

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 660 468**

51 Int. Cl.:

**C12N 15/62** (2006.01)

**C07K 14/33** (2006.01)

**A61K 39/08** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.05.2012 PCT/EP2012/059793**

87 Fecha y número de publicación internacional: **06.12.2012 WO12163811**

96 Fecha de presentación y número de la solicitud europea: **25.05.2012 E 12724950 (6)**

97 Fecha y número de publicación de la concesión europea: **10.01.2018 EP 2714910**

54 Título: **Composición inmunogénica**

30 Prioridad:

**27.05.2011 US 201161490707 P**  
**27.05.2011 US 201161490734 P**  
**27.05.2011 US 201161490716 P**

45 Fecha de publicación y mención en BOPI de la  
traducción de la patente:  
**22.03.2018**

73 Titular/es:

**GLAXOSMITHKLINE BIOLOGICALS SA (100.0%)**  
**Rue de l'Institut 89**  
**1330 Rixensart, BE**

72 Inventor/es:

**CASTADO, CINDY**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 660 468 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

Composición inmunogénica

**Campo técnico**

La presente invención se refiere a antígenos de *Clostridium difficile*. En particular, la invención se refiere a polipéptidos recombinantes que comprenden fragmentos de toxina A y toxina B de *C. difficile*. La invención se refiere adicionalmente a composiciones inmunogénicas o a vacunas que comprenden estos polipéptidos, y al uso de las vacunas y las composiciones inmunogénicas de la invención en profilaxis o terapia. La invención también se refiere a procedimientos de inmunización que usan las composiciones de la invención, y al uso de las composiciones de la invención en la fabricación de un medicamento.

**Antecedentes**

*C. difficile* es la causa más importante de las infecciones intestinales nosocomiales y es la principal causa de colitis pseudomembranosa en seres humanos (Bartlett y col., Am. J. Clin. Nutr. 11 supl:2521-6 (1980)). La tasa de mortalidad global asociada para individuos infectados con *C. difficile* se calculó en 5,99 % dentro de los 3 meses del diagnóstico, con una mortalidad más alta asociada a la edad avanzada, que es 13,5 % en pacientes mayores de 80 años (karas y col., Journal of Infection 561:1-9 (2010)). El tratamiento actual para la infección por *C. difficile* es la administración de antibióticos (metronidazol y vancomicina), sin embargo, ha habido indicio de cepas que son resistentes a estos antibióticos (Shah y col., Expert Rev. Anti Infect. Ther. 8(5), 555-564 (2010)). Belyi Iouri F y col., (2003) FEMS Microbiology Letters, ya no publicado por Elsevier, páginas 325-329. Los documentos WO 00/61762 y WO 2010/017383 desvelan proteínas de fusión de toxinas clostridiales. Por consiguiente, existe la necesidad de composiciones inmunogénicas que puedan inducir anticuerpos y/o una respuesta inmune protectora contra *C. difficile*.

**Breve resumen**

La enterotoxigenicidad de *C. difficile* se debe principalmente a la acción de dos toxinas, la toxina A y la toxina B. Estas son citotoxinas potentes (Lyerly y col., Current Microbiology 21:29-32 (1990)). Los dominios C-terminales de la toxina A y la toxina B comprenden unidades repetitivas, por ejemplo, el dominio C-terminal de la toxina A está constituido por unidades repetitivas contiguas (Dove y col., Infect. Immun. 58:480-499 (1990)), por esta razón, el dominio C-terminal puede denominarse el "dominio de repetición". Estas porciones de repetición se pueden separar aun más en repeticiones cortas (RC) y repeticiones largas (RL) como se describe en Ho y col., (PNAS 102:18373-18378 (2005)).

Se ha determinado la estructura de un fragmento de 127 aa del extremo C del dominio de repetición de la toxina A (Ho y col., PNAS 102:18373-18378 (2005)). Este fragmento formó un pliegue tipo  $\beta$ -solenoides, compuesto predominantemente por cadenas  $\beta$  con una baja proporción de hélices  $\alpha$ .

Se ha demostrado que los fragmentos de la toxina A, en particular los fragmentos del dominio C-terminal, pueden conducir a una respuesta inmune protectora en hamsters (Lyerly y col., Current Microbiology 21:29-32 (1990)), documentos WO96/12802 y WO00/61762.

Se sabe que existe una dificultad de diseño de proteínas de fusión que se plieguen correctamente durante la expresión. Los polipéptidos de la presente invención son proteínas de fusión en las que se mantiene la estructura de tipo  $\beta$  solenoides nativo, y que se ve que proporcionan una respuesta inmune tanto contra la toxina A como la toxina B en ratones.

En un primer aspecto de la invención, se proporciona un polipéptido que comprende un primer fragmento y un segundo fragmento, en el que

(i) el primer fragmento es un fragmento del dominio de repetición de la toxina A y comprende al menos 100 aminoácidos;

(ii) el segundo fragmento es un fragmento del dominio de repetición de la toxina B y comprende al menos 100 aminoácidos;

(iii) el extremo proximal del primer fragmento está situado dentro de una primera porción de repetición;

(iv) el extremo proximal del segundo fragmento está situado dentro de una segunda porción de repetición y

en el que el primer fragmento y el segundo fragmento están separados por menos de o exactamente 5 aminoácidos en la estructura primaria, en la que los polipéptidos provocan anticuerpos que neutralizan la toxina A y la toxina B y en el que la primera porción de repetición y la segunda porción de repetición tienen una identidad de secuencia entre sí superior a 50 %.

En un segundo aspecto de la invención, se proporciona un polinucleótido que codifica el polipéptido de la invención.

En un tercer aspecto de la invención, se proporciona un vector que comprende el polinucleótido de la invención unido a un promotor inducible.

En un cuarto aspecto de la invención, se proporciona una célula huésped que comprende el vector de la invención o el polinucleótido de la invención.

- 5 En un quinto aspecto de la divulgación, se proporciona una composición inmunogénica que comprende el polipéptido de la invención y un excipiente farmacéuticamente aceptable.

En un sexto aspecto de la divulgación, se proporciona una vacuna que comprende la composición inmunogénica de la invención y un excipiente farmacéuticamente aceptable.

- 10 En un séptimo aspecto de la divulgación, se proporciona un uso de la composición inmunogénica de la invención o la vacuna de la invención en el tratamiento o la prevención de la enfermedad de *C.difficile*.

En un octavo aspecto de la divulgación, se proporciona un uso de la composición inmunogénica de la invención o la vacuna de la invención en la preparación de un medicamento para la prevención o el tratamiento de la enfermedad de *C.difficile*.

- 15 En un noveno aspecto de la divulgación, se proporciona un procedimiento para evitar o tratar la enfermedad de *C.difficile* que comprende administrar la composición inmunogénica de la invención o la vacuna de la invención a un paciente.

### **Breve descripción de los dibujos**

Fig. 1-Fig.1- Listas de secuencias de los polipéptidos de la invención.

- 20 Fig. 2- Representación pictórica de los dominios C-terminales de ToxA y ToxB, con las repeticiones RC representadas como recuadros blancos y los recuadros de RL representados como recuadros negro.

Fig. 3- Representación pictórica de una unión entre la tercera RC VIII de ToxA y la cuarta RC II de ToxB usadas en Fusión 1.

Fig. 4- Representación pictórica de una unión entre la segunda RC VIII de ToxA y la tercera RC II de ToxB usadas en Fusión 2.

- 25 Fig. 5- Representación pictórica de una unión entre la RL VII de ToxA y la RL II de ToxB usadas en Fusión 3 (que contiene solo parte de la RL VII de ToxA y parte de la RL II de ToxB).

Fig. 6- Representación pictórica de una unión entre la segunda RC VIII de ToxA y la tercera RC I de ToxB usadas en Fusión 4.

- 30 Fig. 7- Representación pictórica de una unión que comprende un engarce de glicina entre el último resto de la secuencia de la proteína ToxA y el comienzo de la cuarta RC II de ToxB usadas en Fusión 5.

Fig 8.- Gráficos que describen la distribución de las fusiones 1-5 de ToxA-ToxB *C.difficile* como se determina por ultracentrifugación analítica de la velocidad de sedimentación. El estudio de panel a) describe la distribución de Fusión 1, el estudio de panel b) describe la distribución de Fusión 2, el estudio de panel c) describe la distribución de Fusión 3, el estudio de panel d) describe la distribución de Fusión 4 y el estudio de panel e) describe la distribución de Fusión 5.

- 35 Fig. 9 -Gráfico que describe el espectro UV lejano de las Fusiones, 2, 3, 4 y 5, medido usando dicroísmo circular. El espectro para la fusión 2 está representado por una línea con los puntos representados como cuadrados pequeños, el espectro para la fusión 3 está representado por una línea con los puntos representados como pequeñas formas de diamante, la fusión 4 está representada por una línea con los puntos representados como círculos, y la fusión 5 está representada por una línea con los puntos representados como formas en cruz.

- 40 Fig. 10 -Gráfico que describe el espectro UV cercano de las Fusiones 2, 3, 4 y 5 medidas usando dicroísmo circular. El espectro para la fusión 2 está representado por una línea con los puntos representados como formas en cruz, el espectro para la fusión 3 está representado por una línea con los puntos representados como círculos, el espectro para la fusión 4 está representado por una línea con los puntos representados como triángulos, y el espectro para la fusión 5 está representado por una línea con los puntos representados como pequeñas formas de diamante.

- 45 Fig. 11- Gráfico que muestra la inmunogenicidad anti-ToxA en ratones inmunizados con un fragmento del extremo C de la toxina A (aa 2387-2706), un fragmento del extremo C de la toxina B (aa 1750-2360) o las fusiones 1, 2, 3, 4 o 5.

- 50 Fig. 12- Gráfico que muestra la inhibición de la hemaglutinación en ratones inmunizados con un fragmento del

extremo C de la toxina A (aa 2387-2706), un fragmento del extremo C de la toxina B (aa 1750-2360) o las fusiones 1, 2, 3, 4 o 5.

