161(9):4463-6).

50 Otro adyuvante preferido es una saponina o miméticos o derivados de saponina, tales como Quil A, preferiblemente QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), que se puede utilizar en solitario o en combinación con otros adyuvantes. Por ejemplo, un sistema mejorado implica la combinación de un monofosforil lípido A (MPL®) y un derivado de saponina, tal como la combinación de QS21 y 3D-MPL®, como se describe en WO 94/00153, o una composición menos reactogénica, en la que el QS21 se aplaca con colesterol, como se describe en WO 96/33739. 55 Otras formulaciones preferidas comprenden una emulsión de aceite en agua y tocoferol. Una formulación de adyuvante particularmente potente que implica QS21, 3D-MPL®, y tocoferol en una emulsión de aceite en agua, se describe en WO 95/17210. Los adyuvantes de saponina adicionales de uso en la presente invención incluyen QS7 (descrito en WO 96/33739 y WO 96/11711) y QS17 (descrito en la Patente U.S. N° 5.057.540 y en la Patente EP 0 362 279 B1).

Alternativamente, las formulaciones de saponina se pueden combinar con vehículos de vacuna compuestos de quitosano u otros polímeros policatiónicos, partículas de polilactida y de polilactida-co-glicolida, matriz de polímero basada en poli-N-acetil-glucosamina, partículas compuestas de polisacáridos o de polisacáridos químicamente modificados, partículas basadas en liposomas y en lípidos, partículas compuestas de monoésteres de glicerol, etc. Las saponinas también se pueden formular en presencia de colesterol para formar estructuras de partículas, tales como liposomas o ISCOM®s. Además, las saponinas se pueden formular junto con un éter o éster de polioxietileno, ya sea en una disolución o suspensión que no está en partículas, o bien en una estructura en partículas, tal como un liposoma paucilamelar o ISCOM®. Las saponinas también se pueden formular con excipientes, tales como CARBOPOL®, para aumentar la viscosidad, o se pueden formular 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 WO 94/00153, o una composición menos reactogénica, en donde el QS21 se aplaca con liposomas que contienen colesterol, como se describe en WO 96/33739. Otras formulaciones adecuadas comprenden una emulsión de aceite en agua y tocoferol. Otra formulación de adyuvante adecuada que emplea QS21, adyuvante 3D-MPL® y tocoferol en una emulsión de aceite en agua, se describe en WO 95/17210.

Otro sistema adyuvante mejorado incluye la combinación de un oligonucleótido que contiene CpG y un derivado de saponina, particularmente la combinación de CpG y QS21 como se describe en WO 00/09159. De una manera adecuada, 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 de adyuvantes SBAS (SmithKline Beecham, Rixensart, Bélgica), Detox (Corixa), RC-529 (Corixa) y otros aminoalquil-glucosaminida 4-fosfatos (AGPs), tales como los descritos en las solicitudes de Patentes de EE.UU., en tramitación N<sup>os</sup> de Serie 08/853.826 y 09/074.720, cuyas divulgaciones se incorporan en la presente memoria como referencia en su totalidad, y adyuvantes de éter de polioxietileno, tales como los descritos en WO 99/52549A1. SmithKline Beecham y Corixa Corporation son ahora parte de GlaxoSmithKline.

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



en donde, n es de 1-50, A es un enlace o -C(O)-, R es alquilo de 1-50 átomos de carbono o fenil-alquilo de 1-50 átomos de carbono.

Otro adyuvante de interés es la cadena de toxina b shiga, usada, por ejemplo, como se describe en WO2005/112991.

Una realización de la presente invención consiste en una composición inmunogénica que comprende un éter de polioxietileno de la fórmula general (I), en donde n está entre 1 y 50, preferiblemente 4-24, lo más preferiblemente 9; el componente R es alquilo de 1-50 átomos de carbono, preferiblemente alquilo de 4-20 átomos de carbono, y lo más preferiblemente alquilo de 12 átomos de carbono, y A es un enlace. La concentración de los éteres de polioxietileno deberá estar en el intervalo de 0,1-20%, preferiblemente de 0,1-10%, y lo más preferiblemente en el intervalo de 0,1-1%. Los éteres de polioxietileno preferidos se seleccionan a partir del siguiente grupo: polioxietileno 9-lauril éter, polioxietileno 9-estearil éter, polioxietileno 8-estearil éter, polioxietileno 4-lauril éter, polioxietileno 35-lauril éter, y polioxietileno 23-lauril éter. Los éteres de polioxietileno, tales como el polioxietileno lauril éter, se describen en Merck Index (12<sup>a</sup> edición: entrada 7717). Estas moléculas adyuvantes se describen WO 99/52549.

Cualquier composición inmunogénica proporcionada en la presente memoria se puede preparar empleando métodos 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 la presente memoria se pueden administrar como parte de una formulación de liberación sostenida (es decir, una formulación tal como una cápsula, esponja, o gel (compuesto de polisacáridos, por ejemplo), que efectúe una liberación lenta del compuesto después de la administración). Tales formulaciones se pueden preparar generalmente empleando tecnología bien conocida (véase, por ejemplo, Coombes *et al.*, *Vaccine* 14:1429-1438 (1996)), y se administran, por ejemplo, mediante implantación oral, rectal, o subcutáneo, o mediante implantación en el sitio objetivo deseado. Las formulaciones de liberación sostenida pueden contener un polipéptido, polinucleótido, o anticuerpo dispersado en una matriz de vehículo, y/o pueden estar contenidos dentro de un depósito rodeado por una membrana que controla la velocidad.

Los vehículos para su uso dentro de tales formulaciones son biocompatibles, y también pueden ser biodegradables; preferiblemente la formulación proporciona un nivel relativamente constante de liberación del componente activo. Tales vehículos incluyen micropartículas de poli-(láctido-co-glicólido), poliacrilato, látex, almidón, celulosa, dextrano, y similares. Otros vehículos de liberación sostenida incluyen biovectores supramoleculares, que comprenden un núcleo hidrofílico no líquido (por ejemplo, un polisacárido u oligosacárido reticulado), y opcionalmente, una capa externa que comprende un compuesto anfifílico, tal como un fosfolípido (véase, por ejemplo, la Patente U.S. N<sup>o</sup>

5.151.254, y las Solicitudes PCT WO 94/20078, WO/94/23701 y WO 96/06638). La cantidad de compuesto activo contenida dentro de una formulación de liberación sostenida depende del sitio de implantación, de la velocidad y duración esperada de la liberación.

5 Se puede emplear cualquiera de una variedad de vehículos de suministro dentro de las composiciones farmacéuticas y composiciones inmunogénicas, para facilitar la producción de una respuesta inmunitaria específica del antígeno. Los vehículos de suministro incluyen células presentadoras de antígeno (APCs), tales como células dendríticas, macrófagos, células-B, monocitos, y otras células que se puedan modificar genéticamente para ser APCs eficaces. Estas células pueden, pero no necesitan, modificarse genéticamente para aumentar la capacidad para presentar el antígeno, para mejorar la activación y/o el mantenimiento de la respuesta de células T, y/o para ser compatibles inmunológicamente con el receptor (es decir, haplotipo HLA compatible). Las APCs se pueden aislar generalmente a partir de cualquiera de una variedad de fluidos y órganos biológicos, y pueden ser células autólogas, alogénicas, singénicas, o xenogénicas.

15 Ciertas realizaciones preferidas de la presente invención utilizan células dendríticas o progenitoras de las mismas como células presentadoras de antígeno. Las células dendríticas son APCs altamente potentes (Banchereau y Steinman, *Nature* 392:245-251 (1998)), y se ha demostrado que son eficaces como un adyuvante fisiológico para provocar una inmunidad profiláctica o terapéutica (véase Timmerman y Levy, *Ann. Rev. Med.* 50:507-529 (1999)). En general, las células dendríticas se pueden identificar basándose en su forma típica (estrellada *in situ*, con procesos citoplásmicos marcados (dendritas) visibles *in vitro*), su capacidad para absorber, procesar, y presentar antígenos con una alta eficacia, y su capacidad para activar las respuestas de células T sin tratamiento previo. Desde luego, las células dendríticas se pueden modificar genéticamente para expresar receptores o ligandos de superficie celular específicos que comúnmente no se encuentran en células dendríticas *in vivo* o *ex vivo*, y la presente invención contempla tales células dendríticas modificadas. Como una alternativa a las células dendríticas, se pueden utilizar células dendríticas cargadas con antígeno en vesículas secretadas (denominadas exosomas) dentro de una composición inmunogénica (véase Zitvogel *et al.*, *Nature Med.* 4:594-600 (1998)).

25 Las células dendríticas y las progenitoras se pueden obtener a partir de sangre periférica, médula ósea, ganglios linfáticos, bazo, piel, sangre de cordón umbilical, o cualquier otro tejido o fluido adecuado. Por ejemplo, las células dendríticas se pueden diferenciar *ex vivo* mediante la adición de una combinación de citoquinas, tales como GM-CSF, IL-4, IL-13, y/o TNF $\alpha$  a los cultivos de monocitos recogidos de sangre periférica. Alternativamente, las células positivas para CD34 recogidas de sangre periférica, sangre de cordón umbilical, o médula ósea, se pueden diferenciar en células dendríticas mediante la adición al medio de cultivo de combinaciones de GM-CSF, IL-3, TNF $\alpha$ , ligando CD40, LPS, ligando fIt3, y/u otro compuesto o compuestos que inducen la diferenciación, maduración, y proliferación de células dendríticas.

35 Las células dendríticas se clasifican convenientemente como células "inmaduras" y "maduras", lo cual permite una manera simple de discriminar entre dos fenotipos bien caracterizados. Sin embargo, esta nomenclatura no debe interpretarse como que excluye todas las posibles etapas intermedias de diferenciación. Las células dendríticas inmaduras se caracterizan como APC con una alta capacidad para absorber y procesar el antígeno, lo cual se correlaciona con la alta expresión del receptor Fc $\gamma$  y el receptor de manosa. El fenotipo maduro se caracteriza normalmente por una expresión más baja de estos marcadores, pero una alta expresión de moléculas de superficie celular responsables de la activación de células T, tales como MHC clase I y clase II, moléculas de adhesión (por ejemplo, CD54 y CD11), y moléculas co-estimuladoras (por ejemplo, CD40, CD80, CD86, y 4-1BB).

45 Las APCs se pueden transfectar generalmente con un polinucleótido que codifica una proteína (o una porción u otra variante de la misma), de tal manera que el polipéptido se expresa sobre la superficie celular. Tal transfección puede tener lugar *ex vivo*, y entonces se puede utilizar una composición farmacéutica o una composición inmunogénica que comprenda tales células transfectadas, como se describe en la presente memoria. Alternativamente, se puede administrar un vehículo de suministro de genes que se dirige a una célula dendrítica u otra célula presentadora de antígeno a un paciente, dando como resultado que se presente la transfección *in vivo*. La transfección *in vivo* y *ex vivo* de células dendríticas, por ejemplo, se puede llevar a cabo generalmente empleando cualquier método conocido en la técnica, tales como los descritos en WO 97/24447, o el planteamiento de la pistola de genes descrito por Mahvi *et al.*, *Immunology and Cell Biology* 75:456-460 (1997). La carga de antígeno de células dendríticas se puede lograr incubando células dendríticas o células progenitoras con el polipéptido, ADN (desnudo o dentro de un vector de plásmido), o ARN; o con bacterias o virus recombinantes que expresan el antígeno (por ejemplo, vectores de vaccinia, virus de la viruela aviar, adenovirus, o lentivirus). Antes de cargarse, el polipéptido se puede conjugar covalentemente con un compañero inmunológico que proporciona ayuda a las células T (por ejemplo, una molécula transportadora). Alternativamente, una célula dendrítica se puede impulsar con un compañero inmunológico no conjugado, por separado, o en presencia del polipéptido.

60 Las composiciones inmunogénicas y las composiciones farmacéuticas se pueden presentar en recipientes de dosis unitaria o de múltiples dosis, tales como ampollas o viales sellados. Estos recipientes preferiblemente se sellan herméticamente para conservar la esterilidad de la formulación hasta su uso. En general, las formulaciones se pueden almacenar como suspensiones, disoluciones, o emulsiones en vehículos oleosos o acuosos. Alternativamente, una composición inmunogénica o una composición farmacéutica se puede almacenar en un estado liofilizado, que solamente requiere la adición de un vehículo líquido estéril inmediatamente antes de usarse.

En algunas realizaciones, un “cebado” o primera administración de un polipéptido Rv2386c (incluyendo variantes, fragmentos inmunogénicos, o proteínas de fusión), o un polinucleótido que codifica dicho polipéptido, es seguido por una o más administraciones “potenciadoras” o posteriores de un polipéptido Rv2386c (incluyendo variantes, fragmentos inmunogénicos, o proteínas de fusión), o un polinucleótido que codifica dicho polipéptido (método de “cebado y refuerzo”). Por ejemplo, una primera administración con un polipéptido Rv2386c (incluyendo 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 Rv2386c (incluyendo variantes, fragmentos inmunogénicos, o proteínas de fusión), o un polinucleótido que codifica dicho polipéptido.

En una realización, una primera administración con un polipéptido o polinucleótido Rv2386c, es seguida por una o más administraciones posteriores de un polipéptido Rv2386c. En una realización, una primera administración con un polipéptido o polinucleótido Rv2386c, es seguida por una o más administraciones posteriores de un polinucleótido Rv2386c. Usualmente la primera administración o “cebado” y la segunda administración o “refuerzo” se dan con aproximadamente 2-12 semanas de separación, o con hasta 4-6 meses de separación. Las posteriores administraciones de “refuerzo” se dan con aproximadamente 6 meses de separación, o con una diferencia de 1, 2, 3, 4 ó 5 años. El tratamiento de refuerzo convencional (por ejemplo, una administración de cebado de proteína seguida por una administración de refuerzo de proteína) también puede ser útil para prevenir o tratar la tuberculosis (por ejemplo prevenir o tratar la tuberculosis latente, en particular prevenir o retrasar la reactivación de la tuberculosis).

**Ejemplos**

Los siguientes ejemplos se proporcionan sólo a modo de ilustración, y no a modo de limitación. Los expertos en la técnica reconocerán fácilmente una variedad de parámetros no críticos que se podrían cambiar o modificar para proporcionar resultados esencialmente similares.

**Ejemplo 1 – Identificación de Rv2386c como un objetivo de vacuna contra la TB latente**

El gen Rv2386c, también conocido como MbtI, codifica para una proteína que está implicada en la biogénesis de las micobactinas sideróforas que contienen hidroxifeniloxazolina, donde convierte el corismato a salicilato, la unidad de iniciación para la síntesis del sideróforo de micobacterina.

Rv2386c se seleccionó basándose en un análisis genómico completo de los genes de *Mycobacterium tuberculosis* asociados con el mantenimiento de la fase latente y la infectividad, como en Murphy y Brown *BMC. Infect. Dis.* 2007 7:84-99. Se priorizaron los objetivos potenciales del gen de la fase latente en *Mycobacterium tuberculosis* a través de un meta-análisis bioinformático de conjuntos de datos publicados de microarrays de ADN de todo el genoma de la expresión genética bacteriana bajo condiciones latentes simuladas. La localización subcelular de las proteínas de *M. tuberculosis* codificadas por los 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 muy variadas, de modo que se desarrolló un sistema de calificación de cero a cinco para normalizar estos datos, basándose en dos criterios: 1) la relevancia de las condiciones experimentales para el estado latente, y 2) el orden de clasificación de la expresión. La máxima calificación para un conjunto de datos experimentales particular se ajustó basándose en la relevancia potencial para la presentación clínica de infecciones por *M. tuberculosis* en fase latente. La Tabla 1 muestra los conjuntos de datos recopilados para la Etapa 1, junto con las máximas puntuaciones ajustadas para cada conjunto de datos. Se obtuvieron conjuntos de datos adicionales sobre la esencialidad de los genes para el crecimiento a partir de estudios publicados utilizando experimentos de supresión genética basados en transposones (TraSH). Los genes que no tuvieron efecto alguno sobre el crecimiento recibieron una puntuación de cero.

Tabla 1 - Fuentes, modelos experimentales, y criterios de puntuación para la expresión genética del microarray de ADN de *M. tuberculosis* y la supresión genética en todo el genoma (esencialidad de la fase de crecimiento).