Fig. 13- Gráfico que muestra la inmunogenicidad anti-ToxB en ratones inmunizados con un fragmento del extremo C de la toxina A (aa 2387-2706), un fragmento del extremo C de la toxina B (aa 1750-2360) o las fusiones 1,2, 3, 4 o 5.

Fig. 14- Títulos de inhibición de la citotoxicidad de ratones inmunizados con un fragmento del extremo C de la toxina A (aa 2387-2706), un fragmento del extremo C de la toxina B (aa 1750-2360) o las fusiones 1, 2, 3, 4 o 5.

Fig. 15 -Gráficos que describen la distribución de las fusiones F52Nueva, F54Gli, F54Nueva y F5ToxB de ToxA-ToxB *C.difficile* como se determina por ultracentrifugación analítica de la velocidad de sedimentación. El estudio de panel a) describe la distribución de F52Nueva, el estudio de panel b) describe la distribución de F54Gli, el estudio de panel c) describe la distribución de F54Nueva y el estudio de panel d) describe la distribución de F5ToxB.

Fig. 16 -Gráfico que describe el espectro UV lejano de las fusiones F52Nueva, F54Gli, F54Nueva y F5ToxB medido usando dicroísmo circular. El espectro para F52Nueva está representado por una línea con los puntos representados como cruces dobles, el espectro para F54Gli está representado por una línea con los puntos representados como triángulos, F54Nueva está representada por una línea con los puntos representados como cuadrados, y F5ToxB está representada por una línea con los puntos representados como formas en cruz.

Fig. 17 -Gráfico que describe el espectro UV cercano de las fusiones F52Nueva, F54Gli, F54Nueva y F5ToxB medido usando dicroísmo circular. El espectro para F52Nueva está representado por una línea con los puntos representados como cruces dobles, el espectro para F54Gli está representado por una línea con los puntos representados como triángulos, F54Nueva está representada por una línea con los puntos representados como cuadrados, y F5ToxB está representada por una línea con los puntos representados como formas en cruz.

Fig. 18- Gráfico que muestra los resultados de ELISA anti-ToxA para ratones inmunizados con las fusiones F2, F52Nueva, F54Gli, G54Nueva o F5ToxB.

Fig. 19- Gráfico que muestra los resultados de ELISA anti-ToxB para ratones inmunizados con las fusiones F2, F52Nueva, F54Gli, F54Nueva o F5ToxB.

Fig. 20- Gráfico que muestra la inhibición de la hemaglutinación en ratones inmunizados con las fusiones F2, F52Nueva, F54Gli, F54Nueva o F5ToxB.

Fig. 21- Gráfico que muestra los títulos de citotoxicidad en células HT29 de ratones inmunizados con las fusiones F2, F52Nueva, F54Gli, F54Nueva o F5ToxB.

Fig. 22- Gráfico que muestra los títulos de citotoxicidad en células IMR90 de ratones inmunizados con las fusiones F2, F52Nueva, F54Gli, F54Nueva o F5ToxB.

### **Descripción detallada**

#### **POLIPÉPTIDOS**

La invención se refiere a un polipéptido que comprende un primer fragmento y un segundo fragmento, en el que

(i) el primer fragmento es un fragmento del dominio de repetición de la toxina A y comprende al menos 100 aminoácidos;

(ii) el segundo fragmento es un fragmento del dominio de repetición de la toxina B y comprende al menos 100 aminoácidos;

(iii) el extremo proximal del primer fragmento está situado dentro de una primera porción de repetición;

(iv) el extremo proximal del segundo fragmento está situado dentro de una segunda porción de repetición y

en el que el primer fragmento y el segundo fragmento están separados por menos de o exactamente 5 aminoácidos en la estructura primaria, en la que los polipéptidos provocan anticuerpos que neutralizan la toxina A y la toxina B y en el que la primera porción de repetición y la segunda porción de repetición tienen una identidad de secuencia entre sí superior a 50 %.

El término polipéptido se refiere a una secuencia contigua de aminoácidos.

El término "dominio de repetición de la toxina A" se refiere al dominio C-terminal de la proteína de la toxina A de *C.difficile*, que comprende secuencias repetidas. Este dominio se refiere a los aminoácidos 1832-2710 de la toxina A de la cepa VPI10463 (ATCC43255) y sus equivalentes en una cepa diferente, la secuencia de aminoácidos 1832-2710 de la cepa VPI10463 (ATCC43255) corresponde a los aminoácidos 1832-2710 de la SEQ ID NO:1.

El término "dominio de repetición de la toxina B" se refiere al dominio C-terminal de la proteína de la toxina B de *C. difficile*. Este dominio se refiere a los aminoácidos 1834-2366 de la cepa VPI10463 (ATCC43255) y sus equivalentes en una cepa diferente, la secuencia de aminoácidos 1834-2366 de la cepa VPI10463 (ATCC43255) corresponde a los aminoácidos 1834-2366 de la SEQ ID NO:2.

- 5 Las toxinas A y B *C. difficile* son proteínas conservadas, sin embargo, la secuencia difiere una pequeña cantidad entre las cepas, además, la secuencia de aminoácidos para las toxinas A y B en diferentes cepas puede diferir en el número de aminoácidos.

Por lo tanto, la invención contempla el término dominio de repetición de la toxina A y/o dominio de repetición de la toxina B para referirse a una secuencia que es una variante con 90 %, 95 %, 98 %, 99 % o 100 % de identidad de secuencia con los aminoácidos 1832-2710 SEQ ID NO: 1 o una variante con 90 %, 95 %, 98 %, 99 % o 100 % de identidad de secuencia con los aminoácidos 1834-2366 SEQ ID NO: 2. En una realización, una "variante" es un polipéptido que varía de los polipéptidos de referencia por sustituciones de aminoácidos conservativas, por lo que un resto se sustituye por otro con las mismas propiedades físico-químicas. Típicamente, dichas sustituciones están entre Ala, Val, Leu y Ile; entre Ser y T; entre los restos ácidos Asp y Glu; entre Asn y Gin, y entre los restos básicos L y Arg; o restos aromáticos F y Tir. En una realización, un "fragmento" es un polipéptido que comprende una porción contigua de al menos 250 aminoácidos de un polipéptido.

Además, la numeración de aminoácidos puede diferir entre los dominios C-terminal de la toxina A (o toxina B) de una cepa y la toxina A (o toxina B) de otra cepa. Por esta razón, el término "equivalentes en una cepa diferente" se refiere a aminoácidos que corresponden a los de una cepa de referencia (p. ej., *C. difficile* VPI10463), pero que se encuentran en una toxina de una cepa diferente y que por tanto pueden numerarse de manera diferente. Se puede determinar una región de aminoácidos "equivalentes" alineando las secuencias de las toxinas de las diferentes cepas. Los números de aminoácidos proporcionados en todo se refieren a los de la cepa VPI10463.

El término "fragmento" de un polipéptido o proteína se refiere a una porción contigua de al menos 200, 230, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 o 600 aminoácidos de ese polipéptido o proteína. El término "primer fragmento" se refiere a una porción contigua de al menos 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 o 600 aminoácidos del dominio de repetición de la toxina A. El término "segundo fragmento" se refiere a una porción contigua de al menos 200, 230, 250, 280, 300, 350, 400, 450 o 500 aminoácidos del dominio de repetición de la toxina B.

El término "primer extremo proximal" se refiere al extremo del primer fragmento (fragmento ToxA) que está unido covalentemente al segundo fragmento (fragmento ToxB) o unido covalentemente a una secuencia engarzadora entre el primer y el segundo fragmento. El término "segundo extremo proximal" se refiere al extremo del segundo fragmento que está más cerca del primer fragmento en la estructura primaria (secuencia de aminoácidos).

La Figura 2 representa la organización de los dominios C-terminal de ToxA y ToxB. El dominio C-terminal de la toxina A está constituido por 8 porciones de repetición (porción de repetición designada I, porción de repetición II, porción de repetición III, porción de repetición IV, porción de repetición V, porción de repetición VI, porción de repetición VII y porción de repetición VIII), cada una de estas porciones de repetición se puede dividir aun más en repeticiones cortas (RC) que se representan como recuadros blancos en la Figura 2 y repeticiones largas (RL) que se representan como recuadros negros en la Figura 2 (excepto para la porción de repetición VIII de ToxA que no tiene una repetición larga). Cada una de las repeticiones largas tiene cierta similitud estructural y de secuencia con las otras repeticiones largas. De forma similar, las repeticiones cortas tienen cierta similitud de secuencia y estructural entre sí. De forma similar, el dominio C-terminal de la toxina B está constituido por 5 porciones de repetición subdivididas en RC y RL. Cada porción de repetición contiene una RL y entre 2 y 5 RC (excepto para la porción de repetición V de ToxB que no tiene una repetición larga). Para los fines de la divulgación, la expresión "una porción de repetición" se refiere a una de las ocho porciones de repetición de ToxA (designadas I, II, III, IV, V, VI, VII y VIII) o una de las cinco porciones de repetición de ToxB (designadas I, II, III, IV o V). Tal como se usa en el presente documento, el término "primera porción de repetición" se refiere a una porción de repetición (o una porción de repetición parcial) del dominio de repetición de la toxina A. El término "segunda porción de repetición" se refiere a una porción de repetición (o una porción de repetición parcial) del dominio de repetición de la toxina B. Para los fines de la divulgación, el término "repetición larga" se refiere a uno de los dominios de RL representados como recuadros negros en la Figura 2. Para los fines de la divulgación, el término "repetición corta" se refiere a uno de los dominios RC representados como recuadros blancos en la Figura 2.

Así, por ejemplo, la porción de repetición I de ToxA contiene tres RC y una RL, que pueden denominarse la primera RCI de ToxA, la segunda RCI de ToxA, la tercera RCI de ToxA y la RLI de ToxA, respectivamente.

El primer extremo proximal se considera dentro de una "porción de repetición" si el primer fragmento termina en un aminoácido que está dentro de esa porción de repetición (es decir, el primer extremo proximal contiene solo parte de la secuencia de la porción repetida). De forma similar, el segundo extremo proximal se considera dentro de una "porción de repetición" si el segundo fragmento termina en un aminoácido que está dentro de esa porción de repetición. Por ejemplo, el primer extremo proximal está dentro de la porción de repetición I de ToxA si el primer fragmento termina con uno cualquiera de los aminoácidos 1832-1924 (inclusive) de VPI10463 o su equivalente en

otra cepa. El primer extremo proximal está dentro de una "repetición larga" o una "repetición corta" si el primer fragmento termina en un aminoácido que está dentro de una "repetición larga" o una "repetición corta", de forma similar, el segundo extremo proximal está dentro de un "repetición larga" o una "repetición corta" si el segundo fragmento termina en un aminoácido que está dentro de una "repetición larga" o una "repetición corta".

- 5 Las posiciones de los aminoácidos de cada dominio se han definido para la toxina A y la toxina B de la cepa VPI10463 (ATCC43255). Estas son las siguientes

Tabla 1

| Nombre   |     | Posición inicial | Posición final |
|----------|-----|------------------|----------------|
| ToxA_I   | RC1 | 1832             | 1852           |
|          | RC2 | 1853             | 1873           |
|          | RC3 | 1874             | 1893           |
|          | RL  | 1894             | 1924           |
| ToxA_II  | RC1 | 1925             | 1944           |
|          | RC2 | 1945             | 1965           |
|          | RC3 | 1966             | 1986           |
|          | RC4 | 1987             | 2007           |
|          | RC5 | 2008             | 2027           |
|          | RL  | 2028             | 2058           |
| ToxA_III | RC1 | 2059             | 2078           |
|          | RC2 | 2079             | 2099           |
|          | RC3 | 2100             | 2120           |
|          | RC4 | 2121             | 2141           |
|          | RC5 | 2142             | 2161           |
|          | RL  | 2162             | 2192           |
| ToxA_IV  | RC1 | 2193             | 2212           |
|          | RC2 | 2213             | 2233           |
|          | RC3 | 2234             | 2253           |
|          | RC4 | 2254             | 2275           |
|          | RL  | 2276             | 2306           |
| ToxA_V   | RC1 | 2307             | 2326           |
|          | RC2 | 2327             | 2347           |
|          | RC3 | 2348             | 2368           |
|          | RC4 | 2369             | 2389           |
|          | RC5 | 2390             | 2409           |
|          | RL  | 2410             | 2440           |
| ToxA_VI  | RC1 | 2441             | 2460           |
|          | RC2 | 2461             | 2481           |

(continuación)

| Nombre    |     | Posición inicial | Posición final |
|-----------|-----|------------------|----------------|
|           | RC3 | 2482             | 2502           |
|           | RC4 | 2503             | 2522           |
|           | RL  | 2523             | 2553           |
| ToxA_VII  | RC1 | 2554             | 2573           |
|           | RC2 | 2574             | 2594           |
|           | RC3 | 2595             | 2613           |
|           | RL  | 2614             | 2644           |
| ToxA_VIII | RC1 | 2645             | 2664           |
|           | RC2 | 2665             | 2686           |
|           | RC3 | 2687             | 2710           |
| ToxB_I    | RC1 | 1834             | 1854           |
|           | RC2 | 1855             | 1876           |
|           | RC3 | 1877             | 1896           |
|           | RL  | 1897             | 1926           |
| ToxB_II   | RC1 | 1927             | 1946           |
|           | RC2 | 1947             | 1967           |
|           | RC3 | 1968             | 1987           |
|           | RC4 | 1988             | 2007           |
|           | RC5 | 2008             | 2027           |
|           | RL  | 2028             | 2057           |
| ToxB_III  | RC1 | 2058             | 2078           |
|           | RC2 | 2079             | 2099           |
|           | RC3 | 2100             | 2119           |
|           | RC4 | 2120             | 2139           |
|           | RC5 | 2140             | 2159           |
|           | RL  | 2160             | 2189           |
| ToxB_IV   | RC1 | 2190             | 2212           |
|           | RC2 | 2213             | 2233           |
|           | RC3 | 2234             | 2253           |
|           | RC4 | 2254             | 2273           |
|           | RC5 | 2274             | 2293           |
|           | RL  | 2294             | 2323           |
| ToxB_V    | RC1 | 2324             | 2343           |
|           | RC2 | 2344             | 2366           |

Por esta razón, el término "porción de repetición" puede referirse a los aminoácidos 1832-1924, 1925-2058, 2059-2192, 2193-2306, 2307-2440, 2441-2553, 2554-2644 o 2645-2710 de la toxina A (SEQ ID NO: 1), o a los aminoácidos 1834-1926, 1927-2057, 2058-2189, 2190-2323 o 2324-2366 de la toxina B (SEQ ID NO: 2) o a sus equivalentes en una cepa diferente de *C.difficile*.

5 Por esta razón, el término "repetición corta" puede referirse a los aminoácidos 1832-1852, 1853-1873, 1874-1893, 1925-1944 1945-1965, 1966-1986, 1987-2007, 2008-2027, 2059-2078, 2079 -2099, 2100-2120, 2121-2141, 2142-2161, 2193-2212, 2213-2233, 2234-2253, 2254-2275, 2307-2326, 2327-2347, 2348-2368, 2369-2389, 2390-2409, 2441-2460, 2461-2481, 2482-2502, 2503-2522, 2554-2573, 2574-2594, 2595-2613, 2645-2664, 2665-2686 o 2687-2710 de la toxina A (SEQ ID NO: 1) o a los aminoácidos 1834-1854, 1855-1876, 1877-1896, 1927-1946, 1947-1967, 1968-1987, 1988-2007, 2008-2027, 2058-2078, 2079-2099, 2100-2119, 2120-2139, 2140-2159, 2190-2212, 2213-2233, 2234-2253, 2254-2273, 2274-2293, 2324-2343 o 2344-2366 de la toxina B (SEQ ID NO: 2) o a sus equivalentes en una cepa diferente de *C.difficile*.

10 De forma similar, el término "repetición larga" puede referirse a los aminoácidos 1894-1924, 2028-2058, 2162-2192, 2276-2306, 2410-2440, 2523-2553, o 2614-2644 de la toxina A (SEQ ID NO: 1), o a los aminoácidos 1897-1926, 2028-2057, 2160-2189 o 2294-2323 de la toxina B (SEQ ID NO: 2) o a sus equivalentes en una cepa diferente de *C.difficile*.

15 Los polipéptidos de la invención pueden ser parte de una proteína más grande tal como un precursor o una proteína de fusión. A menudo es ventajoso incluir una secuencia de aminoácidos adicional que contenga secuencias que ayuden en la purificación tal como múltiples restos de histidina, o una secuencia adicional para la estabilidad durante la producción recombinante. Además, también se considera la adición de un polipéptido exógeno o una cola lipídica o secuencias de polinucleótidos para aumentar el potencial inmunogénico de la molécula final.

La palabra "adyacente" significa separada por menos de o exactamente 20, 15, 10, 8, 5, 2, 1 o 0 aminoácidos en la estructura primaria.

20 Los fragmentos pueden colocarse de manera que el extremo N del primer fragmento sea adyacente al extremo C del segundo fragmento, como alternativa, el extremo C del primer fragmento puede ser adyacente al extremo N del segundo fragmento, o el extremo C del primer fragmento puede ser adyacente al extremo C del segundo fragmento, o el extremo N del primer fragmento puede estar adyacente al extremo N del segundo fragmento.

Dos secuencias tendrán "similitud de secuencia entre sí" si tienen una identidad de secuencia superior a 50 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o 100 %.

30 El término "identidad" se conoce en la técnica, es una relación entre dos o más secuencias polipeptídicas o dos o más secuencias polinucleotídicas, según sea el caso, como se determina mediante la comparación de las secuencias. En la técnica, "identidad" también significa el grado de relación de secuencia entre las secuencias polipeptídicas o polinucleotídicas, según sea el caso, como se determina mediante el emparejamiento entre cadenas de dichas secuencias. La "identidad" se puede calcular fácilmente por procedimientos conocidos, que incluyen, pero sin limitación, los descritos en (Computational Molecular Biology, Lesk, AM, ed., Oxford University Press, Nueva York, 1988; Biocomputing: Informática y Genoma Projects, Smith, DW, ed., Academic Press, Nueva York, 1993; Computer Analysis of Sequence Data, Parte I, Griffin, AM, y Griffin, HG, eds., Humana Press, Nueva Jersey, 1994; Sequence Analysis in Molecular Biology, von Heine, G., Academic Press, 1987 y Sequence Analysis Primer, Gribskov, M. y Devereux, J., eds., M Stockton Press, Nueva York, 1991; y Carillo, H., y Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Los procedimientos para determinar la identidad están diseñados para proporcionar la mayor coincidencia entre las secuencias ensayadas. Además, los procedimientos para determinar la identidad están codificados en programas informáticos disponibles al público. Los procedimientos de programas informáticos para determinar la identidad entre dos secuencias incluyen, pero sin limitación, el programa Needle BLASTP, BLASTN (Altschul, SF y col., J. Molec. Biol. 215: 403-410 (1990), y FASTA (Pearson y Lipman Proc. Natl. Acad. Sci. EE.UU. 85: 2444-2448 (1988). La familia de programas BLAST está disponible al público por NCBI y otras fuentes (Manual BLAST, Altschul, S., y col., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., y col., J. Mol. Biol. 215 403-410 (1990). El conocido algoritmo de Smith Waterman también puede usarse para determinar la identidad.