Referencia	Modelo experimental	Punto del tiempo: Puntuación máxima <sup>a</sup>
Betts JC et al. <i>Mol. Microbiol.</i> 2002 43:717-731	Inanición bajo O <sub>2</sub> controlado	96h: 3 24h: 2 4h: 1
Hampshire T et al. <i>Tuberculosis.(Edinb.)</i> 2004 84:228-238	Agotamiento de nutrientes bajo O <sub>2</sub> controlado	62 y 75d: 5 49d: 4 18d: 2

Referencia	Modelo experimental	Punto del tiempo: Puntuación máxima <sup>a</sup>
Muttucumaru DG et al. <i>Tuberculosis.(Edinb.)</i> 2004 84:239-246	Modelo de hipoxia de Wayne <sup>#</sup>	14d (NRP-2): 4 7d (NRP-1): 2
Voskuil MI et al. <i>Tuberculosis.(Edinb.)</i> 2004 84:218-227	Modelo de hipoxia de Wayne <sup>#</sup>	30 y 80d: 5 14 y 20d: 4 10 y 12d: 3 6 y 8d: 2
Schnappinger D et al. <i>J. Exp. Med.</i> 2003 198:693-704	Infección de macrófagos de ratón, +/- $\gamma$ -INF	24 y 48h: 5
Karakousis PC et al. <i>J. Exp. Med.</i> 2004 200:647-657	Implante subcutáneo de fibra hueca en ratones	10d: 3
Talaat AM et al. <i>Proc. Natl. Acad. Sci. EE.UU.</i> 2004, 101:4602-4607	Infección de ratones. MTB recogida de pulmón <sup>b</sup>	28d: 3
Sasseti CM et al. <i>Mol. Microbiol.</i> 2003 48:77-84	Bibliotecas mutadas TraSH cultivadas en un medio sólido	14d:5
Rengarajan J et al. <i>Proc. Natl. Acad. Sci. EE.UU.</i> 2005, 102:8327-8332	Infección de macrófagos de ratón, +/- $\gamma$ -INF con bibliotecas mutadas TraSH de <i>M. tuberculosis</i>	7d:5
Sasseti CM et al. <i>Proc. Natl. Acad. Sci. EE.UU.</i> 2003 100:12989-12994	Ratones C57BL/6J infectados con Bibliotecas mutadas TraSH de <i>M. tuberculosis</i>	7, 14, 28 y 56d:5

<sup>a</sup> Puntuación máxima basándose en la relevancia como un modelo de latencia; h = hora; d = día.

<sup>b</sup> Proporción de *M. tuberculosis* a partir de pulmón de Balb/c a MTB en un cultivo aireado durante 28 días.

<sup>#</sup> Wayne LG y Hayes LG *Infect. Immun.* 1996 64:2062-2069.

5 Etapa 2 - Al aplicar el segundo criterio, el orden de clasificación de expresión genética, las puntuaciones de los genes a partir de cada conjunto de datos, se ordenó de la más alta a la más baja, basándose en la proporción de expresión (expresión doble en la condición experimental frente a las células en cultivo líquido en fase logarítmica). El gen de puntuación más alta recibió la puntuación máxima para ese conjunto de datos particular (enumerados en la columna 3 de la Tabla 1. (por ejemplo, 5, 4 ..., 1 punto)). La puntuación se redujo en 0,005 puntos para cada gen en orden hasta cero, o hasta que se alcanzó el final del conjunto de datos. Por tanto, cuando la puntuación máxima fue de 4 puntos, el gen clasificado como 100<sup>o</sup> recibiría una puntuación de 3.500. Para una puntuación máxima de 5 puntos, 1000 genes o el 25% del genoma de *M. tuberculosis* recibió una puntuación. Para los experimentos donde se recolectaron datos de múltiples puntos de tiempo, la puntuación máxima a través de todos los puntos de tiempo se utilizó como la puntuación final.

15 En la Etapa 3, las puntuaciones para cada gen en cada una de las condiciones experimentales se recolectaron en una base de datos Microsoft Access. Se agregaron campos de referencia para facilitar la priorización, tales como Refseq ID (Identificación de Secuencia de Referencia), Función Genbank, Nota Genbank, Clasificación de

Tuberculista, y los vínculos de KEGG y Sanger Center. Mediante la combinación de los datos de diferentes estudios y fuentes, se alcanzó una opinión consensuada sobre los genes y las vías particulares más críticas para la supervivencia en el estado latente.

5 En la Etapa 4, se derivó una lista priorizada de dianas terapéuticas utilizando los 400 genes de puntuación superior (~10% del genoma) complementados mediante análisis computacionales y manuales de expertos de vías bioquímicas, enzimología, tratabilidad del fármaco, homología con los genes humanos, y otros conocimientos previos. La gran mayoría de los genes de alta puntuación provienen del subconjunto donde se cruzan dos o tres grupos.

10 En la Etapa 5, se llevó a cabo en el genoma completo la identificación de la localización subcelular de las proteínas de *M. tuberculosis* codificadas por los genes. La heurística empleada para la predicción de la proteína de membrana se describe en Chalker et al. *J. Bacteriol.* 2001 183:1259-1268. Se generaron los perfiles de hidropatía promedio (H) (von Heijne G *J. Mol. Biol.* 1992 225:487-494) utilizando valores de hidropatía GES (Engelman DM et al., *Annu. Rev. Biophys. Biophys. Chem.* 1986 15:321-353) ponderados utilizando una ventana trapezoidal. Empleando un proceso similar a las etapas iniciales del algoritmo TopPred II (Claros MG et al. *Comput. Appl. Biosci.* 1994 10:685-686), se predijeron los segmentos transmembrana helicoidales (TMS) para cada secuencia peptídica, mediante la selección de 19 aminoácidos centrados en el valor H más alto (MaxH), enmascarándolos de una consideración adicional, y repitiendo el proceso hasta que ya no quedaron picos con una H >0,5. Las localizaciones subcelulares se asignaron basándose en el valor pico MaxH, en el número de segmentos con una H >1,0, y en la distribución y valores H pico de los supuestos TMS. Se seleccionó un punto de corte de MaxH de 1,15 para maximizar la discriminación entre dos conjuntos de datos de prueba de la versión 34 SwissProtein que contienen proteínas transmembrana y citoplásmicas, respectivamente (Boyd D et al. *Protein Sci.* 1998 7:201-205). Las proteínas con una MaxH <1,15 se clasificaron como citoplásmicas, mientras que aquellas con una MaxH >1,15 y al menos tres posibles TMS, se clasificaron como proteínas de membrana. Las proteínas ancladas se definieron por tener exactamente dos TMS, uno empezando antes del aminoácido (aa) 35 y uno con una H >1,15 con el otro teniendo una H no inferior de 0,5. Se utilizó específicamente SignalP con ajustes Gram-positivas para *M. bacterium*, para identificar las proteínas secretadas entre aquellas clasificadas como citoplásmicas o "desconocidas" en el análisis heurístico (Nielsen H et al. *Protein Eng.* 1997 10:1-6).

Rv2386c se clasificó muy alto como un antígeno de vacuna según varios criterios:

30 (i) Rv2386c se regula al alza sistemáticamente a través de todos los modelos de latencia. Entre todo el conjunto de 3999 genes calificados en el meta-análisis, Rv2386c se clasificó como el 120º como uno del 10% de los principales genes sobre-expresados a través de todos los modelos de latencia. La puntuación regulada al alza para Rv2386c fue de 13,165, la cual se comparó favorablemente con la puntuación del gen superior de 22,28. Rv2386c se reguló a la baja mínimamente, obteniéndose una puntuación ligeramente inferior a 0 (-2,316 comparándose con -18,13 para el gen regulado por disminución más amplia).

35 (ii) La eliminación del hierro de los macrófagos es crucial para la supervivencia e infectividad de *M. tuberculosis*. Las moléculas pequeñas, denominadas sideróforos, funcionan en las bacterias para enlazarse al hierro para la absorción intracelular; MbtI cataliza la primera etapa en la formación del único sideróforo de *M. tuberculosis*, micobactina T, cuya biosíntesis depende de la producción de salicilato a partir del corismato (Harrison et al., *J. Bacteriol.* 2006 188:6081-6091). La expresión de MbtI se induce durante la infección por *M. tuberculosis* de macrófagos THP-1 humanos (Gold et al. *Mol. Microbiol.* 2001 42:851-865). La estructura tridimensional de la proteína está disponible, y se ha estudiado bien el mecanismo de la acción enzimática de MbtI (Zwahlen et al. *Biochemistry* 2007 46:954-964). Rv2386c se clasificó alto en la puntuación de esencialidad, y es requerido por *M. tuberculosis* para la supervivencia en modelos de crecimiento *in vitro* (puntuación de 2,6 de una posible puntuación de 5).

45 (iii) La localización subcelular predijo que la proteína Rv2386c se secreta y, por tanto, tiene una exposición extracelular significativa, indicando su idoneidad como un objetivo de vacuna.

Ejemplo 2- Identificación del epítipo Rv2386c

Método

La predicción del epítipo de células T se basó en los siguientes planteamientos:

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 Brusic,V. (2005) "MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides" <i>Nucleic Acids</i>

Predicción	Nombre	URL/Referencias
		Res. 33, W172 - W179.
	SVMHC	sitio web: <a href="http://www-bs.informatik.uni-tuebingen.de/SVMHC">www-bs.informatik.uni-tuebingen.de/SVMHC</a> "Prediction of MHC class I binding peptides, using SVMHC." Pierre Dönnes y Arne Elofsson en: <i>BMC Bioinformatics</i> 2002 3: 25
CD4	ProPred	sitio web: <a href="http://www.imtech.res.in/raghava/propred/">www.imtech.res.in/raghava/propred/</a> Singh,H. y Raghava,G.P.S.(2001) "ProPred: Prediction of HLA-DR binding sites." <i>Bioinformatics</i> ,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." <i>Methods</i> 34 : 468-75
CD8	nHLA	sitio web: <a href="http://www.imtech.res.in/raghava/nhlapred/">www.imtech.res.in/raghava/nhlapred/</a> Bhasin M. y Raghava G P S (2006) "A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes"; <i>J. Biosci.</i> 32:31-42
	NetCTL	sitio web: <a href="http://www.cbs.dtu.dk/services/NetCTL/">www.cbs.dtu.dk/services/NetCTL/</a> "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. <i>European Journal of Immunology.</i> 35(8):2295-303. 2005
	Epijen	sitio web: <a href="http://www.jenner.ac.uk/EpiJen/">www.jenner.ac.uk/EpiJen/</a> Doytchinova, I. A., P. Guan, D. R. Flower. "EpiJen: a server for multi-step T cell epitope prediction." <i>BMC Bioinformatics</i> , 2006, 7, 131.
	Syfpeithi	sitio web: <a href="http://www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm">www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm</a> Hans-Georg Rammensee, Jutta Bachmann, Niels Nikolaus Emmerich, Oskar Alexander Bachor, Stefan Stevanovic: "SYFPEITHI: database for MHC ligands and peptide motifs." <i>Immunogenetics</i> (1999) 50: 213-219
	PredTAP	sitio web: <a href="http://antigen.i2r.a-star.edu.sg/predTAP/">antigen.i2r.a-star.edu.sg/predTAP/</a> 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." <i>Immunome Res.</i> 2(1), 3.
	PAPROC	sitio web: <a href="http://www.paproc2.de/paproc1/paproc1.html">www.paproc2.de/paproc1/paproc1.html</a> C. Kuttler, A.K. Nussbaum, T.P. Dick, H.-G. Rammensee, H. Schild, K.P. Haderler, "An algorithm for the prediction of proteasomal cleavages", <i>J. Mol. Biol.</i> 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", <i>Immunogenetics</i> 53 (2001), 87-94.

ES 2 685 498 T3

Resultados

Tabla 2

Supuestos epítomos de Rv2386c de células T CD4+ humanas

Nº del supuesto epítomo de CD4	Posición de Aminoácido	Secuencia del epítomo	SEQ ID No:	Alelo HLA
1	54	WVLAAGVQA	SEQ ID No: 30	DRB1_0401
2	55	VLAAGVQAM	SEQ ID No: 31	DRB1_0301, DRB1_1301
3	73	VIRDGVTRR	SEQ ID No: 32	DRB1_0301, DRB1_0401
4	125	LAPHTPLAR	SEQ ID No: 33	DRB1_1301
5	148	IRLFDAGIR	SEQ ID No: 34	DRB1_1101, DRB1_1301, DRB1_1501
6	155	IRHREAIDR	SEQ ID No: 35	DRB1_0801, DRB1_1301
7	164	LLATGVREV	SEQ ID No: 36	DRB1_1301
8	187	FRRRVAVAV	SEQ ID No: 37	DRB1_0101, DRB1_0801
9	216	FAIDFPLTY	SEQ ID No: 38	DRB1_0301, DRB1_0401
10	220	FPLTYRLGR	SEQ ID No: 39	DRB1_1101
11	224	YRLGRRHNT	SEQ ID No: 40	DRB1_0801, DRB1_1101, DRB1_1501
12	234	VRSFLLQLG	SEQ ID No: 41	DRB1_1301
13	238	LLQLGGIRA	SEQ ID No: 42	DRB1_1501
14	253	LVTAVRADG	SEQ ID No: 43	DRB1_1301
15	257	VRADGVVIT	SEQ ID No: 44	DRB1_0301, DRB1_0401

ES 2 685 498 T3

Nº del supuesto epitopo de CD4	Posición de Amino-ácido	Secuencia del epitopo	SEQ ID No:	Alelo HLA
16	262	VVITEPLAG	SEQ ID No: 45	DRB1_0301, DRB1_0401, DRB1_1301, DRB1_1501
17	295	VEHAISVR	SEQ ID No: 46	DRB1_1301
18	321	IDFMTVRER	SEQ ID No: 47	DRB1_1301
19	375	FRLDECPRG	SEQ ID No: 48	DRB1_0301, DRB1_0401
20	384	LYSGAVVML	SEQ ID No: 49	DRB1_0301
21	385	YSGAVVMLS	SEQ ID No: 50	DRB1_0401, DRB1_1101
22	389	VVMLSADGG	SEQ ID No: 51	DRB1_0401, DRB1_1301
23	415	WLRAGAGII	SEQ ID No: 52	DRB1_0101

Tabla 3

Supuestos epitopos de Rv2386c de células T CD8+ humanas

Nº del supuesto epitopo de CD4	Posición de Amino-ácido	Secuencia de epitopo	SEQ ID No:	Alelo HLA
1	3	ELSVATGAV	SEQ ID No: 53	A_0201
2	5	SVATGAVST	SEQ ID No: 54	A_0201
3	18	IPMPAGVNP	SEQ ID No: 55	B_0702, B51, Cw_0401
4	19	PMPAGVNPA	SEQ ID No: 56	A_0201
5	21	PAGVNPADL	SEQ ID No: 57	A_2402, B_3501
6	23	GVNPADLAA	SEQ ID No: 58	A3

ES 2 685 498 T3

Nº del supuesto epítipo de CD4	Posición de Amino-ácido	Secuencia de epítipo	SEQ ID No:	Alelo HLA
7	25	NPADLAAEL	SEQ ID No: 59	A24, B_0702, B44, Cw_0401, Cw_0602, B7, B_3501, B51
8	28	DLAAELAAV	SEQ ID No: 60	A_0201
9	33	LAAVVTESV	SEQ ID No: 61	A2, A_0201, B51
10	37	VTESVDEDY	SEQ ID No: 62	A1, A_0101, A_0301
11	38	TESVDEDYL	SEQ ID No: 63	B44
12	40	SVDEDYLLY	SEQ ID No: 64	A1, A_0101, A3
13	48	YECDGQWVL	SEQ ID No: 65	A24, B_4403, B51, Cw_0401, B44
14	52	GQWVLAAGV	SEQ ID No: 66	A2, A_0201
15	55	VLAAGVQAM	SEQ ID No: 67	A2, B8
16	65	ELDSDELRV	SEQ ID No: 68	A_0201
17	67	DSDELRVIR	SEQ ID No: 69	A_0101, A_0301
18	82	QQWSGRPGA	SEQ ID No: 70	A_0201, A_0301
19	86	GRPGAALGE	SEQ ID No: 71	A3, B8
20	87	RPGAALGEA	SEQ ID No: 72	B7, B_0702, B_3501, B51
21	91	ALGEAVDRL	SEQ ID No: 73	A2, A_0201
22	93	GEAVDRLLL	SEQ ID No: 74	B44
23	106	AFGWVAFEF	SEQ ID No: 75	A24
24	124	RLAPHTPLA	SEQ ID No: 76	A2, A_0201, A3
25	125	LAPHTPLAR	SEQ ID No: 77	A_0101, A_0301