Los parámetros para la comparación de las secuencia polipeptídicas incluyen los siguientes:

Algoritmo: Needleman y Wunsch, J. Mol Biol. 48: 443-453 (1970)

50 Matriz de comparación: BLOSSUM62 de Henikoff y Henikoff,  
Proc. Natl. Acad. Sci. EE.UU. 89:10915-10919 (1992)

Penalización por hueco: 10

Penalización por extensión de hueco: 0,5

Un programa útil con estos parámetros está disponible al público como el programa 'needle' del paquete EMBOSS

(Rice P y col., Trends in Genetics 2000 col. 16(6):276-277). Los parámetros anteriormente mencionados son los parámetros predeterminados para las comparaciones de péptidos (junto con la ausencia de penalizaciones para los huecos finales).

5 En una realización de la divulgación, la primera porción repetida y la segunda porción repetida tienen una alta similitud estructural entre sí. Se puede considerar que dos secuencias tienen una alta similitud estructural cuando su porcentaje de identidad es superior al 40 %, 45 %, 50 % o 60 % (M. Marty-Renom y col., Annu. Rev. Biophys. Biomol. Struct. 2000 vol 29:291-325). La presencia de una alta similitud estructural puede determinarse comparando las dos secuencias usando los softwares SwissModel y SwissPDB Viewer.

10 En una realización de la divulgación, el polipéptido de la invención provoca anticuerpos que neutralizan la toxina A o la toxina B. En una realización adicional de la divulgación, el polipéptido provoca anticuerpos que neutralizan la toxina A. En una realización adicional de la divulgación, el polipéptido provoca anticuerpos que neutralizan la toxina B. En una realización adicional, el polipéptido provoca anticuerpos que neutralizan la toxina A y la toxina B. La expresión "provoca anticuerpos neutralizantes" significa que cuando las composiciones se usan para inmunizar a un mamífero, por ejemplo un ratón, un cobaya o un ser humano, el mamífero genera anticuerpos neutralizantes.

15 Se puede medir si un polipéptido provoca anticuerpos neutralizantes contra una toxina inmunizando ratones con una composición inmunogénica que comprende el polipéptido, recogiendo sueros y analizando los títulos de antitoxina de los sueros usando ELISA. Los sueros deben compararse con una muestra de referencia obtenida de ratones que no han sido inmunizados. Un ejemplo de esta técnica se puede encontrar en el Ejemplo 6. El polipéptido de la invención provoca anticuerpos que neutralizan la toxina A si los sueros contra el polipéptido dan una lectura de ELISA mayor  
20 que 10 %, 20 %, 30 %, 50 %, 70 %, 80 %, 90 % o un 100 % más alta que la muestra de referencia.

En una realización adicional, el polipéptido de la invención provoca una respuesta inmune protectora en un huésped mamífero contra cepas de *C.difficile*. La expresión "provocar una respuesta inmune protectora" significa que cuando la composición inmunogénica de la invención se usa para inmunizar un mamífero tal como un ratón, una cobaya o un ser humano, el mamífero genera anticuerpos que pueden proteger al mamífero de la muerte causada por  
25 *C.difficile*. En una realización, el huésped mamífero se selecciona del grupo que consiste en un ratón, un conejo, una cobaya, un mono, un primate no humano o un ser humano. En una realización, el huésped mamífero es un ratón. En una realización adicional, el huésped mamífero es un ser humano.

Se puede determinar si un polipéptido provoca una respuesta inmune protectora en un huésped mamífero contra cepas de *C.difficile* usando un ensayo de desafío. En dicho ensayo, el huésped mamífero se vacuna con el  
30 polipéptido y se desafía por exposición a *C.difficile*, el tiempo en el que el mamífero sobrevive después de que el desafío se compara con el tiempo que sobrevive un mamífero de referencia que no ha sido inmunizado con el polipéptido. Un polipéptido provoca una respuesta inmune protectora si un mamífero inmunizado con el polipéptido sobrevive al menos 10 %, 20 %, 30 %, 50 %, 70 %, 80 %, 90 % o 100 % más que un mamífero de referencia que no ha sido inmunizado después del desafío con *C.difficile*. En una realización, el polipéptido de la invención provoca  
35 una respuesta inmune protectora contra cepas de *C.difficile* en un mamífero seleccionado del grupo que consiste en un ratón, una cobaya, un mono o un ser humano. En una realización, el mamífero es un ratón, en una realización adicional, el mamífero es un ser humano.

La estructura nativa de los dominios C-terminal (repetidos) de las toxinas A y B consiste en una estructura de tipo solenoide  $\beta$  extendida. Esta estructura consiste principalmente en estructuras de lámina  $\beta$ , con una minoría de  
40 estructuras  $\alpha$  helicoidales como se ve en Ho y col., (PNAS 102:18373-18378 (2005)). Las estructuras secundarias presentes se pueden determinar usando dicroísmo circular. Por ejemplo, midiendo la forma y la magnitud de los espectros de CD en la región de UV lejano (190-250 nm) y comparando los resultados con los de las estructuras conocidas. Esto se puede llevar a cabo usando un trazado óptico de 0,01 cm desde 178 hasta 250 nm, con una resolución de 1 nm y un ancho de banda en un espectropolarímetro Jasco J-720, por ejemplo como se ve en el  
45 Ejemplo 5 a continuación.

En una realización, el primer fragmento comprende menos de 28 %, 25 %, 23 %, 20 %, 18 %, 15 %, 10 % o 7 % de estructura secundaria alfa helicoidal. En una realización, el segundo fragmento comprende menos de 28 %, 25 %, 23 %, 20 %, 18 %, 15 %, 10 % o 7 % de estructura secundaria alfa helicoidal. En una realización adicional, tanto el  
50 primer fragmento como el segundo fragmento comprenden menos de 28 %, 25 %, 23 %, 20 %, 18 %, 15 %, 10 % o 7 % de estructura secundaria alfa helicoidal.

En una realización, el primer fragmento comprende más de 20 %, 25 %, 28 %, 30 %, 33 %, 35 %, 38 %, 40 % o 42 % de estructura de lámina beta. En una realización, el segundo fragmento comprende más de 20 %, 25 %, 28 %, 30 %, 33 %, 35 %, 38 %, 40 % o 42 % de estructura de lámina beta. En una realización adicional, tanto el primer  
55 fragmento como el segundo fragmento comprenden más de 20 %, 25 %, 28 %, 30 %, 33 %, 35 %, 38 %, 40 % o 42 % de estructura de lámina beta.

En una realización, el primer extremo proximal está dentro de la porción de repetición V (aminoácidos 2307-2440 de SEQ ID NO: 1 o sus equivalentes en una cepa diferente), VI (aminoácidos 2441-2553 de SEQ ID NO: 1 o sus equivalentes en una cepa diferente), VII (aminoácidos 2554-2644 de SEQ ID NO: 1 o sus equivalentes en una cepa

diferente) o VIII (aminoácidos 2645-2710 de SEQ ID NO: 1 o sus equivalentes en una cepa diferente) de la toxina A. En una realización adicional, el primer extremo proximal está dentro de la porción de repetición VII (aminoácidos 2554-2644 de SEQ ID NO: 1 o sus equivalentes en una cepa diferente) de la toxina A. En una realización adicional, el primer extremo proximal está dentro de la porción de repetición VIII (aminoácidos 2645-2710 de SEQ ID NO: 1 o sus equivalentes en una cepa diferente) de la toxina A.

En una realización, el segundo extremo proximal está dentro de la porción de repetición I (aminoácidos 1834-1926 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente), II (aminoácidos 1927-2057 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente) o III (aminoácidos 2058-2189 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente) de la toxina B. En una realización adicional, el segundo extremo proximal está dentro de la porción de repetición II (aminoácidos 1927-2057 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente) de la toxina B. En una realización adicional, el segundo extremo proximal está dentro de la porción de repetición I (aminoácidos 1834-1926 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente) de la toxina B.

En una realización, el primer extremo proximal está dentro de una repetición larga. El primer extremo proximal puede estar dentro de la repetición larga V de la toxina A (aminoácidos 2410-2440 de la SEQ ID NO: 1 o sus equivalentes en una cepa diferente), o dentro de la repetición larga VI de la toxina A (aminoácidos 2523-2553 de la SEQ ID NO: 1 o sus equivalentes en una cepa diferente), o dentro de la repetición larga VII de la toxina A (aminoácidos 2614-2644 de la SEQ ID NO: 1 o sus equivalentes en una cepa diferente).

En una realización, el segundo extremo proximal está dentro de una repetición larga. El segundo extremo proximal puede estar dentro de la repetición larga I de la toxina B (aminoácidos 1897-1926 de la SEQ ID NO: 2 o sus equivalentes en una cepa diferente), o dentro de la repetición larga II de la toxina B (aminoácidos 2028-2057 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente), o dentro de una repetición larga III de la toxina B (aminoácidos 2160-2189 de la SEQ ID NO: 2 o sus equivalentes en una cepa diferente).

En una realización adicional, el primer extremo proximal y el segundo extremo proximal están ambos dentro de repeticiones largas. En una realización, el primer extremo proximal está dentro de la repetición larga V de la toxina A (aminoácidos 2410-2440 de SEQ ID NO: 1 o sus equivalentes en una cepa diferente), o dentro de la repetición larga VI de la toxina A (aminoácidos 2523-2553 de SEQ ID NO: 1 o sus equivalentes en una cepa diferente), o dentro de la repetición larga VII de la toxina A (aminoácidos 2614-2644 de SEQ ID NO: 1 o sus equivalentes en una cepa diferente) y el segundo extremo proximal está dentro de la repetición larga I de la toxina B (aminoácidos 1897-1926 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente), o dentro de la repetición larga II de la toxina B (aminoácidos 2028-2057 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente), o dentro de la repetición larga III de la toxina B (aminoácidos 2160-2189 de la SEQ ID NO: 2 o sus equivalentes en una cepa diferente). En una realización, el primer extremo proximal está dentro de la repetición larga VII de la toxina A (aminoácidos 2614-2644 de la SEQ ID NO: 1 o sus equivalentes en una cepa diferente) y el segundo extremo proximal está dentro de la repetición larga II de la toxina B (aminoácidos 2028-2057 de SEQ ID NO: 2 o sus equivalentes en una cepa diferente).