ES 2 685 498 T3

Nº del supuesto epítipo de CD4	Posición de Amino-ácido	Secuencia epítipo	SEQ ID No:	Alelo HLA
26	126	APHTPLARV	SEQ ID No: 78	B7, B_3501, B51
27	127	PHTPLARVF	SEQ ID No: 79	A24, B44
28	130	PLARVFSR	SEQ ID No: 80	A_0101, A3, A_0301
29	133	RVFSRTRI	SEQ ID No: 81	B7
30	134	VFSRTRIM	SEQ ID No: 82	A24, B8
31	143	VSEKEIRLF	SEQ ID No: 83	A1, A24
32	153	AGIRHREAI	SEQ ID No: 84	B8, B51
33	164	LLATGVREV	SEQ ID No: 85	A2, A_0201
34	170	REVPQRSRV	SEQ ID No: 86	B44
35	172	VPQRSVDV	SEQ ID No: 87	B7, B8
36	180	VSDDPGFR	SEQ ID No: 88	A_0101, A_0301
37	183	DPSGFRRV	SEQ ID No: 89	B_3501, B51
38	185	SGFRRVAV	SEQ ID No: 90	B8, B51
39	186	GFRRVAVA	SEQ ID No: 91	A_0301, B8
40	187	FRRRVAVAV	SEQ ID No: 92	A3, B8, B51
41	190	RVAVAVDEI	SEQ ID No: 93	A24, B7
42	198	IAAGRYHKV	SEQ ID No: 94	B8, B51
43	200	AGRYHKVIL	SEQ ID No: 95	B7, B8
44	202	RYHKVILSR	SEQ ID No: 96	A3, A24
45	206	VILSRCEV	SEQ ID No: 97	A2, A_0201
46	214	VPFAIDFPL	SEQ ID No: 98	B7, B_3501, B51

ES 2 685 498 T3

Nº del supuesto epítipo de CD4	Posición de Amino-ácido	Secuencia de epítipo	SEQ ID No:	Alelo HLA
47	216	FAIDFPLTY	SEQ ID No: 99	A1, A_0101, B_3501, B51
48	218	IDFPLTYRL	SEQ ID No: 100	B44
49	226	LGRRHNTPV	SEQ ID No: 101	B8, B51
50	228	RRHNTPVRS	SEQ ID No: 102	A3, B8
51	231	NTPVRSFLL	SEQ ID No: 103	A_0101, A24
52	233	PVRSFLLQL	SEQ ID No: 104	A_0201, B7
53	245	RALGYSPPEL	SEQ ID No: 105	A2, A_0201, B_3501, B51
54	246	ALGYSPPELV	SEQ ID No: 106	A_0101, A2, A_0201
55	249	YSPPELVTAV	SEQ ID No: 107	A2
56	250	SPPELVTAVR	SEQ ID No: 108	B_0702, B_3501, B51
57	256	AVRADGVVI	SEQ ID No: 109	A3, A_0301, B7
58	264	ITEPLAGTR	SEQ ID No: 110	A_0101, A_0301
59	266	EPLAGTRAL	SEQ ID No: 111	B7, B8, B_3501, B51
60	270	GTRALGRGP	SEQ ID No: 112	A3, B8
61	272	RALGRGPAI	SEQ ID No: 113	A24, B7, B51
62	288	LESNSKEIV	SEQ ID No: 114	B44
63	294	EIVEHAISV	SEQ ID No: 115	A_0201
64	295	IVEHAISVR	SEQ ID No: 116	A_0101
65	298	HAISVRSSL	SEQ ID No: 117	B7, B8, B_3501
66	301	SVRSSLEEI	SEQ ID No: 118	B7
67	305	SLEEITDIA	SEQ ID No: 119	A2

ES 2 685 498 T3

Nº del supuesto epítipo de CD4	Posición de Amino-ácido	Secuencia de epítipo	SEQ ID No:	Alelo HLA
68	311	DIAEPGSAA	SEQ ID No: 120	A_0101, A_0301
69	313	AEPGSAAVI	SEQ ID No: 121	B44
70	327	RERGSVQHL	SEQ ID No: 122	B7, B44
71	331	SVQH LGSTI	SEQ ID No: 123	A24, B7
72	342	RLDPSSDRM	SEQ ID No: 124	A2, A_0201
73	344	DPSSDRMAA	SEQ ID No: 125	B7, B_3501
74	346	SSDRMAALE	SEQ ID No: 126	A1, B8
75	348	DRMAALEAL	SEQ ID No: 127	B7, Cw_0401, Cw_0602
76	349	RMAALEALF	SEQ ID No: 128	A24
77	351	AALEALFPA	SEQ ID No: 129	A2, A_0201, B_3501
78	352	ALEALFPAV	SEQ ID No: 130	A_0101, A2, A_0201
79	353	LEALFPAVT	SEQ ID No: 131	B8, B_4403
80	357	FPAVTASGI	SEQ ID No: 132	B7, B8, B_3501, B51
81	359	AVTASGIPK	SEQ ID No: 133	A3, A_0301, B7
82	360	VTASGIPKA	SEQ ID No: 134	A_0201
83	364	GIPKAAGVE	SEQ ID No: 135	A_0301, B8
84	365	IPKAAGVEA	SEQ ID No: 136	B7, B_3501
85	367	KAAGVEAIF	SEQ ID No: 137	A24, B_3501
87	369	AGVEAIFRL	SEQ ID No: 138	A_0201, B51
87	376	RLDECPRGL	SEQ ID No: 139	A2, A_0201
88	380	CPRGLYSGA	SEQ ID No: 140	B7, B_3501

Nº del supuesto epítipo de CD4	Posición de Amino-ácido	Secuencia de epítipo	SEQ ID No:	Alelo HLA
89	383	GLYSGAVVM	SEQ ID No: 141	A2, A_0201, A_0301, A3
90	384	LYSGAVVML	SEQ ID No: 142	A24, B44
91	390	VMLSADGGL	SEQ ID No: 143	A2, A_0201, A24
92	397	GLDAALTLR	SEQ ID No: 144	A_0101, A3, A_0301
93	398	LDAALTLRA	SEQ ID No: 145	A_0201, B_4403
94	400	AALTLRAAY	SEQ ID No: 146	A_0101, A_0301, B_3501
95	402	LTLRAAYQV	SEQ ID No: 147	A_0201
96	407	AYQVGGRTW	SEQ ID No: 148	A24
97	414	TWLRAGAGI	SEQ ID No: 149	A24
98	424	EESEPEREF	SEQ ID No: 150	B44
99	430	REFEETCEK	SEQ ID No: 151	B44
100	438	KLSTLTPYL	SEQ ID No: 152	A2, A_0201
101	440	STLTPYLVA	SEQ ID No: 153	A_0101, A_0201
102	441	TLTPYLVAR	SEQ ID No: 154	A3, A_0301

Como se puede ver a partir de las Tablas 2 y 3, Rv2386c contiene un número de epítomos de células T CD4+ y CD8 predichos. Además, esta información sugiere que la proteína lleva epítomos que pueden ser reconocidos por los HLAs que se presentan en todo el mundo (es decir, los HLAs de individuos caucásicos, africanos, asiáticos, o latinoamericanos – véase el sitio web en [www.allelefrequencies.net](http://www.allelefrequencies.net)).

5

Ejemplo 3 – Homólogos de H37Rv

Se identificaron las secuencias de Rv2386c a partir de varias cepas de *M. tuberculosis* y BCG utilizando búsquedas BLASTP de GenBank (secuencia de referencia H37Rv, número de acceso YP\_177877.1):

Cepa	Número de Acceso	% de Identidad
CDC1551	NP_336935.1	100

Cepa	Número de Acceso	% de Identidad
F11	YP_001288342.1	100
Haarlem	ZP_02247771.1	100
C	ZP_00878160.1	99
BCG	CAL72388.1	100

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

### Ensayos biológicos

Cuantificación de respuestas de células T a Rv2386c

- 5 Los polipéptidos se pueden cribar por su capacidad para activar células T (inducción de la proliferación y/o producción de citoquinas) en células mononucleares de sangre periférica (PBMC) o en preparaciones de sangre entera a partir de los individuos infectados (tal como los infectados de una manera latente).

10 Los individuos infectados de una manera latente normalmente se identifican mediante una prueba cutánea que tiene un diámetro mayor de 10 mm y sin síntomas, sin cultivo positivo de Mtb, con un esputo negativo, y sin lesión (como se detecta mediante radiografía de tórax).

15 Se pueden utilizar una gama de ensayos *in vitro*, basándose en muestras de PBMC o en sangre entera: después de la re-estimulación en presencia del antígeno (o variante/fragmento inmunogénico del mismo, según corresponda), se puede determinar la proliferación de las células (medida por CFSE/citometría de flujo), o se puede cuantificar la producción de citoquinas (presentes en las células del sobrenadante de cultivo y medidas mediante ELISA, o, después de la tinción intracelular de las células T CD4 y CD8 y del análisis mediante citometría de flujo).

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

### 20 Proliferación de Células T

La respuesta inmunitaria específica se puede caracterizar llevando a cabo el análisis de la proliferación de linfocitos utilizando la timidina tritiada. Esta técnica evalúa la expansión celular después de la estimulación *in vitro* a un antígeno. En la práctica, 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.

25 Más adecuadamente, la proliferación de linfocitos se puede llevar a cabo utilizando el éster de succinimidil de diacetato de carboxifluoresceína (CFSE). El CFSE se acopla de manera espontánea e irreversible tanto con las proteínas intracelulares como con las de la superficie celular, mediante reacción con las cadenas laterales de lisina y otros grupos amina disponibles. Cuando las células linfocíticas se dividen, el marcaje con CFSE se distribuye igualmente entre las células hijas, que, por lo tanto, son la mitad de fluorescentes que las progenitoras. Como resultado, la mitad de la intensidad de fluorescencia celular marca cada generación sucesiva en una población de células proliferantes, y se sigue fácilmente mediante citometría de flujo (para más detalles, véase Hodgkins, PD *et al. J. Exp. Med.* 1996 184:277-281).

35 Prácticamente, después de la descongelación, las PBMC se pueden lavar y teñir con CFSE antes de cultivarse ( $2 \times 10^5$  células) durante 72 horas con 10  $\mu\text{g/ml}$  de antígeno en el medio de cultivo (RPMI-1640 complementado con glutamina, aminoácido no esencial, piruvato, y suero AB humano inactivado por calor). Las células entonces se pueden cosechar, y se puede caracterizar su fenotipo utilizando tinción superficial para identificar las células T CD8 y CD4+ de memoria. Posteriormente, se puede utilizar el análisis de citometría de flujo para indicar el grado de proliferación de linfocitos en respuesta a cada antígeno (proporción de células con una intensidad de CFSE reducida después de la estimulación *in vitro*).

40

Producción de Citoquinas

La producción de IFN- $\gamma$  (o la producción de otras citoquinas, tales como, por ejemplo, IL2, TNF-alfa, IL5, IL12, etc.), se puede medir utilizando un ensayo inmunoabsorbente ligado a enzimas (ELISA). Las placas ELISA se pueden recubrir con un anticuerpo monoclonal de ratón dirigido a IFN- $\gamma$  humano (PharMingen, San Diego, CA) en PBS (tampón fosfato salino) durante cuatro horas a temperatura ambiente. Entonces los pocillos se bloquean con PBS que contiene leche descremada y deshidratada al 5% (peso/volumen) durante 1 hora a temperatura ambiente. Las placas se lavan entonces, por ejemplo, seis veces, en PBS/TWEEN-20 al 0,2%, y las muestras se diluyen a 1:2 en el medio de cultivo en las placas ELISA, y se incuban durante la noche a temperatura ambiente. Las placas se lavan nuevamente, y se puede agregar a cada pocillo un suero policlonal de IFN- $\gamma$  de conejo anti-humano, por ejemplo, diluido a 1:3000 en PBS/suero normal de cabra al 10%. Las placas se incuban entonces durante dos horas a temperatura ambiente, se lavan, y se puede agregar IgG de anti-conejo acoplada con peroxidasa de rábano picante (Sigma Chemical So., St. Louis, MO), por ejemplo, a una dilución de 1:2000 en PBS/leche descremada y deshidratada al 5%. Después de una incubación adicional durante dos horas a temperatura ambiente, las placas se lavan, y se agrega el sustrato de TMB. La reacción se puede interrumpir después de 20 minutos con ácido sulfúrico 1 N. Entonces se puede determinar la densidad óptica a 450 nm, utilizando 570 nm como la longitud de onda de referencia. Típicamente, las fracciones que dan como resultado que ambas réplicas den una DO dos veces mayor que la DO media a partir de las células cultivadas en el medio en solitario, se pueden considerar positivas.

Ejemplo 4- Inmunogenicidad en ratones CB6F1

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

Los ratones CB6F1 se inmunizaron intramuscularmente tres veces (en el día 0, día 14, y día 28) con 0,5 ug o 2 ug del antígeno de proteína en combinación con el sistema adyuvante AS01E (una formulación adyuvante liposomal que comprende 3D-MPL y QS21).

El diseño experimental fue como sigue:

Grupo	Día 0	Día 14	Día 28
1	2 ug Rv2386c/AS01E	2 ug Rv2386c/AS01E	2 ug Rv2386c/AS01E
2	0,5 ug Rv2386c/AS01E	0,5 ug Rv2386c/AS01E	0,5 ug Rv2386c/AS01E

Se utilizaron un total de 24 ratones en cada grupo del protocolo.

Se recolectaron, y se agruparon linfocitos de sangre periférica (PBL) 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 células T CD4 y CD8 específicas del antígeno (determinadas por las células T CD4 o CD8 que producen IL-2 y/o IFN-gamma y/o TNF-alfa) mediante citometría de flujo después de la re-estimulación *in vitro* durante la noche con grupos de péptidos 15mer que cubren las secuencias de interés.

La detección de células T de ratón que expresan IL-2 y/o IFN-gamma y/o TNF-alfa, se hizo utilizando amplificación *in vitro* impulsada por antígeno a corto plazo de la expresión de citoquina.

En resumen, se agregó la solución PharmLyse (BD-Pharmingen) a la sangre periférica de ratón heparinizada para lisar los glóbulos rojos sanguíneos. Los PBLs (linfocitos de sangre periférica) obtenidos se lavaron, y luego se incubaron en presencia de un grupo de péptidos 15-mer superpuestos por 11 aminoácidos que cubren la secuencia del antígeno de interés, y de 1 ug/ml de anticuerpos para CD28 y CD49d (BD-Pharmingen). Cada péptido 15-mer se utilizó en una concentración final de 1 ug/ml. Los controles del medio también se estimularon con anticuerpos para CD28 y CD49d.

La secreción de citoquina que bloquea el compuesto de brefeldina-A (BD-Pharmingen) se agregó 2 horas después del inicio de los cultivos a 37°C, con CO<sub>2</sub> al 5%, y las células se mantuvieron a 37°C, con CO<sub>2</sub> al 5% durante 4 horas adicionales, seguido de una incubación durante la noche a +4°C.

Después las células se cosecharon y se tiñeron con anti-CD4 acoplado con Pacific Blue (BD – clon RM4-5, BD-Pharmingen), y anticuerpos anti-CD8 alfa (clon 53-6.7, BD-Pharmingen) acoplados con proteína de peridina clorofila A (PerCp)-cianina 5,5 (Cy5.5).

Luego las células se lavaron, se fijaron, se permeabilizaron (kit Cytotfix-cytoperm, BD-Pharmingen), y se tiñeron con

anticuerpos anti-IFN-g acoplados con alofocianina (clon XMG1.2, BDPHarmingen), anticuerpos anti-IL-2 acoplados con isotiocianato de fluoresceína (FITC) (clon JES 6-5H4, Beckman Coulter), y anticuerpos anti-TNF alfa acoplados con ficoeritrina (PE) (clon MP6-XT22, 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 el subconjunto CD8+.

Para más antecedentes, véase Walzer T et al, *Cell Immunol.* 2000 206(1):16-25 y Maecker HT et al, *J. Immunol. Methods* 2001 255(1-2):27-40.

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

En cada punto de tiempo y para cada grupo, los datos se recogieron a partir de 4 grupos de 6 ratones cada uno. Los datos se presentan más adelante como el % de células T CD4 o CD8 que producen IL-2 y/o IFN-gamma y/o TNF-alfa. Se traza cada grupo individual de ratones (triángulos), así como el valor promedio del grupo (barra).

La Figura 1 muestra que en el día 21 (es decir, 7 días después de la segunda inmunización), las respuestas de células T CD4 y CD8 específicas de Rv2386c se detectan en los ratones inmunizados con cualquier dosis de Rv2386c/AS01E. Los niveles de las respuestas de células T CD4 específicas de Rv2386c son similares, independientemente de la dosis de inmunización de Rv2386. Por el contrario, los ratones inmunizados con 2 ug de Rv2386c/AS01E exhibieron respuestas de células T CD8 específicas de Rv2386c más altas que los ratones inmunizados con 0,5 ug de Rv2386c/AS01E.

La Figura 2 muestra el perfil de citoquina de la respuesta de células T CD4 a partir de PBL estimulados por el grupo del péptido de Rv2386c (medio sin eliminar) el día 21 (es decir, 7 días después de la segunda inmunización).

La Figura 3 muestra el perfil de citoquina de la respuesta de células T CD8 a partir de PBL estimulados por el grupo del péptido de Rv2386c (medio sin eliminar) 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), se detectan respuestas de células T CD4 y CD8 específicas de Rv2386c en los ratones inmunizados con cualquier dosis de Rv2386c/AS01E. Los niveles de las respuestas de células T CD4 específicas de Rv2386c son similares, independientemente de la dosis de inmunización de Rv2386. Por el contrario, los ratones inmunizados con 2 ug de Rv2386c/AS01E exhibieron respuestas de células T CD8 específicas de Rv2386c más altas que los ratones inmunizados con 0,5 ug de Rv2386c/AS01E. La tercera dosis de inmunización aumenta la respuesta de células T CD4 pero no la respuesta de células T CD8.

La Figura 5 muestra el perfil de citoquina de la respuesta de células T CD4 a partir de PBL estimulados por el grupo del péptido de Rv2386c (medio sin eliminar) el día 35 (es decir, 7 días después de la tercera inmunización).

La Figura 6 muestra el perfil de citoquina de la respuesta de células T CD8 a partir de PBL estimulados por el grupo del péptido de Rv2386c (medio sin eliminar) el día 35 (es decir, 7 días después de la tercera inmunización).

#### Ejemplo 5 – Inmunogenicidad en ratones C57BL/6

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

Los ratones C57BL/6 se inmunizaron intramuscularmente tres veces (en el día 0, el día 14, y el día 28) con 1 ug o 4 ug del antígeno de proteína en combinación con un Sistema Adyuvante AS01E (una formulación de adyuvante liposomal que comprende 3D-MPL y QS21).

El diseño experimental fue como sigue:

Grupo	Día 0	Día 14	Día 28
1	4 ug Rv2386c/AS01E	4 ug Rv2386c/AS01E	4 ug Rv2386c/AS01E
2	1 ug Rv2386c/AS01E	1 ug Rv2386c/AS01E	1 ug Rv2386c/AS01E

Los linfocitos de sangre periférica (PBL) se recolectaron y se agruparon 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 células T CD4 y CD8 específicas del antígeno (como se determina por las células T CD4 o CD8 que

producen IL-2 y/o IFN-gamma y/o TNF-alfa) mediante citometría de flujo después de la re-estimulación *in vitro* durante la noche con los grupos de péptidos 15mer que cubren las secuencias de interés. El procedimiento seguido fue como se describe anteriormente.

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

10 En cada punto de tiempo y para cada grupo, los datos se recogieron a partir de 4 grupos de 6 ratones cada uno. Los datos se presentan más adelante como el % de células T CD4 o CD8 que producen IL-2 y/o IFN-gamma y/o TNF-alfa. Se traza cada grupo individual de ratones (triángulos), así como el valor promedio del grupo (barra).

15 La Figura 7 muestra que en el día 21 (es decir, 7 días después de la segunda inmunización), se detectan respuestas de células T CD4 y CD8 específicas de Rv2386c en ratones inmunizados con cualquier dosis de Rv2386c/AS01E. Los niveles de las respuestas de células T CD4 específicas de Rv2386c son similares independientemente de la dosis de inmunización de Rv2386c. Por el contrario, los ratones inmunizados con 4 ug de Rv2386c/AS01E exhibieron respuestas de células T CD8 específicas de Rv2386c más altas que los ratones inmunizados con 1 ug de Rv2386c/AS01E. Debido a dificultades técnicas, no están disponibles los datos para la el primer grupo de células a la dosis de 4 ug.

20 La Figura 8 muestra el perfil de citoquina de la respuesta de células T CD4 a partir de PBL estimulados por el grupo del péptido de Rv2386c (medio sin eliminar) en el día 21 (es decir, 7 días después de la segunda inmunización). Debido a dificultades técnicas, no están disponibles los datos para el primer grupo de células a la dosis de 4 ug.

La Figura 9 muestra el perfil de citoquina de la respuesta de células T CD8 a partir de PBL estimulados por el grupo del péptido de Rv2386c (medio sin eliminar) en el día 21 (es decir, 7 días después de la segunda inmunización). Debido a dificultades técnicas, no están disponibles los datos para el primer de células a la dosis de 4 ug.

Los datos inmunológicos del día 35 aún no estaban disponibles en el momento en que se preparó esta solicitud.

25 En conclusión, se puede observar que el antígeno Rv2386c es capaz de provocar una respuesta inmunitaria tanto en ratones CB6F1 como en ratones C57BL/6. Además, el perfil de producción de citoquina indica que una gran proporción de células T específicas del antígeno expresan una pluralidad de citoquinas asociadas con Th1 (es decir, se provoca una respuesta de células T polifuncional). Es importante que después de la inmunización estén presentes células T específicas del antígeno, tanto CD4 como CD8, las células CD8 pueden ser particularmente importantes en un escenario de TB latente. Por lo tanto, se puede esperar que Rv2386c tenga un valor sustancial en la prevención, el tratamiento, y el diagnóstico de la infección por tuberculosis latente.

30 A través de toda la memoria y de las siguientes reivindicaciones, a menos que el contexto lo requiera de otra manera, se entenderá que la palabra "comprenden", y variaciones tales como "comprende" y "compuesto de", se entenderán por implicar la inclusión de un entero, etapa, grupo de enteros, o grupo de etapas , pero no la exclusión de cualquier otro entero, etapa, grupo de enteros, o grupo de etapas.

**Listado de secuencias**

40 <110> GlaxoSmithKline Biologicals s.a.  
Glaxo Group Limited  
Mettens, Pascal  
Brown, James  
Murphy, Dennis

45 <120> COMPOSICIONES Y MÉTODOS NOVEDOSOS

<130> VB63088PCT

<150> US61/083699

50 <151> 2008-07-25

<160> 156

<170> PatentIn versión 3.5

55 <210> 1

<211> 450

<212> PRT

<213> Mycobacterium tuberculosis

<220>  
 <221> CARACTERÍSTICAS\_VARIAS  
 <223> Ceba H37Rv

5  
 <400> 1

Met Ser Glu Leu Ser Val Ala Thr Gly Ala Val Ser Thr Ala Ser Ser  
 1 5 10 15  
 10 Ser Ile Pro Met Pro Ala Gly Val Asn Pro Ala Asp Leu Ala Ala Glu  
 20 25 30

Leu Ala Ala Val Val Thr Glu Ser Val Asp Glu Asp Tyr Leu Leu Tyr  
 35 40 45  
 15 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  
 20 Arg Gln Gln Trp Ser Gly Arg Pro Gly Ala Ala Leu Gly Glu Ala Val  
 85 90 95  
 Asp Arg Leu Leu Leu 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  
 25 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  
 30 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  
 35 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  
 40 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  
 45 Glu Ser Asn Ser Lys Glu Ile Val Glu His Ala Ile Ser Val Arg Ser  
 290 295 300  
 Ser Leu Glu Glu Ile Thr Asp Ile Ala Glu Pro Gly Ser Ala Ala Val  
 305 310 315 320  
 50 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  
 60 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  
 65 Arg Gln

ES 2 685 498 T3

450

<210> 2  
 <211> 1353  
 5 <212> ADN  
 <213> Mycobacterium tuberculosis

<220>  
 <221> características\_varias  
 10 <223> Cepa H37Rv

<400> 2

15 atgtccgagc tcagcgtcgc gacaggcgcc gtcagcaccg cgtcgtcgtc catcccgatg 60  
 cccgccggtg tcaaccccgc cgacctggca gcgagctgg cggcgggtgt taccgagtcc 120  
 gtcgacgagg attacctgct ctacgagtgc gacggccaat gggctctggc cgccggtgtg 180  
 caggcgatgg tggagctaga cagcgacgaa ctgcgctca tccgtgatgg cgttacgcgg 240  
 cgacagcaat ggtcgggtcg cccgggagcg gccctggcg aagccgtcga tccgctgttg 300  
 20 ctggaacccg atcaagcttt tggctgggtc gccttgaat tcggcgtgca ccgctatggg 360  
 ttgcagcagc ggctggcgcc gcacaccca ctggcccggg tgtttcgcc ccgaaccgg 420  
 atcatggtga gcgaaaagga gattcgctg ttcgatgctg ggattcgcca ccgcgaggcc 480  
 atcgaccgat tactcgccac cggggtgcga gaggtgccgc agtcccgtc cgtcgacgtc 540  
 tccgacgac catccggctt ccgctcgg gtggcgtag ccgtcgatga aatcgtgccc 600  
 25 ggccgctacc acaaggtgat tctgtcccgt tgtgtcgaag tgccttcgc gatcgacttt 660  
 ccgtgacct accggctggg gcgtcggcac aacaccccgg tgaggtcgtt tttgtgcag 720  
 ttgggggaa tccgtgctct gggttacagc cccgaactcg tcacggcggg gcgcgcccac 780  
 ggagtgtga tcaccgagcc gttggccgtt acccgccct tggccgtgg tcccgcatt 840  
 gaccgactgg ctcgatga cctggaatca aactcctaaag aaattgtcga gcacgccatt 900  
 30 tcagtgcgtc ctgcctga ggagattacc gacatcgccg aaccaggag tgctgcgtc 960  
 atcatttca tgacggtgcg cgagcggcg agtgtgcagc acctcggctc caccatcaga 1020  
 gcacggttg atccatcgag cgaccggatg gccgcctgg aagcccttt tctgctgtc 1080  
 actgatccg gaatcccga agcagctggc gttgaggcca tcttcgcct cgtatgagtc 1140  
 ccacgtgggc tgtattccgg tgcggtgtg atgcttcgg cggatggcgg gctagacgcc 1200  
 35 gcgctgacgc tgcgggccc ataccagtc ggcgggccc cttggctgcg ggccggcgcc 1260  
 ggcatcatc aagaatcggg gccagagcgc gaattcgagg agacttgcga aaagctatcc 1320  
 acattgacgc ctatctggt tgcacgccag taa 1353

<210> 3  
 40 <211> 450  
 <212> PRT  
 <213> Mycobacterium tuberculosis

<220>  
 45 <221> CARACTERÍSTICAS\_VARIAS  
 <223> Cepa CDC1551

<400> 3

50 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

55 Leu Ala Ala Val Val Thr Glu Ser Val Asp Glu Asp Tyr Leu Leu Tyr  
 35 40 45  
 Glu Cys Asp Gly Gln Trp Val Leu Ala Ala Gly Val Gln Ala Met Val  
 50 55 60  
 Glu Leu Asp Ser Asp Glu Leu Arg Val Ile Arg Asp Gly Val Thr Arg  
 60 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  
 65 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  
 5 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 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  
 10 195 200 205  
 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  
 15 225 230 235 240  
 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  
 20 Ala Leu Gly Arg Gly Pro Ala Ile Asp Arg Leu Ala Arg Asp Asp Leu  
 275 280 285  
 Glu Ser Asn Ser Lys Glu Ile Val Glu His Ala Ile Ser Val Arg Ser  
 290 295 300  
 Ser Leu Glu Glu Ile Thr Asp Ile Ala Glu Pro Gly Ser Ala Ala Val  
 25 305 310 315 320  
 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  
 30 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  
 35 385 390 395 400  
 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  
 40 Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala  
 435 440 445  
 Arg Gln  
 450  
  
 45 <210> 4  
 <211> 450  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 50 <220>  
 <221> CARACTERÍSTICAS\_VARIAS  
 <223> Cepa F11  
  
 <400> 4  
  
 55 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  
 60 Leu Ala Ala Val Val Thr Glu Ser Val Asp Glu Asp Tyr Leu Leu Tyr  
 35 40 45  
 Glu Cys Asp Gly Gln Trp Val Leu Ala Ala Gly Val Gln Ala Met Val  
 50 55 60  
 Glu Leu Asp Ser Asp Glu Leu Arg Val Ile Arg Asp Gly Val Thr Arg  
 65 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  
 5 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  
 10 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 Pro Ser Gly Phe Arg Arg Arg Val Ala  
                   180                  185                  190  
 15 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  
 20 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  
 25 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  
 30 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  
 35 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  
 40 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  
 45 Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala  
                   435                  440                  445  
 Arg Gln  
                   450  
 50 <210> 5  
                   <211> 450  
                   <212> PRT  
                   <213> Mycobacterium tuberculosis  
 55 <220>  
                   <221> CARACTERÍSTICAS\_VARIAS  
                   <223> Cepa Haarlem A  
 60 <400> 5  
  
 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  
 65 Leu Ala Ala Val Val Thr Glu Ser Val Asp Glu Asp Tyr Leu Leu Tyr

ES 2 685 498 T3

35 40 45  
 Glu Cys Asp Gly Gln Trp Val Leu Ala Ala Gly Val Gln Ala Met Val  
 50 55 60  
 5 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  
 10 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  
 15 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  
 20 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  
 25 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  
 30 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  
 35 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  
 40 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  
 45 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  
 50 Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala  
 435 440 445  
 Arg Gln  
 450  
 55 <210> 6  
 <211> 450  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 60 <220>  
 <221> CARACTERÍSTICAS\_VARIAS  
 <223> Cepa C  
 65 <400> 6

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  
 5 Leu Ala Ala Val Val Thr Glu Ser Val Asp Glu Asp Tyr Leu Leu Tyr  
 35 40 45  
 Glu Cys Asp Gly Gln Trp Val Leu Ala Ala Gly Val Gln Ala Met Val  
 50 55 60  
 10 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 Leu  
 100 105 110  
 15 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  
 20 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  
 25 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  
 30 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  
 35 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  
 40 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  
 45 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  
 50 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  
 55 Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala  
 435 440 445  
 Arg Gln  
 450  
 60 <210> 7  
 <211> 450  
 <212> PRT  
 <213> Mycobacterium bovis  
 65 <220>

<221> CARACTERÍSTICAS\_VARIAS

<223> Cepa BCG

<400> 7

5 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  
 10 Leu Ala Ala Val Val Thr Glu Ser Val Asp Glu Asp Tyr Leu Leu Tyr  
 35 40 45  
 Glu Cys Asp Gly Gln Trp Val Leu Ala Ala Gly Val Gln Ala Met Val  
 50 55 60  
 15 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  
 20 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  
 25 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 Pro Ser Gly Phe Arg Arg Arg Val Ala  
 180 185 190  
 30 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  
 35 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  
 40 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  
 45 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  
 50 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  
 55 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  
 60 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  
 65 450

<210> 8  
 <211> 110  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 5  
 <220>  
 <221> péptido\_mat  
 <222> (29)..(110)  
 10 <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  
 15       -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  
 20 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  
 25 Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr  
       70          75          80  
  