En una realización, el primer extremo proximal está dentro de los aminoácidos 2620-2660 de la toxina A. En una realización, el segundo extremo proximal está dentro de los aminoácidos 2030-2050 de la toxina B. En una realización adicional, el primer extremo proximal está dentro de los aminoácidos 2620-2660 de toxina A y el segundo extremo proximal está dentro de los aminoácidos 2030-2050 de la toxina B.

En una realización, el primer fragmento comprende al menos 100, 150, 180, 200, 240, 250, 280, 300, 330, 350, 380, 400, 430, 450, 480, 500 o 530 aminoácidos. En una realización, el segundo fragmento comprende al menos 100, 130, 150, 180, 200, 230, 250, 270, 300, 330, 350, 390 o 400 aminoácidos.

En una realización, el polipéptido comprende además un engarce. Este engarce puede estar entre el primer extremo proximal y el segundo extremo proximal, como alternativa, el engarce puede unir los extremos distales del primer fragmento y/o el segundo fragmento a una secuencia adicional de aminoácidos.

En un aspecto adicional de la divulgación, puede emplearse una secuencia de engarce peptídico para separar el primer fragmento y el segundo fragmento. Dicha secuencia de engarce peptídico se incorpora a la proteína de fusión usando técnicas convencionales bien conocidas en la técnica. Las secuencias de engarce peptídico adecuadas pueden elegirse basándose en los siguientes factores: (1) su capacidad para adoptar una conformación extendida flexible; (2) su incapacidad para adoptar una estructura secundaria que pudiera interactuar con epítopos funcionales en el primer fragmento y/o los segundos fragmentos; y (3) la ausencia de restos hidrófobos o cargados que pudieran reaccionar con los epítopos funcionales de ToxA y/o ToxB. Las secuencias del engarce de péptidos pueden contener restos de Gli, Asn y Ser. Otros aminoácidos prácticamente neutros, tales como T y Ala, también pueden usarse en la secuencia engarzadora. Las secuencias de aminoácidos que pueden ser útiles empleadas como engarces incluyen aquellas desveladas en Maratea y col., Gene 40:39-46 (1985); Murphy y col., Proc. Natl. Acad. Sci. EE.UU. 83: 8258-8262 (1986); Patente de EE.UU. n.º 4.935.233 y Patente de EE.UU. n.º 4.751.180. La secuencia engarzadora puede tener, generalmente, de 1 a aproximadamente 50 aminoácidos de longitud.

En una realización de la divulgación, el engarce comprende entre 1-20, 1-15, 1-10, 1-5, 1-2, 5-20, 5-15, 5-10, 10-20,

o 10-15 aminoácidos. En una realización, el engarce es un engarce de glicina, el engarce puede comprender restos de glicina contiguos múltiples (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, o 20), o como alternativa, el engarce puede comprender algunos restos de glicina y algunos restos de otros aminoácidos tales como alanina. En una realización adicional, el engarce comprende un único resto de glicina.

- 5 En una realización, el polipéptido de la invención es parte de una proteína de fusión más grande. Las proteínas de fusión pueden comprender además aminoácidos que codifican una porción inmunogénica de un antígeno de proteína adicional. Por ejemplo, la proteína de fusión puede comprender además una porción inmunogénica de un antígeno de proteína obtenido o procedente de una bacteria seleccionada del grupo que consiste en *S.pneumoniae*, *H.influenzae*, *N.meningitidis*, *E.coli*, *M.cattarhalis*, *C.tentanii*, *C.diphtheriae*, *B.pertussis*, *S.epidermidis*, enterococos, *S.aureus*, y *Pseudomonas aeruginosa*. En este caso, el engarce puede estar entre el primer fragmento o el segundo fragmento y una porción inmunogénica adicional de un antígeno de proteína.

- 10 El término "porción inmunogénica del mismo" o 'fragmento inmunogénico' se refiere a un fragmento de un polipéptido en el que el fragmento comprende un epítipo que es reconocido por los linfocitos T citotóxicos, linfocitos T auxiliares o células B. De manera adecuada, la porción inmunogénica comprenderá al menos el 30 %, adecuadamente al menos el 50 %, especialmente al menos el 75 % y en particular al menos el 90 % (p. ej., 95 % o 98 %) de los aminoácidos en la secuencia de referencia. La porción inmunogénica adecuadamente comprenderá todas las regiones epítipo de la secuencia de referencia.

### POLINUCLEÓTIDOS

- 20 La invención proporciona además un polinucleótido que codifica el polipéptido de la invención. Para los fines de la invención, el término 'polinucleótido (s)' generalmente se refiere a cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser ARN o ADN no modificado o ARN o ADN modificado incluyendo regiones/formas monocatenarias y bicatenarias.

- 25 El término "polinucleótido que codifica un péptido" tal como se usa en el presente documento abarca polinucleótidos que incluyen una secuencia que codifica un péptido o un polipéptido de la invención. El término también abarca polinucleótidos que incluyen una única región continua o regiones discontinuas que codifican el péptido o polipéptido (por ejemplo, polinucleótidos interrumpidos por fago integrado, una secuencia de inserción integrada, una secuencia de vector integrada, una secuencia de transposón integrada, o debido a la edición de ARN o a la reorganización del ADN genómico) junto con regiones adicionales, que también pueden contener secuencias codificantes y/o no codificantes.

- 30 Se apreciará por los expertos habituales 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 el presente documento. Algunos de estos polinucleótidos llevan una homología mínima con la secuencia de nucleótidos de cualquier gen nativo (es decir, de origen natural). Sin embargo, los polinucleótidos que varían debido a las diferencias en el uso de codones se contemplan específicamente por la presente invención, por ejemplo los polinucleótidos que están optimizados para la selección de codones de seres humanos y/o primates y/o *e. coli*.

- 35 Las secuencias que codifican un polipéptido deseado pueden sintetizarse, en su totalidad o en parte, usando procedimientos químicos bien conocidos en la técnica (véase Caruthers, MH y col., Nucl. Acids Res. Symp. Ser. pág. 215-223 (1980), Horn y col., Nucl. Acids Res. Symp. Ser. pág. 225-232 (1980)). Como alternativa, la propia proteína puede producirse usando procedimientos químicos para sintetizar la secuencia de aminoácidos de un polipéptido, o una porción del mismo. Por ejemplo, la síntesis de péptidos puede realizarse usando diversas técnicas de fase sólida (Roberge y col., Science 269:202-204 (1995)) y la síntesis automática puede conseguirse, por ejemplo, usando el sintetizador de péptidos ASI 431 A (Perkin Elmer, Palo Alto, CA).

- 40 Además, las secuencias de polinucleótidos de la presente invención pueden modificarse por ingeniería usando procedimientos generalmente conocidos en la técnica con el fin de alterar las secuencias codificantes de polipéptidos para una diversidad de razones, incluyendo, pero sin limitación, alteraciones que modifican la clonación, procesamiento y/o expresión del producto génico. Por ejemplo, el barajado de ADN mediante fragmentación aleatoria y el reensamblaje por RCP de fragmentos de genes y oligonucleótidos sintéticos pueden ser usados para diseñar por ingeniería las secuencias de nucleótidos. Además, la mutagénesis dirigida al sitio se puede usar para insertar nuevos sitios de restricción, alterar patrones de glicosilación, cambiar la preferencia de codones, producir variantes de corte y empalme, o introducir mutaciones, y así sucesivamente.

### VECTORES

En un aspecto adicional de la invención, la presente invención se refiere a un vector que opcionalmente comprende un polinucleótido de la invención unido a un promotor inducible de manera que cuando se induce el promotor, un polipéptido codificado por el polinucleótido se expresa.

- 55 Un aspecto adicional de la invención comprende dicho vector en el que el promotor inducible se activa mediante la adición de una cantidad suficiente de IPTG (isopropil  $\beta$ -D-1-tiogalactopiranosido) al medio de crecimiento. Opcionalmente esto es a una concentración de entre 0,1 y 10 mM, 0,1 y 5 mM, 0,1 y 2,5 mM, 0,2 y 10 mM, 0,2 y 5

mM, 0,2 y 2,5 mM, 0,4 y 10 mM, 1 y 10 mM, 1 y 5 mM, 2,5 y 10 mM, 2,5 y 5 mM, 5 y 10 mM. Como alternativa, el promotor puede ser inducido mediante un cambio de temperatura o pH.

#### CÉLULAS HUÉSPED

5 Para la producción recombinante de los polipéptidos de la invención, las células huésped pueden modificarse por ingeniería genética para incorporar sistemas de expresión o porciones de los mismos o polinucleótidos de la invención. La introducción de un polinucleótido en la célula huésped puede efectuarse por procedimientos descritos en muchos manuales de laboratorio convencional, tal como Davis, y col., PROCEDIMIENTOS BÁSICOS EN BIOLOGÍA MOLECULAR, (1986) y Sambrook, y col., CLONACIÓN MOLECULAR: UN MANUAL DE LABORATORIO, 2ª Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989), tales como, 10 transfección con fosfato de calcio, transfección mediada por DEAE-dextrano, transvección, microinyección, transfección mediada por lípidos catiónicos, electroporación, conjugación, transducción, raspado de carga, introducción balística e infección.