 <210> 9  
 <211> 97  
 <212> PRT  
 30 <213> Mycobacterium tuberculosis  
  
 <400> 9  
  
 Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser  
 35   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  
 40 Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val Ala Ala Ala Lys  
       50          55          60  
 Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala  
       65          70          75          80  
 45 Gly Thr Tyr Val Ala Ala Asp Ala Ala Ala Ser Thr Tyr Thr Gly  
       85          90          95  
 Phe  
  
 <210> 10  
 <211> 94  
 50 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 10  
  
 Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met  
 55   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  
 60       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  
 65 Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala  
       85          90

<210> 11  
 <211> 132  
 <212> PRT  
 5 <213> Mycobacterium tuberculosis

<400> 11

10 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  
 15 Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val  
 50 55 60  
 Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val  
 65 70 75 80  
 Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala  
 20 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  
 25 Gly Pro Pro Ala  
 130

<210> 12  
 <211> 195  
 30 <212> PRT  
 <213> Mycobacterium tuberculosis

<400> 12

35 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  
 40 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  
 45 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  
 50 115 120 125  
 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  
 55 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  
 60 Ala Ala Ser  
 195

<210> 13  
 <211> 391  
 <212> PRT  
 65 <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

5 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

10 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

15 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

20 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

25 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

30 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

35 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

40 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

45 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

50 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

60

<400> 14

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

65 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  
 5 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  
 10 100 105 110  
 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  
 15 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  
 20 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  
 25 210 215 220  
 Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr  
 225 230 235 240  
 Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala  
 245 250 255  
 30 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  
 35 290 295 300  
 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  
 40 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  
 45 370 375 380  
 Lys Val Leu Val Arg Asn Val Val  
 385 390  
  
 <210> 15  
 <211> 423  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 15  
 55 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 Trp Asp  
 20 25 30  
 60 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 65 70 75 80  
 Ala Leu Ala Lys Glu Thr Ala Thr Gln Ala Arg Ala Ala Ala Glu Ala

ES 2 685 498 T3

85            90            95  
 Phe Gly Thr Ala Phe Ala Met Thr Val Pro Pro Ser Leu Val Ala Ala  
           100            105            110

5    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  
 10    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  
 15    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  
 20    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  
 25    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  
 30    305            310            315            320  
 Gly Val Gly His Ala Ala Leu Val Gly Ala Leu Ser Val Pro His Ser  
       325            330            335  
 Trp Thr Thr Ala Ala Pro Glu Ile Gln Leu Ala Val Gln Ala Thr Pro  
       340            345            350  
 35    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  
 40    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

45    <210> 16  
       <211> 95  
       <212> PRT  
       <213> Mycobacterium tuberculosis

50    <220>  
       <221> MET\_INICIAL  
       <222> (1)..(1)

55    <220>  
       <221> péptido\_mat  
       <222> (2)..(95)

60    <400> 16

60    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  
 65    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  
 5 Gln Ala Met Ala Ser Thr Glu Gly Asn Val Thr Gly Met Phe Ala  
 80 85 90  
  
 <210> 17  
 <211> 338  
 10 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <220>  
 <221> péptido\_mat  
 15 <222> (43)..(338)  
  
 <400> 17  
  
 20 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  
 25 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  
 30 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  
 35 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  
 40 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  
 45 Ser Gly Leu Leu Asp Pro Ser Gln Ala Met Gly Pro Thr Leu Ile Gly  
 155 160 165  
 Leu Ala Met Gly Asp Ala Gly Gly Tyr Lys Ala Ser Asp Met Trp Gly  
 170 175 180  
 Pro Lys Glu Asp Pro Ala Trp Gln Arg Asn Asp Pro Leu Leu Asn Val  
 185 190 195  
 50 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  
 55 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  
 60 Leu Gln Arg Ala Leu Gly Ala Thr Pro Asn Thr Gly Pro Ala Pro Gln  
 280 285 290  
 Gly Ala  
 295  
  
 65 <210> 18  
 <211> 325

<212> PRT  
 <213> Mycobacterium tuberculosis

<220>  
 5 <221> péptido\_mat  
 <222> (41)..(325)

<400> 18

10 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 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  
 15 -5 -1 1 5  
 Glu Tyr Leu Gln Val Pro Ser Pro Ser Met Gly Arg Asp Ile Lys Val  
 10 15 20  
 Gln Phe Gln Ser Gly Gly Asn Asn Ser Pro Ala Val Tyr Leu Leu Asp  
 25 30 35 40  
 20 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  
 25 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  
 30 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  
 35 155 160 165  
 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  
 40 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  
 45 235 240 245  
 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  
 50 Ser Leu Gly Ala Gly  
 285

<210> 19  
 <211> 144  
 55 <212> PRT  
 <213> Mycobacterium tuberculosis

<220>  
 60 <221> MET\_INICIAL  
 <222> (1)..(1)

<220>  
 <221> péptido\_mat  
 <222> (2)..(144)

65 <400> 19

Met Ala Thr Thr Leu Pro Val Gln Arg His Pro Arg Ser Leu Phe Pro  
 -1 1 5 10 15  
 5 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  
 10 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  
 15 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

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

25 <220>  
 <221> péptido\_mat  
 <222> (24)..(228)

30 <400> 20

Met Arg Ile Lys Ile Phe Met Leu Val Thr Ala Val Val Leu Leu Cys  
 -20 -15 -10  
 35 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  
 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  
 45 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  
 50 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  
 55 155 160 165  
 Thr Asn Asp Gly Val Ile 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

60 Ser Met Leu Ala  
 205

<210> 21  
 <211> 355  
 65 <212> PRT  
 <213> Mycobacterium tuberculosis

<220>  
 <221> péptido\_mat  
 <222> (33)..(355)  
 5 <400> 21  
 Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser  
 -30 -25 -20  
 10 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  
 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  
 20 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  
 25 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  
 30 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  
 35 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  
 40 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  
 45 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  
 50 Ala Leu Asn Gly 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  
 55 <210> 22  
 <211> 323  
 <212> PRT  
 <213> Secuencia Artificial  
 60 <220>  
 <223> Ser/Ala mutante de Mtb32A maduro  
 <400> 22  
 65 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  
 5 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  
 10 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  
 15 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  
 20 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  
 25 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  
 30 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  
 35 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  
 40 Pro Pro Ala

<210> 23

<211> 96

<212> PRT

45 <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  
 50 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  
 55 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  
 60 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

65 <213> Secuencia Artificial

<220>

<223> Mtb72f

<400> 24

5 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  
 10 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  
 15 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  
 20 100 105 110  
 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  
 25 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  
 30 180 185 190  
 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  
 35 Gln Val Arg Val 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  
 40 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  
 45 290 295 300  
 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  
 50 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  
 55 370 375 380  
 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  
 60 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  
 65 450 455 460

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  
 5 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  
 10 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  
 15 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  
 20 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  
 25 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  
 30 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

35 <210> 25  
 <211> 723  
 <212> PRT  
 <213> Secuencia Artificial

40 <220>  
 <223> M72

<400> 25

45 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  
 50 Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg  
                  50            55            60

55 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  
 60 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  
 65 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  
 5            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  
 10 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  
 15 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  
 20 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  
 25 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  
 30 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  
 35 Ser Ser Gly Leu 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  
 40 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  
 45 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  
 50 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  
 55 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  
 60 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  
 65 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

ES 2 685 498 T3

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  
 5 705            710            715            720  
 Ala Ala Ser  
  
 <210> 26  
 <211> 702  
 10 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Mtb71f  
 15 <400> 26  
  
 Asp Pro Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val  
 1            5            10            15  
 20 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  
  
 25 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  
 30 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  
 35 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  
 40 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  
 45 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  
  
 50 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 Ser  
 260            265            270  
 Thr Tyr Thr Gly Phe Asp Ile Met Asp Phe Gly Leu Leu Pro Pro Glu  
 275            280            285  
 55 Val Asn Ser Ser Arg Met Tyr Ser Gly Pro Gly Pro Glu Ser Met Leu  
 290            295            300  
 Ala Ala Ala Ala Trp Asp Gly Val Ala Ala Glu Leu Thr Ser Ala  
 305            310            315            320  
 60 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  
 65 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  
 5            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  
 10            435            440            445  
 Pro Val Gln Gly Thr Gly Pro Ala Gly Pro Ala Ala Ala Ala Ala  
           450            455            460  
 Thr Gln Ala Ala Gly Ala Gly Ala Val Ala Asp Ala Gln Ala Thr Leu  
 15            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  
 20            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  
 25            545            550            555            560  
 Gly Leu Tyr Gly Asn Ala Gly Gly Leu Gly Pro Thr Gln Gly His Pro  
           565            570            575  
 Leu Ser Ser Ala Thr Asp Glu Pro Glu Pro His Trp Gly Pro Phe Gly  
           580            585            590  
 Gly Ala Ala Pro Val Ser Ala Gly Val Gly His Ala Ala Leu Val Gly  
 30            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  
 35            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  
 40            675            680            685  
 Val Val Ile Arg Glu Gln Pro Pro Pro Gly Asn Pro Pro Arg  
           690            695            700  
  
 <210> 27  
 45 <211> 920  
       <212> PRT  
       <213> Secuencia Artificial  
  
 <220>  
 50 <223> M72-Mtb9.9-Mtb9.8  
  
 <400> 27  
  
 Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly  
 55    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  
 60 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    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  
 5 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  
 10 165 170 175  
 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  
 15 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  
 20 245 250 255

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  
 25 275 280 285  
 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  
 30 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  
 35 355 360 365  
 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  
 40 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  
 45 435 440 445

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  
 50 465 470 475 480  
 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  
 55 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  
 60 545 550 555 560  
 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  
 65 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  
 5 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  
 10 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  
 15 705 710 715 720  
 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  
 20 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  
 25 785 790 795 800  
 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  
 30 Val Ala Ser Gln Ser Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His  
 835 840 845  
 Thr Ile Gly Gln Ala Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His  
 850 855 860  
 Gln Gly Glu Ser Ser Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val  
 35 865 870 875 880  
 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  
 40 Ser Thr Tyr Thr Gly Phe Pro Trp  
 915 920  
  
 <210> 28  
 <211> 1010  
 45 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> M103  
 50  
 <400> 28  
  
 Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly  
 1 5 10 15  
 55 Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg  
 20 25 30  
 Ser Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu  
 35 40 45  
  
 60 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  
 65 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  
 5 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  
 10 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  
 15 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  
  
 20 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  
 25 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 Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met  
 325 330 335  
 30 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  
 35 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  
 40 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  
  
 45 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  
 50 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  
 55 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  
 60 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  
 65 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

ES 2 685 498 T3

610            615            620  
 Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly  
 625            630            635            640  
 5 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  
 10 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  
 15 705            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  
 20 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  
 25 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  
 30 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  
 35 Pro Gln Gln Phe Ile Tyr Ala Gly Ser Leu Ser Ala Leu Leu Asp Pro  
           865            870            875            880  
 Ser Gln Gly Met Gly Pro Ser Leu Ile Gly Leu Ala Met Gly Asp Ala  
           885            890            895  
 Gly Gly Tyr Lys Ala Ala Asp Met Trp Gly Pro Ser Ser Asp Pro Ala  
           900            905            910  
 40 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  
 45 945            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  
 50 Gly Ala Gln Leu Asn Ala Met Lys Gly Asp Leu Gln Ser Ser Leu Gly  
           995            1000            1005  
 Ala Gly  
           1010  
  
 55 <210> 29  
       <211> 1148  
       <212> PRT  
       <213> Secuencia Artificial  
  
 60 <220>  
       <223> M114  
  
       <400> 29  
  
 65 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  
 5 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  
 10 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  
 15  
 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  
 20 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  
 25 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 Tyr Glu Thr Ala Tyr Gly Leu Thr  
 225 230 235 240  
 30 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  
 35 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  
 40  
 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  
 45 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  
 50 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  
 55 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  
 60 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  
 65 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  
 5    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  
 10    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  
 15    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  
 20    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  
  
 25    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  
 30    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  
 35    Pro Ala Ala Ala Met 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  
 40    820            825            830  
 Ser Leu Val Ala Ala Asn Arg Ser Arg Leu Met Ser Leu Val Ala Ala  
 835            840            845  
 Asn Ile Leu Gly Gln Asn Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu  
 850            855            860  
 45    Tyr Ala Glu Met Trp Ala Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu  
 865            870            875            880  
 Gly Ala Ser Ala Ala Ala Ser Ala Leu Pro Pro Phe Thr Pro Pro Val  
 885            890            895  
  
 50    Gln Gly Thr Gly Pro Ala Gly Pro Ala Ala Ala Ala Ala Thr Gln  
 900            905            910  
 Ala Ala Gly Ala Gly Ala Val Ala Asp Ala Gln Ala Thr Leu Ala Gln  
 915            920            925  
 Leu Pro Pro Gly Ile Leu Ser Asp Ile Leu Ser Ala Leu Ala Ala Asn  
 55    930            935            940  
 Ala Asp Pro Leu Thr Ser Gly Leu Leu Gly Ile Ala Ser Thr Leu Asn  
 945            950            955            960  
 Pro Gln Val Gly Ser Ala Gln Pro Ile Val Ile Pro Thr Pro Ile Gly  
 965            970            975  
 60    Glu Leu Asp Val Ile Ala Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser  
 980            985            990  
 Ile Ala Leu Ala Ile Thr Asn Thr Ala Arg Pro Trp His Ile Gly Leu  
 995            1000            1005  
 Tyr Gly Asn Ala Gly Gly Leu Gly Pro Thr Gln Gly His Pro Leu  
 65    1010            1015            1020  
 Ser Ser Ala Thr Asp Glu Pro Glu Pro His Trp Gly Pro Phe Gly

1025            1030            1035  
 Gly Ala Ala Pro Val Ser Ala Gly Val Gly His Ala Ala Leu Val  
 1040            1045            1050  
 5 Gly Ala Leu Ser Val Pro His Ser Trp Thr Thr Ala Ala Pro Glu  
 1055            1060            1065  
 Ile Gln Leu Ala Val Gln Ala Thr Pro Thr Phe Ser Ser Ser Ala  
 1070            1075            1080  
 Gly Ala Asp Pro Thr Ala Leu Asn Gly Met Pro Ala Gly Leu Leu  
 1085            1090            1095  
 10 Ser Gly Met Ala Leu Ala Ser Leu Ala Ala Arg Gly Thr Thr Gly  
 1100            1105            1110  
 Gly Gly Gly Thr Arg Ser Gly Thr Ser Thr Asp Gly Gln Glu Asp  
 1115            1120            1125  
 15 Gly Arg Lys Pro Pro Val Val Val Ile Arg Glu Gln Pro Pro Pro  
 1130            1135            1140  
 Gly Asn Pro Pro Arg  
 1145

20 <210> 30  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

25 <400> 30  
 Trp Val Leu Ala Ala Gly Val Gln Ala  
 1            5

30 <210> 31  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

35 <400> 31  
 Val Leu Ala Ala Gly Val Gln Ala Met  
 1            5

40 <210> 32  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

45 <400> 32  
 Val Ile Arg Asp Gly Val Thr Arg Arg  
 1            5

50 <210> 33  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

55 <400> 33  
 Leu Ala Pro His Thr Pro Leu Ala Arg  
 1            5

60 <210> 34  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

65 <400> 34  
 Ile Arg Leu Phe Asp Ala Gly Ile Arg

1            5  
 <210> 35  
 <211> 9  
 5 <212> PRT  
    <213> Mycobacterium tuberculosis  
  
    <400> 35  
 10 Ile Arg His Arg Glu Ala Ile Asp Arg  
    1            5  
  
    <210> 36  
    <211> 9  
 15 <212> PRT  
    <213> Mycobacterium tuberculosis  
  
    <400> 36  
 20 Leu Leu Ala Thr Gly Val Arg Glu Val  
    1            5  
  
    <210> 37  
    <211> 9  
 25 <212> PRT  
    <213> Mycobacterium tuberculosis  
  
    <400> 37  
 30 Phe Arg Arg Arg Val Ala Val Ala Val  
    1            5  
    <210> 38  
    <211> 9  
    <212> PRT  
 35 <213> Mycobacterium tuberculosis  
  
    <400> 38  
  
 40 Phe Ala Ile Asp Phe Pro Leu Thr Tyr  
    1            5  
  
    <210> 39  
    <211> 9  
    <212> PRT  
 45 <213> Mycobacterium tuberculosis  
  
    <400> 39  
  
 50 Phe Pro Leu Thr Tyr Arg Leu Gly Arg  
    1            5  
  
    <210> 40  
    <211> 9  
    <212> PRT  
 55 <213> Mycobacterium tuberculosis  
  
    <400> 40  
  
 60 Tyr Arg Leu Gly Arg Arg His Asn Thr  
    1            5  
  
    <210> 41  
    <211> 9  
    <212> PRT  
 65 <213> Mycobacterium tuberculosis

<400> 41  
 Val Arg Ser Phe Leu Leu Gln Leu Gly  
 1 5  
 5  
 <210> 42  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 10  
 <400> 42  
 Leu Leu Gln Leu Gly Gly Ile Arg Ala  
 1 5  
 15  
 <210> 43  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 20  
 <400> 43  
 Leu Val Thr Ala Val Arg Ala Asp Gly  
 1 5  
 25  
 <210> 44  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 30  
 <400> 44  
 Val Arg Ala Asp Gly Val Val Ile Thr  
 1 5  
 35  
 <210> 45  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 40  
 <400> 45  
 Val Val Ile Thr Glu Pro Leu Ala Gly  
 1 5  
 45  
 <210> 46  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 50  
 <400> 46  
 Ile Val Glu His Ala Ile Ser Val Arg  
 1 5  
 55  
 <210> 47  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 60  
 <400> 47  
 Ile Asp Phe Met Thr Val Arg Glu Arg  
 1 5  
 65  
 <210> 48

<211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 5 <400> 48  
 Phe Arg Leu Asp Glu Cys Pro Arg Gly  
 1 5  
 10 <210> 49  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 15 <400> 49  
 Leu Tyr Ser Gly Ala Val Val Met Leu  
 1 5  
 20 <210> 50  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 25 <400> 50  
 Tyr Ser Gly Ala Val Val Met Leu Ser  
 1 5  
 30 <210> 51  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 35 <400> 51  
 Val Val Met Leu Ser Ala Asp Gly Gly  
 1 5  
 40 <210> 52  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 45 <400> 52  
 Trp Leu Arg Ala Gly Ala Gly Ile Ile  
 1 5  
 50 <210> 53  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 55 <400> 53  
 Glu Leu Ser Val Ala Thr Gly Ala Val  
 1 5  
 60 <210> 54  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 65 <400> 54

Ser Val Ala Thr Gly Ala Val Ser Thr  
 1           5  
  
 5   <210> 55  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
  
     <400> 55  
 10   Ile Pro Met Pro Ala Gly Val Asn Pro  
       1           5  
  
     <210> 56  
 15   <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
  
     <400> 56  
 20   Pro Met Pro Ala Gly Val Asn Pro Ala  
       1           5  
  
     <210> 57  
 25   <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
  
     <400> 57  
 30   Pro Ala Gly Val Asn Pro Ala Asp Leu  
       1           5  
  
     <210> 58  
 35   <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
  
     <400> 58  
 40   Gly Val Asn Pro Ala Asp Leu Ala Ala  
       1           5  
  
     <210> 59  
 45   <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
     <400> 59  
 50   Asn Pro Ala Asp Leu Ala Ala Glu Leu  
       1           5  
     <210> 60  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
  
     <400> 60  
 55   Asp Leu Ala Ala Glu Leu Ala Ala Val  
       1           5  
 60   <210> 61  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
  
 65   <400> 61

Leu Ala Ala Val Val Thr Glu Ser Val  
 1           5  
 5   <210> 62  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
 10   <400> 62  
  
 Val Thr Glu Ser Val Asp Glu Asp Tyr  
 1           5  
 15   <210> 63  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
 20   <400> 63  
  
 Thr Glu Ser Val Asp Glu Asp Tyr Leu  
 1           5  
 25   <210> 64  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
 30   <400> 64  
  
 Ser Val Asp Glu Asp Tyr Leu Leu Tyr  
 1           5  
 35   <210> 65  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
 40   <400> 65  
  
 Tyr Glu Cys Asp Gly Gln Trp Val Leu  
 1           5  
 45   <210> 66  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
 50   <400> 66  
  
 Gly Gln Trp Val Leu Ala Ala Gly Val  
 1           5  
 55   <210> 67  
     <211> 9  
     <212> PRT  
     <213> Mycobacterium tuberculosis  
 60   <400> 67  
  
 Val Leu Ala Ala Gly Val Gln Ala Met  
 1           5  
 65   <210> 68  
     <211> 9

<212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 68  
 5  
 Glu Leu Asp Ser Asp Glu Leu Arg Val  
 1 5  
  
 <210> 69  
 10 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 69  
 15  
 Asp Ser Asp Glu Leu Arg Val Ile Arg  
 1 5  
  
 <210> 70  
 20 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 70  
 25  
 Gln Gln Trp Ser Gly Arg Pro Gly Ala  
 1 5  
  
 <210> 71  
 30 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 71  
 35  
 Gly Arg Pro Gly Ala Ala Leu Gly Glu  
 1 5  
  
 <210> 72  
 40 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 72  
 45  
 Arg Pro Gly Ala Ala Leu Gly Glu Ala  
 1 5  
  
 <210> 73  
 50 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 73  
 55  
 Ala Leu Gly Glu Ala Val Asp Arg Leu  
 1 5  
  
 <210> 74  
 60 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 74  
 65  
 Gly Glu Ala Val Asp Arg Leu Leu Leu

1            5  
 <210> 75  
 <211> 9  
 5 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 75

10 Ala Phe Gly Trp Val Ala Phe Glu Phe  
 1            5  
 <210> 76  
 <211> 9  
 15 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 76

20 Arg Leu Ala Pro His Thr Pro Leu Ala  
 1            5  
 <210> 77  
 <211> 9  
 25 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 77

30 Leu Ala Pro His Thr Pro Leu Ala Arg  
 1            5  
 <210> 78  
 <211> 9  
 35 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 78

40 Ala Pro His Thr Pro Leu Ala Arg Val  
 1            5  
 <210> 79  
 <211> 9  
 45 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 79

50 Pro His Thr Pro Leu Ala Arg Val Phe  
 1            5  
 <210> 80  
 <211> 9  
 55 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 80

60 Pro Leu Ala Arg Val Phe Ser Pro Arg  
 1            5  
 <210> 81  
 <211> 9  
 65 <212> PRT  
 <213> Mycobacterium tuberculosis

<400> 81  
 5 Arg Val Phe Ser Pro Arg Thr Arg Ile  
 1 5  
 <210> 82  
 <211> 9  
 <212> PRT  
 10 <213> Mycobacterium tuberculosis  
 <400> 82  
 15 Val Phe Ser Pro Arg Thr Arg Ile Met  
 1 5  
 <210> 83  
 <211> 9  
 <212> PRT  
 20 <213> Mycobacterium tuberculosis  
 <400> 83  
 25 Val Ser Glu Lys Glu Ile Arg Leu Phe  
 1 5  
 <210> 84  
 <211> 9  
 <212> PRT  
 30 <213> Mycobacterium tuberculosis  
 <400> 84  
 35 Ala Gly Ile Arg His Arg Glu Ala Ile  
 1 5  
 <210> 85  
 <211> 9  
 <212> PRT  
 40 <213> Mycobacterium tuberculosis  
 <400> 85  
 45 Leu Leu Ala Thr Gly Val Arg Glu Val  
 1 5  
 <210> 86  
 <211> 9  
 <212> PRT  
 50 <213> Mycobacterium tuberculosis  
 <400> 86  
 55 Arg Glu Val Pro Gln Ser Arg Ser Val  
 1 5  
 <210> 87  
 <211> 9  
 <212> PRT  
 60 <213> Mycobacterium tuberculosis  
 <400> 87  
 65 Val Pro Gln Ser Arg Ser Val Asp Val  
 1 5

5 <210> 88  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 88  
 Val Ser Asp Asp Pro Ser Gly Phe Arg  
 1 5  
 10 <210> 89  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 15 <400> 89  
 Asp Pro Ser Gly Phe Arg Arg Arg Val  
 1 5  
 20 <210> 90  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 25 <400> 90  
 Ser Gly Phe Arg Arg Arg Val Ala Val  
 1 5  
 30 <210> 91  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 35 <400> 91  
 Gly Phe Arg Arg Arg Val Ala Val Ala  
 1 5  
 40 <210> 92  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 45 <400> 92  
 Phe Arg Arg Arg Val Ala Val Ala Val  
 1 5  
 50 <210> 93  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 55 <400> 93  
 Arg Val Ala Val Ala Val Asp Glu Ile  
 1 5  
 60 <210> 94  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 65 <400> 94

Ile Ala Ala Gly Arg Tyr His Lys Val  
 1 5

5 <210> 95  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

10 <400> 95

Ala Gly Arg Tyr His Lys Val Ile Leu  
 1 5

15 <210> 96  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

20 <400> 96

Arg Tyr His Lys Val Ile Leu Ser Arg  
 1 5

25 <210> 97  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

30 <400> 97

Val Ile Leu Ser Arg Cys Val Glu Val  
 1 5

35 <210> 98  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

40 <400> 98

Val Pro Phe Ala Ile Asp Phe Pro Leu  
 1 5

45 <210> 99  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

50 <400> 99

Phe Ala Ile Asp Phe Pro Leu Thr Tyr  
 1 5

55 <210> 100  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis

60 <400> 100

Ile Asp Phe Pro Leu Thr Tyr Arg Leu  
 1 5

65 <210> 101  
 <211> 9

<212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 101  
 5 Leu Gly Arg Arg His Asn Thr Pro Val  
 1 5  
  
 <210> 102  
 10 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 102  
 15 Arg Arg His Asn Thr Pro Val Arg Ser  
 1 5  
  
 <210> 103  
 20 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 103  
 25 Asn Thr Pro Val Arg Ser Phe Leu Leu  
 1 5  
  
 <210> 104  
 30 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 104  
 35 Pro Val Arg Ser Phe Leu Leu Gln Leu  
 1 5  
  
 <210> 105  
 40 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 105  
 45 Arg Ala Leu Gly Tyr Ser Pro Glu Leu  
 1 5  
  
 <210> 106  
 50 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 106  
 55 Ala Leu Gly Tyr Ser Pro Glu Leu Val  
 1 5  
  
 <210> 107  
 60 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 107  
 65 Tyr Ser Pro Glu Leu Val Thr Ala Val

1            5  
 <210> 108  
 <211> 9  
 5    <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 108  
 10   Ser Pro Glu Leu Val Thr Ala Val Arg  
      1            5  
  
      <210> 109  
      <211> 9  
 15   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 109  
 20   Ala Val Arg Ala Asp Gly Val Val Ile  
      1            5  
  
      <210> 110  
      <211> 9  
 25   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 110  
 30   Ile Thr Glu Pro Leu Ala Gly Thr Arg  
      1            5  
  
      <210> 111  
      <211> 9  
 35   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 111  
 40   Glu Pro Leu Ala Gly Thr Arg Ala Leu  
      1            5  
  
      <210> 112  
      <211> 9  
 45   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 112  
 50   Gly Thr Arg Ala Leu Gly Arg Gly Pro  
      1            5  
  
      <210> 113  
      <211> 9  
 55   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 113  
 60   Arg Ala Leu Gly Arg Gly Pro Ala Ile  
      1            5  
  
      <210> 114  
      <211> 9  
 65   <212> PRT  
      <213> Mycobacterium tuberculosis

<400> 114  
 Leu Glu Ser Asn Ser Lys Glu Ile Val  
 5 1 5  
 <210> 115  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 10  
 <400> 115  
 Glu Ile Val Glu His Ala Ile Ser Val  
 15 1 5  
 <210> 116  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 20  
 <400> 116  
 Ile Val Glu His Ala Ile Ser Val Arg  
 25 1 5  
 <210> 117  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 30  
 <400> 117  
 His Ala Ile Ser Val Arg Ser Ser Leu  
 35 1 5  
 <210> 118  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 40  
 <400> 118  
 Ser Val Arg Ser Ser Leu Glu Glu Ile  
 45 1 5  
 <210> 119  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 50  
 <400> 119  
 Ser Leu Glu Glu Ile Thr Asp Ile Ala  
 55 1 5  
 <210> 120  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 60  
 <400> 120  
 Asp Ile Ala Glu Pro Gly Ser Ala Ala  
 65 1 5  
 <210> 121  
 <211> 9

<212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 121  
 5 Ala Glu Pro Gly Ser Ala Ala Val Ile  
 1 5  
  
 <210> 122  
 10 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 122  
 15 Arg Glu Arg Gly Ser Val Gln His Leu  
 1 5  
  
 <210> 123  
 20 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 123  
 25 Ser Val Gln His Leu Gly Ser Thr Ile  
 1 5  
  
 <210> 124  
 30 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 124  
 35 Arg Leu Asp Pro Ser Ser Asp Arg Met  
 1 5  
  
 <210> 125  
 40 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 125  
 45 Asp Pro Ser Ser Asp Arg Met Ala Ala  
 1 5  
  
 <210> 126  
 50 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 126  
 55 Ser Ser Asp Arg Met Ala Ala Leu Glu  
 1 5  
  
 <210> 127  
 60 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 127  
 65 Asp Arg Met Ala Ala Leu Glu Ala Leu

1            5  
 <210> 128  
 <211> 9  
 5    <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 128  
 10   Arg Met Ala Ala Leu Glu Ala Leu Phe  
      1            5  
  
      <210> 129  
      <211> 9  
 15   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 129  
 20   Ala Ala Leu Glu Ala Leu Phe Pro Ala  
      1            5  
  
      <210> 130  
      <211> 9  
 25   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 130  
 30   Ala Leu Glu Ala Leu Phe Pro Ala Val  
      1            5  
  
      <210> 131  
      <211> 9  
 35   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 131  
 40   Leu Glu Ala Leu Phe Pro Ala Val Thr  
      1            5  
  
      <210> 132  
      <211> 9  
 45   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 132  
 50   Phe Pro Ala Val Thr Ala Ser Gly Ile  
      1            5  
  
      <210> 133  
      <211> 9  
 55   <212> PRT  
      <213> Mycobacterium tuberculosis  
  
      <400> 133  
 60   Ala Val Thr Ala Ser Gly Ile Pro Lys  
      1            5  
  
      <210> 134  
      <211> 9  
 65   <212> PRT  
      <213> Mycobacterium tuberculosis

<400> 134  
 Val Thr Ala Ser Gly Ile Pro Lys Ala  
 1 5  
 5  
 <210> 135  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 10  
 <400> 135  
 Gly Ile Pro Lys Ala Ala Gly Val Glu  
 1 5  
 15  
 <210> 136  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 20  
 <400> 136  
 Ile Pro Lys Ala Ala Gly Val Glu Ala  
 1 5  
 25  
 <210> 137  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 30  
 <400> 137  
 Lys Ala Ala Gly Val Glu Ala Ile Phe  
 1 5  
 35  
 <210> 138  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 40  
 <400> 138  
 Ala Gly Val Glu Ala Ile Phe Arg Leu  
 1 5  
 45  
 <210> 139  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 50  
 <400> 139  
 Arg Leu Asp Glu Cys Pro Arg Gly Leu  
 1 5  
 55  
 <210> 140  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 60  
 <400> 140  
 Cys Pro Arg Gly Leu Tyr Ser Gly Ala  
 1 5  
 65  
 <210> 141

<211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 5 <400> 141  
 Gly Leu Tyr Ser Gly Ala Val Val Met  
 1 5  
 10 <210> 142  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 15 <400> 142  
 Leu Tyr Ser Gly Ala Val Val Met Leu  
 1 5  
 20 <210> 143  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 25 <400> 143  
 Val Met Leu Ser Ala Asp Gly Gly Leu  
 1 5  
 30 <210> 144  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 35 <400> 144  
 Gly Leu Asp Ala Ala Leu Thr Leu Arg  
 1 5  
 40 <210> 145  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 45 <400> 145  
 Leu Asp Ala Ala Leu Thr Leu Arg Ala  
 1 5  
 50 <210> 146  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 55 <400> 146  
 Ala Ala Leu Thr Leu Arg Ala Ala Tyr  
 1 5  
 60 <210> 147  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 65 <400> 147