Ejemplos representativos de huéspedes adecuados incluyen células bacterianas gram negativas, tales como células de, *E. coli*, *Acinetobacter*, *Actinobacillus*, *Bordetella*, *Brucella*, *Campylobacter*, *Cianobacteria*, *Enterobacter*, *Erwinia*, 15 *Francisella*, *Helicobacter*, *hemophilus*, *Klebsiella*, *Legionella*, *Moraxella*, *Neisseria*, *Pasteurella*, *Proteus*, *Pseudomonas*, *Salmonella*, *Serratia*, *Shigella*, *Treponema*, *Vibrio*, *Yersinia*. En otras realizaciones, la célula huésped es una célula de *Escherichia coli*. Como alternativa, las células bacterianas gram positivas también se pueden usar. Una gran diversidad de sistemas de expresión se puede usar para producir los polipéptidos de la invención. En una realización, el vector se obtiene de plásmidos bacterianos. En general, cualquier sistema o vector adecuado para 20 mantener, propagar o expresar polinucleótidos y/o para expresar un polipéptido en un huésped se puede usar para la expresión en este sentido. La secuencia de ADN adecuada puede insertarse en el sistema de expresión mediante cualquiera de una diversidad de técnicas bien conocidas y habituales, tales como, por ejemplo, las expuestas en Sambrook y col., CLONACIÓN MOLECULAR, UN MANUAL DE LABORATORIO, anteriormente mencionados.

#### COMPOSICIONES INMUNOGÉNICAS Y VACUNAS

25 Se proporciona además una composición inmunogénica que comprende un polipéptido de la invención y un excipiente farmacéuticamente aceptable.

En una realización, la composición inmunogénica comprende además un adyuvante. La elección de un adyuvante adecuado para ser mezclado con toxinas bacterianas o conjugados fabricados usando los procedimientos de la invención está dentro del conocimiento del experto en la técnica. Los adyuvantes adecuados incluyen una sal de 30 aluminio tal como gel de hidróxido de aluminio o fosfato de aluminio o alumbre, pero también pueden ser otras sales metálicas tales como las de calcio, magnesio, hierro o zinc, o puede ser una suspensión insoluble de tirosina acilada, o azúcares acilados, sacáridos derivatizados catiónicos o aniónicos o polifosfocenos.

En una realización, la composición inmunogénica comprende además antígenos adicionales. En una realización, los antígenos adicionales son antígenos procedentes de una bacteria seleccionada de entre el grupo que consiste en 35 *S.pneumoniae*, *H.influenzae*, *N.meningitidis*, *E. coli*, *M.cattarhalis*, tétanos, difteria, tos ferina, *S.epidermidis*, enterococos, *S. aureus* y *Pseudomonas aeruginosa*. En una realización adicional, la composición inmunogénica de la invención puede comprender además antígenos de *C.difficile*, por ejemplo las proteínas de la capa S (documento WO01/73030). Opcionalmente, la composición inmunogénica comprende además un sacárido de *C.difficile*.

Se proporciona además una vacuna que comprende la composición inmunogénica, esta vacuna puede comprender 40 además un excipiente farmacéuticamente aceptable. En un aspecto adicional de la invención, se proporciona una vacuna que comprende la composición inmunogénica de la invención y un adyuvante.

Las preparaciones de vacuna que contienen las composiciones inmunogénicas de la presente invención pueden usarse para proteger a un mamífero susceptible de la infección por *C.difficile* o tratar a un mamífero con una 45 infección de *C.difficile*, mediante la administración de dicha vacuna a través de la vía sistémica o mucosa. Estas administraciones pueden incluir la inyección a través de la vía intramuscular, intraperitoneal, intradérmica o subcutánea; o a través de la administración mucosa a los tractos oral/alimentario, respiratorio, genitourinario. Aunque la vacuna de la invención puede administrarse como una dosis única, los componentes de la misma también pueden ser co-administrados juntos al mismo tiempo o en momentos diferentes (por ejemplo conjugados 50 neumocócicos de sacáridos podrían ser administrados por separado, al mismo tiempo o 1-2 semanas después de la administración del componente de cualquier proteína bacteriana de la vacuna para la coordinación de las respuestas inmunes entre sí). Además de una sola vía de administración, pueden usarse 2 vías diferentes de administración. Por ejemplo, sacáridos o conjugados de sacáridos se pueden administrar por vía intramuscular (IM) o intradérmica (ID) y las proteínas bacterianas se pueden administrar por vía intranasal (IN) o por vía intradérmica (ID). Además, las vacunas de la invención pueden administrarse IM para la sensibilización de las dosis y IN para dosis de refuerzo.

55 El contenido de toxinas en la vacuna típicamente estará en el intervalo de 1-250 µg, preferentemente 5-50 µg, más típicamente en el intervalo de 5-25 µg. Después de una vacunación inicial, los sujetos pueden recibir una o varias inmunizaciones de refuerzo espaciadas adecuadamente. La preparación de vacunas se describe en general en Vaccine Design ("The subunit and adjuvant approach" (eds Powell MF y Newman MJ) (1995) Plenum Press Nueva

York). La encapsulación dentro de liposomas se describe por Fullerton, Patente de EE.UU. n.º 4.235.877.

En un aspecto de la divulgación se proporciona un kit de vacuna, que comprende un vial que contiene una composición inmunogénica de la invención, opcionalmente en forma liofilizada, y que comprende además un vial que contiene un adyuvante como se describe en el presente documento. Se prevé que en este aspecto de la invención, el adyuvante se use para reconstituir la composición inmunogénica liofilizada.

Un aspecto adicional de la divulgación es un procedimiento para evitar o tratar la infección por *C.difficile* que comprende administrar al huésped una dosis inmunoprotectora de la composición inmunogénica o la vacuna o el kit de la invención. En una realización, se proporciona un procedimiento para evitar o tratar los episodios primarios y/o de reaparición de la infección por *C.difficile* que comprende administrar al huésped una dosis inmunoprotectora de la composición inmunogénica o la vacuna o el kit de la divulgación.

Un aspecto adicional de la divulgación es una composición inmunogénica o una vacuna o un kit de la divulgación para su uso en el tratamiento o la prevención de la enfermedad de *C.difficile*. En una realización, se proporciona una composición inmunogénica o una vacuna o un kit de la divulgación para su uso en el tratamiento o prevención de episodios primarios y/o de reaparición de la enfermedad de *C.difficile*.

Un aspecto adicional de la divulgación es el uso de la composición inmunogénica o la vacuna o el kit de la divulgación en la fabricación de un medicamento para el tratamiento o la prevención de la enfermedad de *C.difficile*. En una realización, se proporciona una composición inmunogénica o una vacuna o un kit de la divulgación para su uso en la fabricación de un medicamento para el tratamiento o la prevención de episodios primarios y/o de reaparición de la enfermedad de *C.difficile*.

Alrededor de" o "aproximadamente" se definen como dentro del 10 % más o menos de la cifra proporcionada para los fines de la invención.

Las expresiones "que comprende", "comprender", "comprende" en el presente documento están destinadas por los inventores a ser opcionalmente sustituibles por los términos "que consiste en", "consistir en" y "consiste en", respectivamente, en todos los casos. El término "comprende" significa "incluye". Por tanto, a menos que el contexto requiera otra cosa, la palabra "comprende" y variaciones tales como "comprender" y "que comprende" se entenderá que implican la inclusión de un compuesto o composición indicada (p. ej., ácido nucleico, polipéptido, antígeno) o una etapa o un grupo de compuestos o etapas, pero no la exclusión de cualquier otro compuesto, composición, etapas o grupos de los mismos. La abreviatura, "p. ej.," se obtiene del latín *exempli gratia*, y se usa en el presente documento para indicar un ejemplo no limitante. Por tanto, la abreviatura "p ej.," es sinónimo del término "por ejemplo".

Las realizaciones en el presente documento relacionadas con las "composiciones de vacuna" de la invención también se aplican a realizaciones relacionadas con "composiciones inmunogénicas" de la invención, y viceversa.

A menos que se explique de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que entiende comúnmente un experto habitual en la materia a la que pertenece la presente divulgación. Se pueden encontrar definiciones de términos comunes en biología molecular en Benjamin Lewin, Genes V, publicado por Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew y col., (eds.), The Encyclopedia of Molecular Biology, publicada por Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); y Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, publicado por VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

Los términos singulares "un", "una", y "el", "la" incluyen los referentes en plural a menos que el contexto indique claramente otra cosa. De forma análoga, la palabra "o" pretende incluir la palabra "y", a menos que el contexto indique claramente otra cosa. El término "pluralidad" se refiere a dos o más. Se debe entender además que todos los tamaños de bases o tamaños de aminoácidos, y todos los valores de pesos moleculares o de masas moleculares, proporcionados para los ácidos nucleicos o polipéptidos son aproximados, y se proporcionan para la descripción. Además, las limitaciones numéricas proporcionadas con respecto a las concentraciones o niveles de una sustancia, tal como un antígeno, pueden ser aproximadas.

Con el fin de que la presente invención pueda comprenderse mejor, se exponen los siguientes ejemplos. Estos ejemplos son con fines de ilustración únicamente.

## Ejemplos

### Ejemplo 1: Diseño de cinco fusiones de Tox A-Tox B *C.difficile*

Se diseñaron proteínas de fusión que contienen los fragmentos de los dominios de repetición C-terminal de ToxA y ToxB. Estas fusiones contenían un fragmento del dominio de repetición C-terminal de ToxA y un fragmento del dominio de repetición C-terminal de ToxB, y una unión entre el extremo C-terminal del fragmento de ToxA y el extremo N terminal del fragmento de ToxB. Se diseñaron dos estrategias, en la primera estrategia; la fusión fue diseñada de tal manera que la larga estructura de solenoide se mantuvo en la unión entre los dos fragmentos. En la

segunda estrategia, los dos fragmentos de las fusiones están separados por un engarce para permitir su correcto plegamiento independiente.

La parte C-terminal de ToxA y B está compuesta por secuencias repetidas: repeticiones cortas (RC) y repeticiones largas (RL) (PNAS 2005 vol 102: 18373-18378).

- 5 La conocida estructura 3D parcial para el dominio C-terminal de ToxA (PNAS 2005 Greco y col., vol 102: 18373-18378; Nature Structural & Molecular Biology 2006 vol 13 (5): 460-461; ódigos PDB: 2F6E, 2G7C y 2QJ6).

Los inventores predijeron que hay dos tipos de interacciones importantes entre los restos de la parte C-terminal de ToxA y ToxB. La primera interacción se produce entre los restos contenidos en una RL y su RC anterior y es importante para mantener la estructura de tipo solenoide. El segundo tipo de interacción se produce entre los restos contenidos en una RL y la siguiente RC y esta interacción media la función de unión a carbohidratos de la toxina.

Se definió una nueva repetición RC-RL-RC "estructural-funcional". La estructura de esta repetición se mantuvo intacta en las fusiones diseñadas por los investigadores.

La Figura 2 representa los dominios C-terminales de ToxA y ToxB y el recuadro "RC-RL-RC" definido.

15 En la Tabla 1 se presentan las posiciones de las repeticiones cortas (RC) y las repeticiones largas (RL) de ToxA y ToxB.

En la Tabla 2 se presenta una lista de los recuadros "RC-RL-RC" contenidos en el dominio C-terminal de ToxA y ToxB.

Tabla 2

| Nombre | Posición inicial | Posición final |
|--------|------------------|----------------|
| ToxA_1 | 1874             | 1944           |
| ToxA_2 | 2008             | 2078           |
| ToxA_3 | 2142             | 2212           |
| ToxA_4 | 2254             | 2326           |
| ToxA_5 | 2390             | 2460           |
| ToxA_6 | 2503             | 2573           |
| ToxA_7 | 2595             | 2664           |
| ToxB_1 | 1877             | 1946           |
| ToxB_2 | 2008             | 2078           |
| ToxB_3 | 2140             | 2212           |
| ToxB_4 | 2274             | 2343           |

20 Finalmente, el número de las RC entre dos RL se mantendrá en las fusiones diseñadas para mantener la larga estructura tipo solenoide.

25 Antes del diseño de las uniones para las fusiones, se definieron dos hipótesis de trabajo: primera hipótesis, cuanto más cortas son las fusiones, mejor será la probabilidad de que las fusiones sean sobre expresadas de forma estable; segunda hipótesis, de acuerdo con el concepto de recuadros "RC-RL-RC", la posición de inicio tiene que ser elegida con el fin de garantizar un correcto plegamiento de la primera RC de este recuadro RC-RL-RC anteriormente definido. Por tanto las fusiones comienzan a principios de la RC que precede al recuadro RC-RL-RC. Usando estas dos hipótesis, se analizaron tres posiciones de inicio: los restos 2370, 2234 y 2121 de ToxA.

La posición de inicio 2370 fue excluida. La posición de inicio de 2234 también fue excluida debido a que uno de los restos de las interacciones importantes para la estabilidad estructural de la proteína no se conserva. Por lo tanto, se decidió que toda fusión diseñada comenzará en el resto 2121 de ToxA.

30 Todas las fusiones terminarán en el último resto de ToxB.

Se diseñaron cuatro fusiones (F1-4) con el fin de mantener toda la fusión en una larga estructura de tipo solenoide entre los dos fragmentos de fusión.

Las fusiones 1 (F1) y 2 (F2) fueron diseñadas usando la misma hipótesis. Todas las secuencias de proteínas RC de ToxA y ToxB habían sido comparadas usando un software de alineamiento múltiple (ClustalW-Thompson JD y col., (1994) Nucleic Acids Res., 22, 4673-4680). Las secuencias más similares fueron la tercera RC VIII de ToxA y la tercera RC II de ToxB y la tercera RC III de ToxB. Con el fin de elegir entre estas dos RC de ToxB, se realizó un modelado de homología estructural (usando el SwissModel interfaz-Arnold K y col., (2006) Bioinformática, 22, 195-201) en la parte C-terminal de ToxB usando la estructura 3D conocida del dominio C-terminal parcial de ToxA (código PDB: 2QJ6). Usando la tercera RC VIII de ToxA, la mejor superposición estructural local (se realizó usando SwissPDBViewer-Guex N y col., (1997), Electroforesis 18, 2714-2723) se obtuvo con la tercera RC II de ToxB. Por lo tanto, se diseñaron dos uniones: la primera está entre la tercera RC VIII de ToxA y la cuarta RC II de ToxB (F1) y la segunda está entre la segunda RC VIII de ToxA y la tercera RC II de ToxB (F2). Estas uniones se presentan en la Figura 3 y 4 respectivamente.

Para diseñar la fusión 3 (F3), se realizó una superposición estructural global entre la estructura conocida del dominio C-terminal parcial de ToxA y la estructura prevista del dominio C-terminal de ToxB (usando los softwares SwissModel y SwissPDBViewer). La mejor superposición se encontró entre la RL VII de ToxA y la RL II de ToxB. Por lo tanto, se decidió hacer una unión en esta RL similar. La unión se realizó en primer lugar en una región en la que se conserva la secuencia entre ToxA y ToxB, después de eso, para mantener en la parte de ToxA de la fusión, los restos en interacción con la RC anterior y, por último, para mantener en la parte de ToxB, los restos en interacción con la siguiente RC. Esta unión se presenta en la Figura 5.

Para el diseño de la fusión 4 (F4), el dominio C-terminal de ToxB se dividió en 4 fragmentos y se realizó en ellos un modelado de homología más preciso (SwissModel). La división se realizó con el fin de mantener intactos los recuadros "RC-RL-RC" (cada dominio termina al final de la RC que sigue a una RL). Se realizó una superposición estructural entre las estructuras previstas de estos fragmentos y la estructura tridimensional conocida de ToxA y se obtuvo la mejor superposición estructural de la tercera RC de ToxB (RC I) y la última RC de ToxA (tercera RC VIII). Por lo tanto, la unión se realizó entre la segunda RC VIII de ToxA y la tercera RC I de ToxB. Este diseño se presenta en la Figura 6.

La última fusión (F5) se diseñó con el fin de permitir un plegamiento correcto independiente de los dos fragmentos de la fusión. El engarce se añadió entre el último resto de la secuencia de proteína ToxA y el comienzo de la cuarta RC II de ToxB (siempre teniendo en cuenta la importancia de un recuadro "RC-RL-RC" intacto). Solo un resto exógeno (glicina) se añadió como engarce y se ubicó entre dos glicinas existentes. Por tanto, el engarce también puede describirse que está compuesto por 3 glicinas que rodean la cadena beta conocida (para ToxA) y prevista (para ToxB). Este último diseño se muestra en la Figura 7.

## Ejemplo 2: Clonación expresión y purificación de las proteínas de fusión

### Plásmido de expresión y cepa recombinante

Los genes que codifican las proteínas de fusión de los dominios C-terminal parciales de ToxA y ToxB (SEQ ID NO: 3, 4, 5, 6 y 7) y una etiqueta His se clonaron en el vector de expresión pET24b (+) (Novagen) usando sitios de restricción NdeI/XhoI con procedimientos convencionales. La construcción final se generó mediante la transformación de la cepa de *E. coli* BLR (DE3) con el vector de expresión recombinante de acuerdo con el procedimiento convencional con células tratadas con CaCl<sub>2</sub> (Hanahan D. «Transformación de plásmido por Simanis. » en Glover, DM (Ed), Clonación de ADN. IRL Press London. (1985): págs. 109-135).

### Cepa huésped:

BLR (DE3). BLR es un derivado recA de BL21. Las cepas que tienen la designación (DE3) son lisogénicas para un profago  $\lambda$  que contiene una ARN polimerasa de T7 inducible por IPTG. Los lisógenos de DE3  $\lambda$  están diseñados para la expresión de proteínas a partir de vectores pET Esta cepa también es deficiente en las proteasas lon y ompT.

Genotipo: cepa *E.coli* BLR:: DE3, F<sup>-</sup> *ompT hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm* (DE3)  $\Delta$ (srl-recA)306::Tn10 (Tet<sup>R</sup>)

### Expresión de las proteínas recombinantes:

Un transformante de *E.coli* se extrajo de una placa de agar y se usó para inocular 200 ml de caldo LBT  $\pm$  1 % (p/v) de glucosa + kanamicina (50  $\mu$ g/ml) para obtener una D.O.<sub>600nm</sub> entre 0,1-0,2. Los cultivos se incubaron durante una noche a 37 °C, 250 RPM.

Este cultivo de una noche se diluyó a 1:20 en 500 ml de medio LBT que contenía kanamicina (50  $\mu$ g/ml) y se cultivó a 37 °C a una velocidad de agitación de 250 rpm hasta que la D.O.<sub>620nm</sub> alcanzó 0,5/0,6.

A una D.O.<sub>600nm</sub> de alrededor de 0,6, el cultivo se enfrió antes de inducir la expresión de la proteína recombinante mediante la adición de isopropil  $\beta$ -D-1-tiogalactopiranosido 1 mM (IPTG; EMD Chemicals Inc., número de catálogo: 5815) y se incubó durante una noche a 23 °C, 250 RPM.

Después de la inducción durante una noche (alrededor de 16 horas), se evaluó la D.O.<sub>600nm</sub> después de la inducción

y el cultivo se centrifugó a 14 000 RPM durante 15 minutos y los microgránulos se congelaron a -20 °C por separado.

#### Purificación:

5 El sedimento bacteriano se volvió a suspender en tampón de bicina 20 mM (pH 8,0) que contenía NaCl 500 mM y una mezcla de inhibidor de proteasa (Complete, Roche). Las bacterias se lisaron usando un sistema French Press 138 MPa. Los componentes solubles (sobrenadante) e insolubles (microgránulos) se separaron por centrifugación, por ejemplo, a 20 000 g durante 30 minutos a 4 °C.