Leu Thr Leu Arg Ala Ala Tyr Gln Val  
 1 5  
 5 <210> 148  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 148  
 10 Ala Tyr Gln Val Gly Gly Arg Thr Trp  
 1 5  
 15 <210> 149  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 149  
 20 Thr Trp Leu Arg Ala Gly Ala Gly Ile  
 1 5  
 25 <210> 150  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 150  
 30 Glu Glu Ser Glu Pro Glu Arg Glu Phe  
 1 5  
 35 <210> 151  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 151  
 40 Arg Glu Phe Glu Glu Thr Cys Glu Lys  
 1 5  
 45 <210> 152  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 152  
 50 Lys Leu Ser Thr Leu Thr Pro Tyr Leu  
 1 5  
 55 <210> 153  
 <211> 9  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 <400> 153  
 60 Ser Thr Leu Thr Pro Tyr Leu Val Ala  
 1 5  
 65 <210> 154  
 <211> 9  
 <212> PRT

<213> Mycobacterium tuberculosis

<400> 154

5 Thr Leu Thr Pro Tyr Leu Val Ala Arg  
1 5

<210> 155

<211> 1053

10 <212> PRT

<213> Mycobacterium tuberculosis

<400> 155

15 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  
20 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  
25 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  
30 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  
35 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  
40 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  
45 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  
50 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  
55 305 310 315 320

Phe Gly Leu Phe Asn Ser Gly Ser Gly Asn Val Gly Phe Phe Asn Ser  
325 330 335  
60 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  
65 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  
 5 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  
 10 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  
 15 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  
 20 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  
 25 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  
 30 Pro Ala Asn Ile Thr Val Ser Gly Phe Gln Leu Pro Pro Leu Ser Ile  
 625 630 635 640  
 Pro Ser Val Ala Ile Pro Pro Val Thr Val Pro Pro Ile Thr Val Gly  
 645 650 655  
 35 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 Ile 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  
 40 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  
 45 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  
 50 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  
 55 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  
 60 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  
 65 Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr

900 905 910  
 Thr Pro Pro Leu Thr Ile Glu Pro Ile Gly Val Gly Gly Phe Thr Thr  
 915 920 925  
 5 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  
 10 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  
 15 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  
 20  
 <210> 156  
 <211> 324  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
 25  
 <400> 156  
  
 Met Ser Asp Gln Val Pro Lys Pro His Arg His His Ile Trp Arg Ile  
 1 5 10 15  
 30 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  
 35 Asp Thr Leu Pro Ala Ile Glu Lys Ser Ala Leu Ser Thr Ala His Ser  
 65 70 75 80  
 Phe Phe Ser Pro Ser Val Val Asn Glu Ile Ile Glu Pro Thr Ile Gly  
 85 90 95  
 40 Asp Ile Thr Asn Asn Ala Arg Gly Glu Val Ala Ser Leu Gly Phe Leu  
 100 105 110  
 Ile Ser Leu Trp Ala Gly Ser Ser Ala Ile Ser Ala Phe Val Asp Ala  
 115 120 125  
 Val Val Glu Ala His Asp Gln Thr Pro Leu Arg His Pro Val Arg Gln  
 45 130 135 140  
 Arg Phe Phe Ala Leu Phe Leu Tyr Val Val Met Leu Val Phe Leu Val  
 145 150 155 160  
 Ala Thr Ala Pro Val Met Val Val Gly Pro Arg Lys Val Ser Glu His  
 165 170 175  
 50 Ile Pro Glu Ser Leu Ala Asn Leu Leu Arg Tyr Gly Tyr Tyr Pro Ala  
 180 185 190  
 Leu Ile Leu Gly Leu Thr Val Gly Val Ile Leu Leu Tyr Arg Val Ala  
 195 200 205  
 Leu Pro Val Pro Leu Pro Thr His Arg Leu Val Leu Gly Ala Val Leu  
 55 210 215 220  
 Ala Ile Ala Val Phe Leu Ile Ala Thr Leu Gly Leu Arg Val Tyr Leu  
 225 230 235 240  
 Ala Trp Ile Thr Arg Thr Gly Tyr Thr Tyr Gly Ala Leu Ala Thr Pro  
 245 250 255  
 60 Ile Ala Phe Leu Leu Phe Ala Phe Phe Gly Gly Phe Ala Ile Met Leu  
 260 265 270  
 Gly Ala Glu Leu Asn Ala Ala Val Gln Glu Glu Trp Pro Ala Pro Ala  
 275 280 285  
 Thr His Ala His Arg Leu Gly Asn Trp Leu Lys Ala Arg Ile Gly Val  
 65 290 295 300  
 Gly Thr Thr Thr Tyr Ser Ser Thr Ala Gln His Ser Ala Val Ala Ala

ES 2 685 498 T3

305            310            315            320  
Glu Pro Pro Ser

5

**REIVINDICACIONES**

1. Un polipéptido aislado que comprende:
- (i) una secuencia proteica de Rv2386c;
- 5 (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; para el uso en:
- (a) el tratamiento de la tuberculosis;
- (b) el tratamiento de la tuberculosis latente;
- 10 (c) la prevención de la tuberculosis latente;
- (d) la prevención de la reactivación de la tuberculosis; o
- (e) el retraso de la reactivación de la tuberculosis.
2. El polipéptido para el uso según la reivindicación 1, que comprende una secuencia proteica de Rv2386c.
- 15 3. El polipéptido para el uso según la reivindicación 1, que comprende una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma.
4. El polipéptido para el uso según la reivindicación 1, que comprende un fragmento inmunogénico de una secuencia proteica de Rv2386c.
5. El polipéptido para el uso según cualquiera de las reivindicaciones 1 a 4., en donde la proteína Rv2386c tiene la secuencia de SEQ ID NO: 1.
- 20 6. El polipéptido para el uso según cualquiera de las reivindicaciones 1 a 5, en donde dicho uso es el tratamiento de la tuberculosis latente.
7. El polipéptido para el uso según cualquiera de las reivindicaciones 1 a 5, en donde dicho uso es la prevención de la tuberculosis latente.
- 25 8. El polipéptido para el uso según cualquiera de las reivindicaciones 1 a 5, en donde dicho uso es la prevención de la reactivación de la tuberculosis.
9. El polipéptido para el uso según cualquiera de las reivindicaciones 1 a 5, en donde dicho uso es el retraso de la reactivación de la tuberculosis.
10. Un polinucleótido aislado que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende:
- 30 (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c
- 35 en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6;
- para el uso en:
- (a) el tratamiento de la tuberculosis latente;
- (b) la prevención de la tuberculosis latente;
- (c) la prevención de la reactivación de la tuberculosis; o
- 40 (d) el retraso de la reactivación de la tuberculosis
- 11.. El polinucleótido para el uso según la reivindicación 10, que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una secuencia proteica de Rv2386c.

12. El polinucleótido para el uso según la reivindicación 10, que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma.
- 5 13. El polinucleótido para el uso según la reivindicación 10, que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende un fragmento inmunogénico de una secuencia proteica de Rv2386c.
14. El polinucleótido para el uso según cualquiera de las reivindicaciones 10 a 13, en donde la proteína de Rv2386c tiene la secuencia de SEQ ID NO: 1.
15. El polinucleótido para el uso según cualquiera de las reivindicaciones 10 a 14, en donde dicho uso es el tratamiento de la tuberculosis latente.
- 10 16. El polinucleótido para el uso según cualquiera de las reivindicaciones 10 a 14, en donde dicho uso es la prevención de la tuberculosis latente.
17. El polinucleótido para el uso según cualquiera de las reivindicaciones 10 a 14, en donde dicho uso es la prevención de la reactivación de la tuberculosis.
- 15 18. El polinucleótido para el uso según cualquiera de las reivindicaciones 10 a 14, en donde dicho uso es el retraso de la reactivación de la tuberculosis.
19. Una composición que comprende:
- (a) un polipéptido que comprende:
- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- 20 en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6;
- (b) un polinucleótido que comprende una secuencia de ácido nucleico que codifica el polipéptido de (a); y un componente M72;
- 25 en donde el componente M72 tiene la secuencia de SEQ ID NO: 25.
20. Una composición que comprende:
- (a) un polipéptido que comprende:
- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- 30 en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; y
- (b) monofosforil lípido A 3-des-O-acilado (3D-MPL®), y QS21 en una formulación liposomal.
21. Una composición inmunogénica que comprende un polipéptido que comprende:
- 35 (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO:1 y 6;
- 40 y un potenciador de la respuesta inmunitaria no específica.
22. El uso de un polipéptido que comprende:

- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene al menos el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- 5 en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; en la fabricación de un medicamento para:
- (a) el tratamiento de la tuberculosis;
- (b) el tratamiento de la tuberculosis latente;
- (c) la prevención de la tuberculosis latente;
- 10 (d) la prevención de la reactivación de la tuberculosis; o
- (e) el retraso de la reactivación de la tuberculosis.
23. El uso de un polinucleótido que comprende una secuencia de ácido nucleico que codifica un polipéptido que comprende:
- (i) una secuencia proteica de Rv2386c;
- 15 (ii) una variante de una secuencia proteica de Rv2386c que tiene el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; en la fabricación de un medicamento para:
- 20 (a) el tratamiento de la tuberculosis latente;
- (b) la prevención de la tuberculosis latente;
- (c) la prevención de la reactivación de la tuberculosis; o
- (d) el retraso de la reactivación de la tuberculosis.
- 25 24. Una proteína de fusión que comprende un polipéptido que comprende:
- (i) una secuencia proteica de Rv2386c;
- (ii) una variante de una secuencia proteica de Rv2386c que tiene el 90% de identidad con la misma; o
- (iii) un fragmento inmunogénico de una secuencia proteica de Rv2386c;
- 30 en donde la secuencia proteica de Rv2386c se selecciona de SEQ ID NO: 1 y 6; para el uso en:
- (a) el tratamiento de la tuberculosis;
- (b) el tratamiento de la tuberculosis latente;
- (c) la prevención de la tuberculosis latente;
- (d) la prevención de la reactivación de la tuberculosis; o
- 35 (e) el retraso de la reactivación de la tuberculosis.

Figura 1

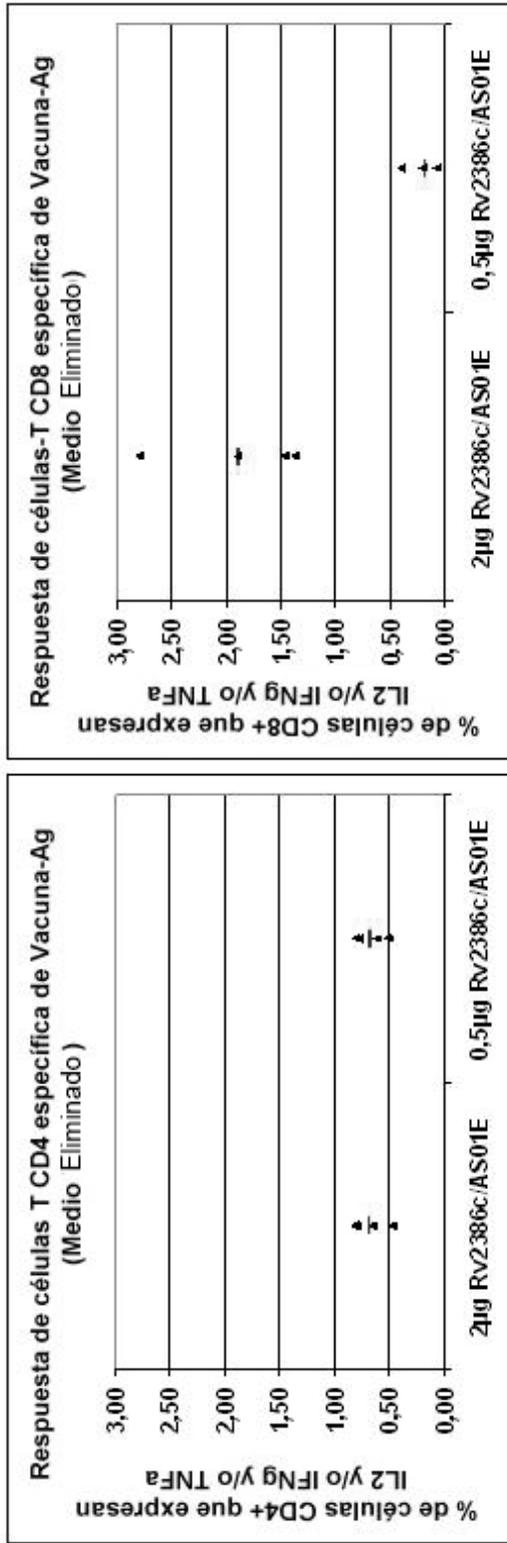


Figura 2

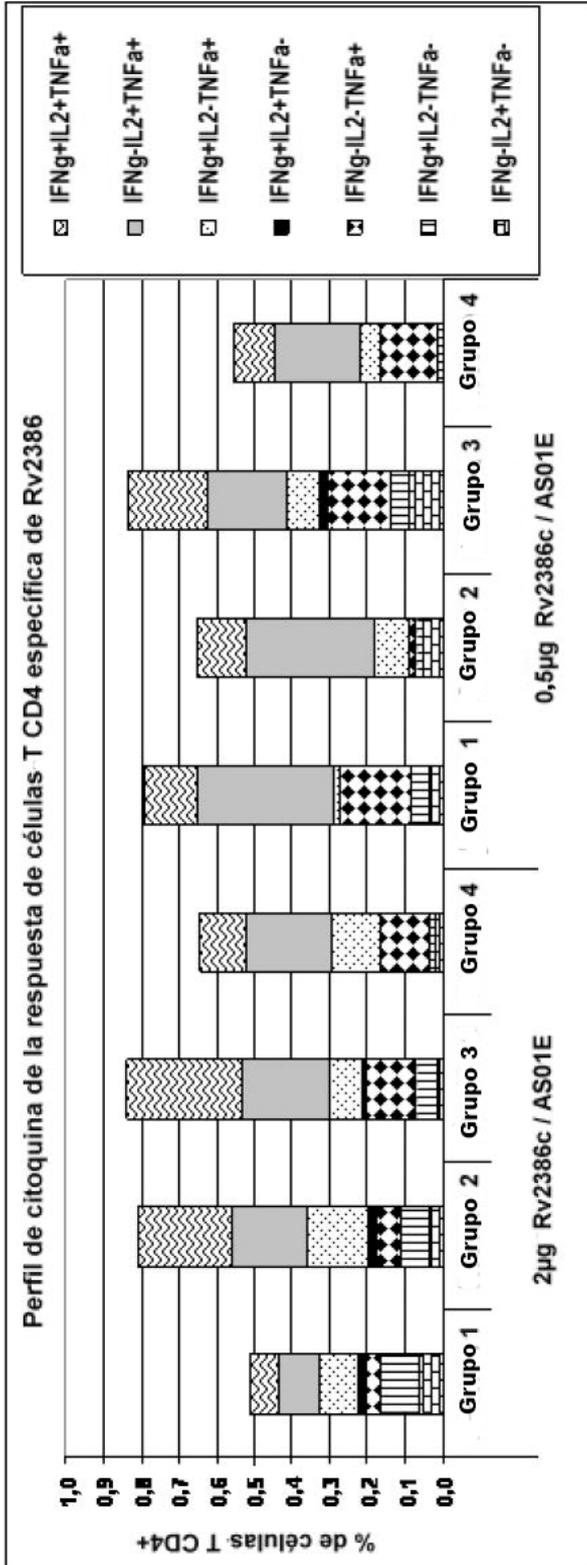


Figura 3

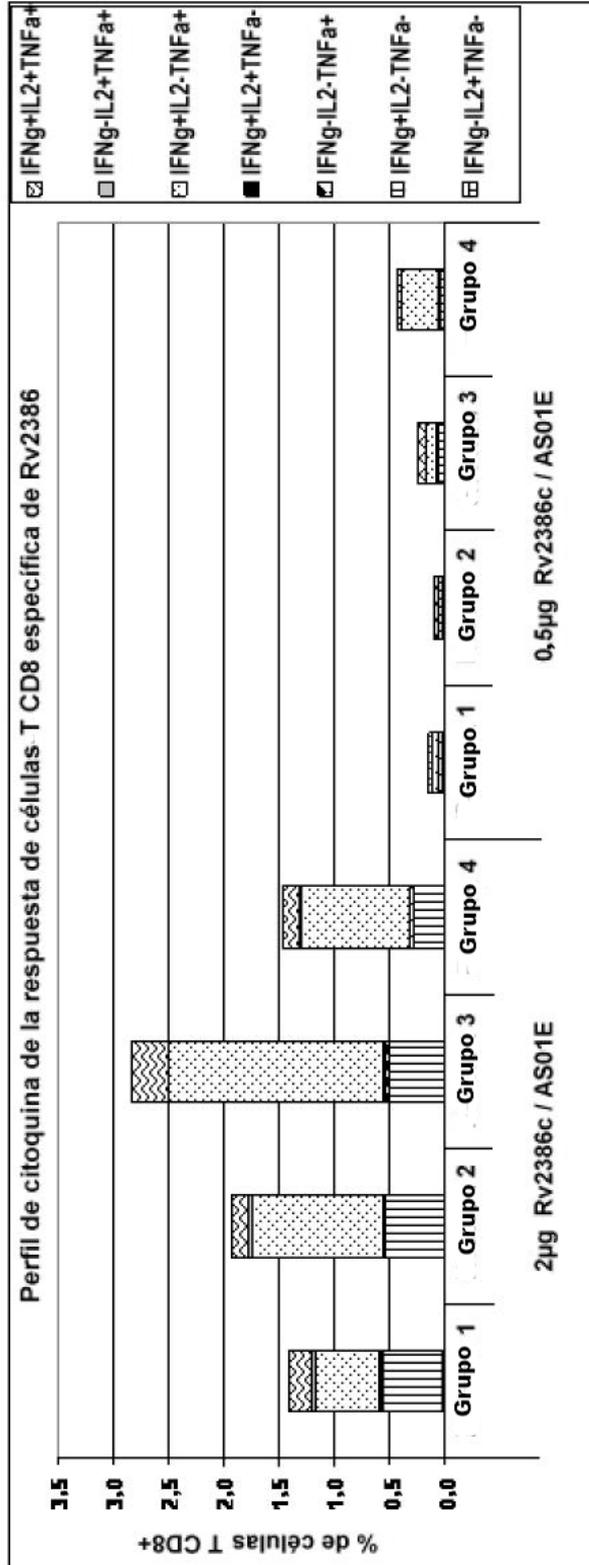


Figura 4

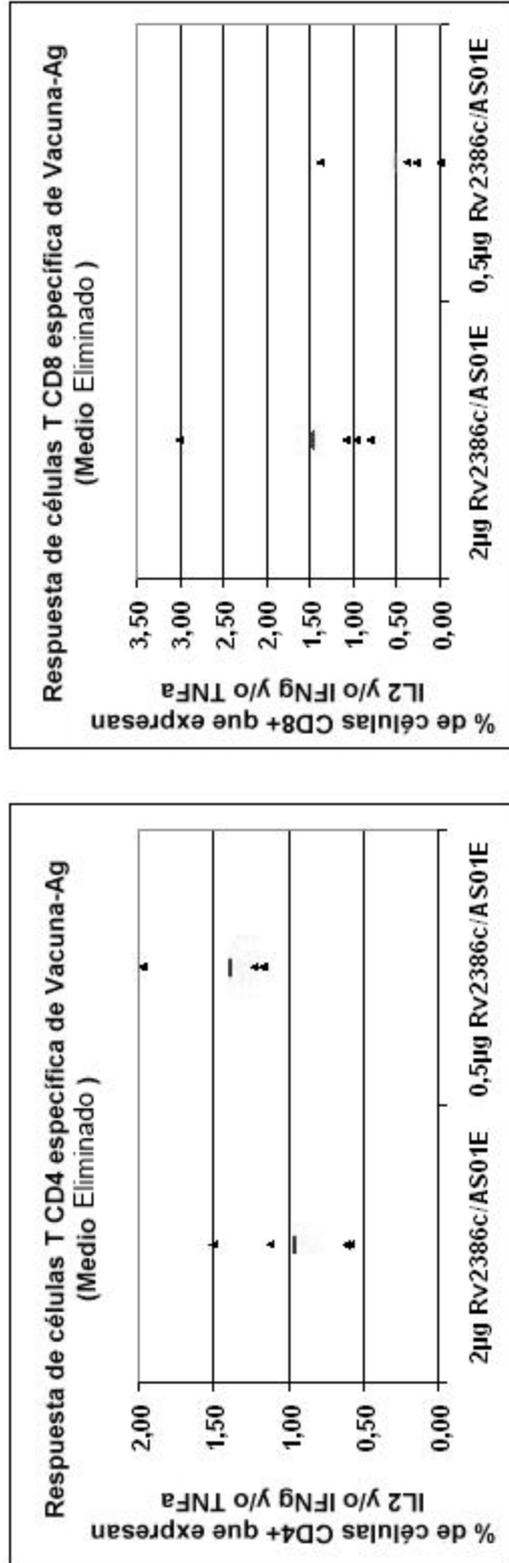


Figura 5

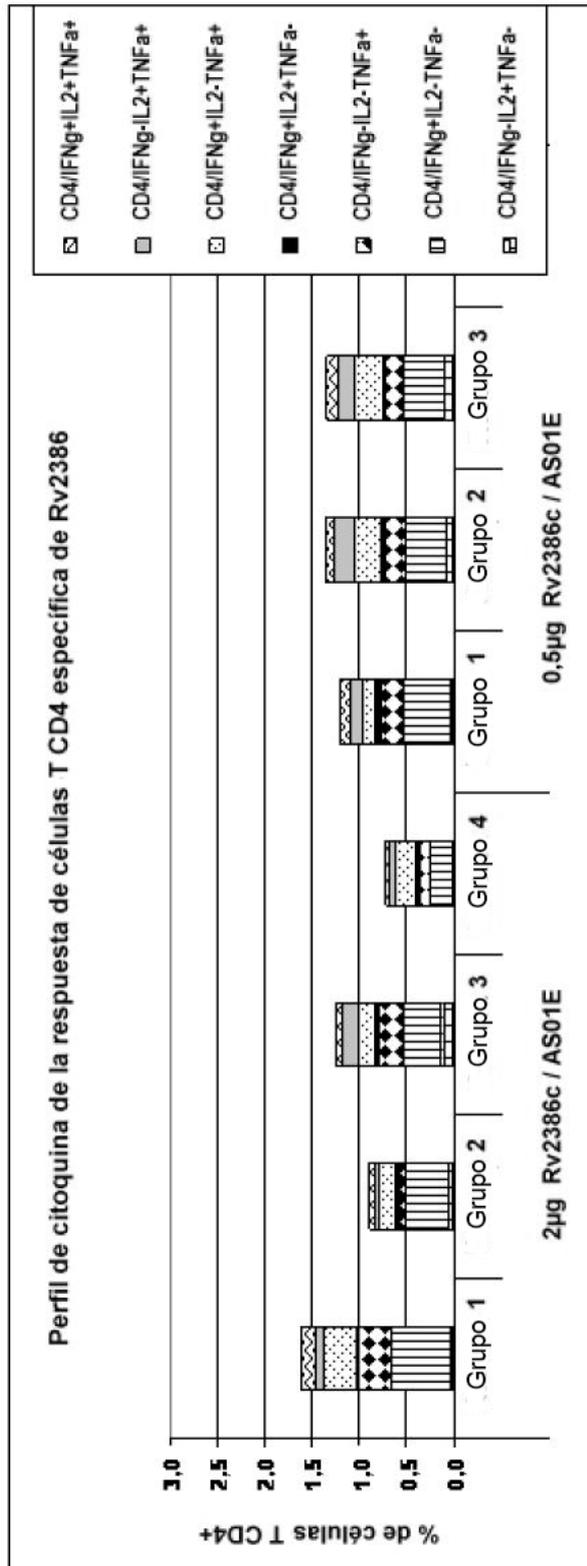


Figura 6

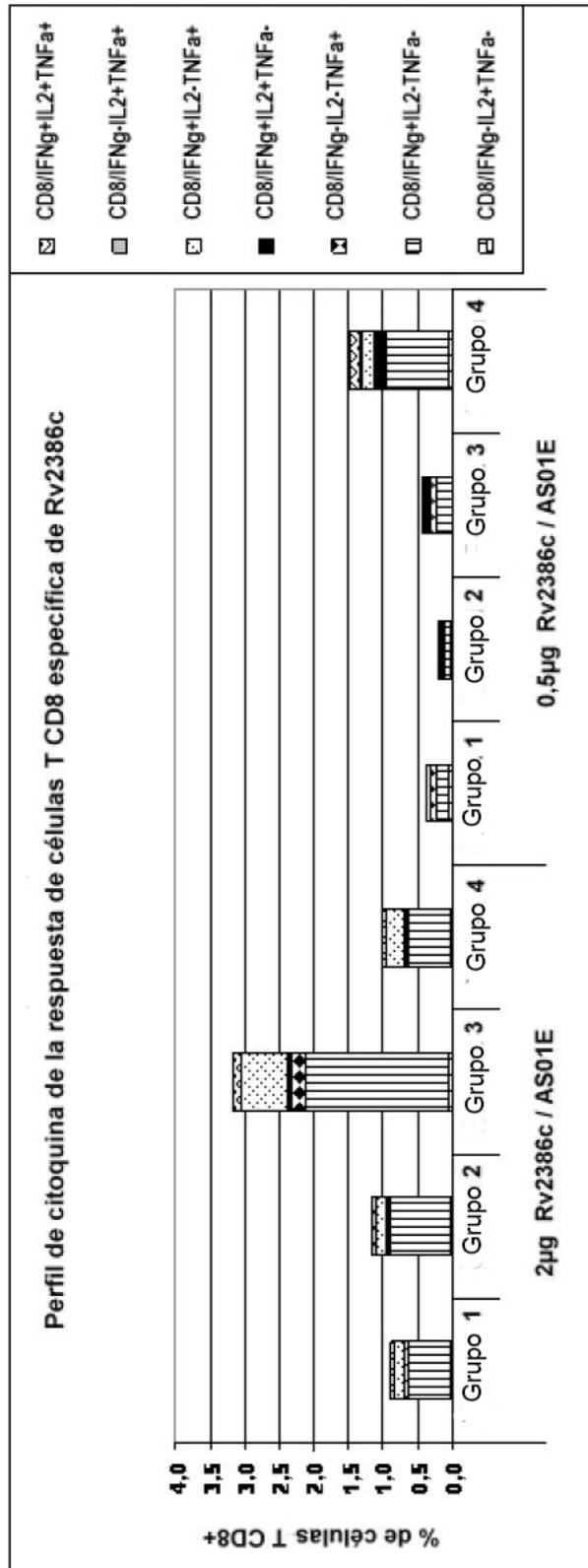


Figura 7

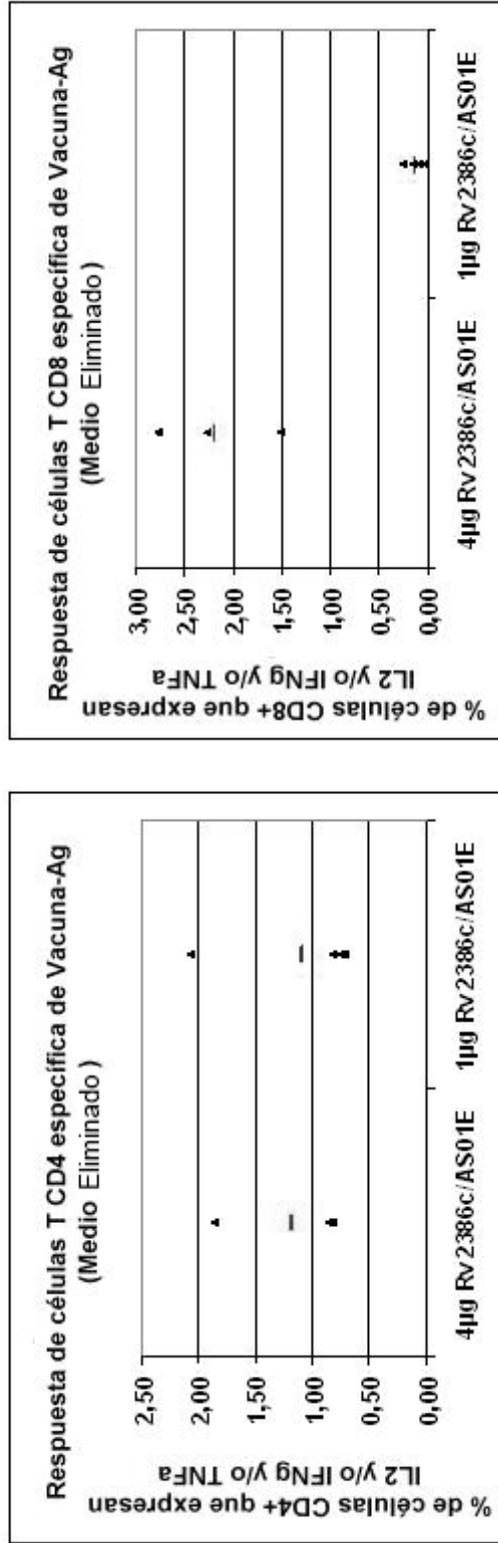


Figura 8

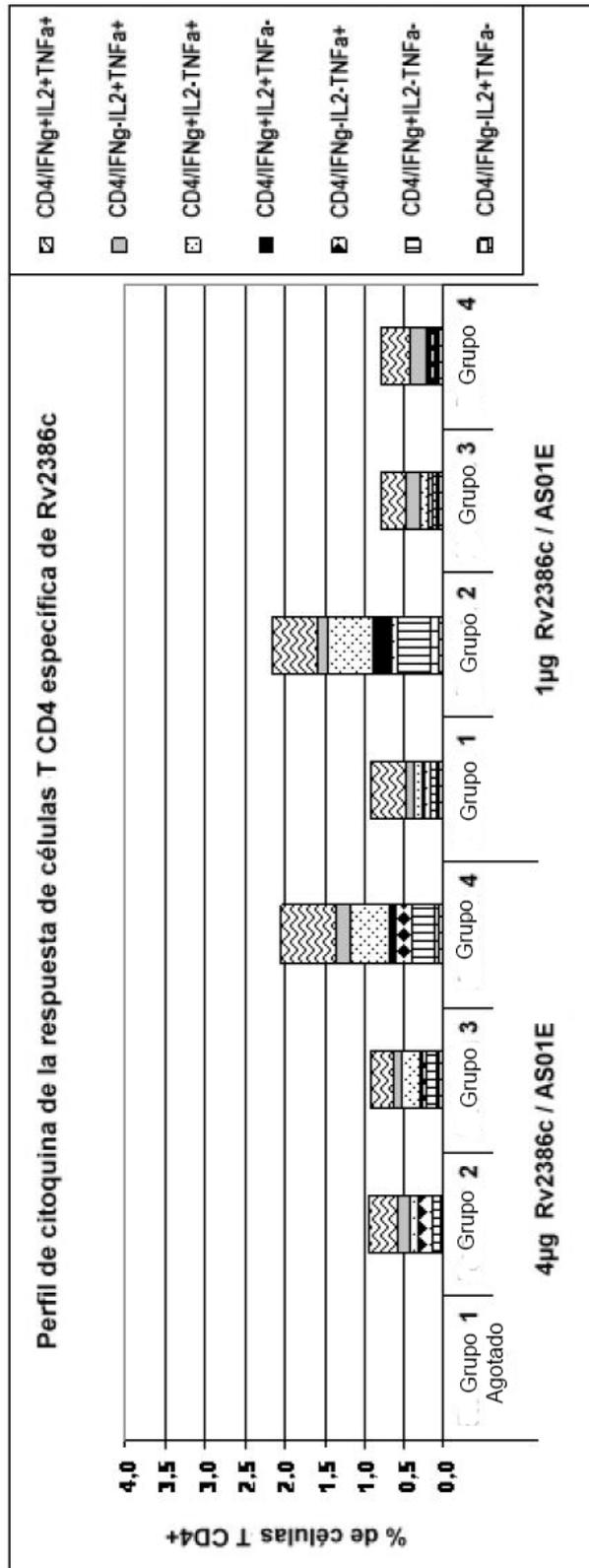


Figura 9