10 La proteína marcada con 6-His se purificó en condiciones nativas en IMAC. Los componentes solubles se cargaron en una columna GE (15 ml, por ejemplo) (con carga de Ni) equilibrada previamente con el mismo tampón usado para la resuspensión bacteriana. Después de cargar en la columna, la columna se lavó con el mismo tampón. La elución se realizó usando un tampón de bicina 20 mM (pH 8,0) que contenía NaCl 500 mM y diferentes concentraciones de imidazol (5-600 mM). Después del análisis en gel, se seleccionaron las fracciones más puras, se concentraron y se cargaron en cromatografía SEC para la etapa de purificación adicional.

15 Las fracciones que contienen las proteínas de fusión se seleccionaron sobre la base de la pureza mediante SDS-PAGE y se dializaron frente a tampón de bicina (Bicina 20 mM, NaCl 150 mM, con o sin EDTA 5 mM, pH 8,0). La concentración de proteína se determinó usando el Ensayo de Proteína DC de BioRad. De este modo, las proteínas se combinaron, se esterilizaron por filtración a 0,22 µm, y se almacenaron a -80 °C.

20 Como alternativa, la purificación de IMAC fue precedida por una etapa de purificación de DEAE usando tampón de bicina 2 mM (pH 8,0) para la carga y el lavado, y se eluyó usando un gradiente con el mismo tampón pero con NaCl 1M añadido.

### **Ejemplo 3- Clonación expresión y purificación de los fragmentos separados ToxA y ToxB *C.difficile***

#### Plásmido de expresión y cepa recombinante.

25 Los genes que codifican los fragmentos de proteína ToxA y ToxB (SEQ ID NO: 8 y SEQ ID NO: 9) y una etiqueta His se clonaron en el vector de expresión pET24b(+) (Novagen) usando los sitios de restricción NdeI/XhoI con procedimientos convencionales. La construcción final se generó mediante la transformación de la cepa de *E. coli* BLR (DE3) con el vector de expresión recombinante de acuerdo con el procedimiento convencional con células tratadas con CaCl<sub>2</sub> (Hanahan D. «transformación de plásmido por Simanis. » en Glover, DM (Ed), Clonación de ADN. IRL Press London. (1985): págs. 109-135).

#### Cepa huésped:

30 BLR (DE3). BLR es un derivado recA de BL21. Las cepas que tienen la designación (DE3) son lisogénicas para un profago λ que contiene una ARN polimerasa de T7 inducible por IPTG. Los lisógenos de DE3 λ están diseñados para la expresión de proteínas a partir de vectores pET Esta cepa también es deficiente en las proteasas lon y ompT.

Genotipo: cepa *E.coli* BLR:: DE3, F<sup>-</sup> *ompT* hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) *gal dcm* (DE3) Δ(srl-recA)306::Tn 10 (Tet<sup>R</sup>) Expresión de las proteínas recombinantes:

35 Un transformante de *E.coli* se extrajo de una placa de agar y se usó para inocular 200 ml de caldo LBT ± 1 % (p/v) de glucosa + kanamicina (50 µg/ml) para obtener una D.O.<sub>600nm</sub> entre 0,1-0,2. Los cultivos se incubaron durante una noche a 37 °C, 250 RPM.

Este cultivo de una noche se diluyó a 1:20 en 500 ml de medio LBT que contenía kanamicina (50 µg/ml) y se cultivó a 37 °C a una velocidad de agitación de 250 rpm hasta que la D.O.<sub>620nm</sub> alcanzó 0,5/0,6.

40 A una D.O. a 600 nm de alrededor de 0,6, el cultivo se enfrió antes de inducir la expresión de la proteína recombinante mediante la adición de isopropil β-D-1-tiogalactopiranosido 1 mM (IPTG; EMD Chemicals Inc., número de catálogo: 5815) y se incubó durante una noche a 23 °C, 250 RPM.

45 Después de la inducción durante una noche (alrededor de 16 horas), se evaluó la D.O. a 600nm después de la inducción y el cultivo se centrifugó a 14 000 RPM durante 15 minutos y los microgránulos se congelaron a -20 °C por separado.

#### Purificación:

50 El sedimento bacteriano se volvió a suspender en tampón de bicina 20 mM (pH 8,0) que contenía NaCl 500 mM suplementado con una mezcla de inhibidor de proteasa (Complete, sin EDTA, Roche cat 11873580001) y benzonasa. (Roche cat 1.01695.0001). Las bacterias se lisaron usando un sistema French Press 2 x 138 MPa. Los componentes solubles (sobrenadante) e insolubles (microgránulos) se separaron por centrifugación a 34 000 g o 48 000 g durante 25-30 minutos a 4 °C. El sobrenadante se recogió y se filtró en un filtro de 0,22 µm.

La proteína marcada con 6-His se purificó en condiciones nativas en IMAC. Los componentes solubles se cargaron en una columna GE (por ejemplo, 15 ml) (con carga de Ni) equilibrada previamente con el mismo tampón usado para la resuspensión bacteriana. Después de la carga, la columna se lavó con el mismo tampón.

#### Para Tox A:

- 5 La elución se realizó usando un tampón de bicina 20 mM (pH 8,0) que contenía NaCl 500 mM y diferentes concentraciones de imidazol (5-100 mM). Después del análisis en gel, se seleccionaron las fracciones más puras, se concentraron y se cargaron en cromatografía SEC (SUPERDEX™ 75) para la etapa de purificación adicional en el mismo tampón sin imidazol.

#### Para Tox B:

- 10 Se realizó un segundo lavado con tampón de bicina 20 mM (pH 8,0) que contenía NaCl 500 mM y desoxicolato al 0,5 N % o el mismo tampón con NaCl 150 mM. La elución se realizó usando un tampón de bicina 20 mM (pH 8,0) que contenía NaCl 500 mM y diferentes concentraciones de imidazol (10-500 mM). Después del análisis en gel, se seleccionaron las fracciones más puras, se suplementaron con EDTA 5 mM y se cargaron en cromatografía SEC (SUPERDEX™200) para una etapa de purificación adicional en el mismo tampón con EDTA 5 mM.
- 15 Las fracciones que contienen fragmentos ToxA o ToxB se seleccionaron sobre la base de la pureza mediante SDS-PAGE y se dializaron frente a tampón de bicina (Bicina 20 mM, NaCl 150 mM, pH 8,0), la concentración de proteína se determinó usando el ensayo de proteína RCDC de BioRad. De este modo, las proteínas se combinaron, se esterilizaron por filtración a 0,22 µm, y se almacenaron a -80 °C.

#### **Ejemplo 4- Evaluación del peso molecular de las cinco fusiones ToxA-ToxB *C.difficile***

- 20 La ultracentrifugación analítica se usa para determinar la homogeneidad y la distribución del tamaño en las solución de las diferentes especies dentro de una muestra de proteína midiendo la velocidad a la que las moléculas se mueven en respuesta a una fuerza centrífuga. Esto se basa en el cálculo de los coeficientes de sedimentación de las diferentes especies que se obtienen por el experimento de la velocidad de sedimentación, que dependen de su forma molecular y masa.
- 25 1. Las muestras de proteínas se centrifugan en una ultracentrífuga analítica Beckman-Coulter PROTEOMELAB™ XL-1 a 42 000 RPM después de equilibrar el rotor AN-60Ti a 15 °C.
- a. Proteína de fusión F1, 500 µg/ml, Bicina 20 mM, NaCl 150 mM, pH 8,0
  - b. Proteína de fusión F2, 500 µg/ml, Bicina 20 mM, NaCl 150 mM, pH 8,0
  - c. Proteína de fusión F3, 500 µg/ml, Bicina 20 mM, NaCl 150 mM, pH 8,0
  - 30 d. Proteína de fusión F4, 500 µg/ml, Bicina 20 mM, NaCl 150 mM, pH 8,0
  - e. Proteína de fusión F5, 500 µg/ml, Bicina 20 mM, NaCl 150 mM, pH 8,0
2. Para la recopilación de datos, se registraron 160 exploraciones a 280 nm cada 5 minutos.
3. El análisis de los datos se realizó usando el programa SEDFIT para la determinación de la distribución C(S). La determinación del volumen específico parcial de las proteínas se realizó con el software SEDNTERP a partir de su secuencia de aminoácidos. SEDNTERP también se usó para determinar la viscosidad y la densidad del
- 35 tampón.
4. El peso molecular de las diferentes especies se determinó a partir del diagrama de distribución C(S) (concentración frente a coeficiente de sedimentación), considerando que es una mejor representación de los datos brutos que la distribución C(M) (concentración frente a peso molecular) para caracterizar la distribución de
- 40 tamaño de una mezcla.

La Figura 8 describe la distribución de las fusiones ToxA-ToxB como se determina por ultracentrifugación analítica de la velocidad de sedimentación.

- El peso molecular de la especie principal detectada a partir de la distribución C(S) de las cinco proteínas de fusión ToxA-ToxB corresponde a su forma monomérica. Las mejores relaciones de fricción ajustadas determinadas para
- 45 las cinco fusiones están todas entre 2 y 2,2. Esto puede indicar que las proteínas están presentes en solución como una forma elongada, que sería consistente con la estructura de la proteína.

#### **Ejemplo 5- Evaluación de las estructuras secundarias y terciarias de las fusiones ToxA-ToxB *C.difficile* mediante dicroísmo circular y espectroscopia de fluorescencia**

- El dicroísmo circular se usa para determinar la composición de la estructura secundaria de una proteína midiendo la
- 50 diferencia en la absorción de la luz polarizada a mano izquierda frente a la luz polarizada a mano derecha que se

debe a la asimetría estructural. La forma y la magnitud de los espectros de CD en la región de UV lejano (190-250 nm) son diferentes tanto si una proteína muestra una estructura de lámina beta, alfa-hélice o helicoidal al azar. La abundancia relativa de cada tipo de estructura secundaria en una muestra de proteína determinada se puede calcular por comparación con los espectros de referencia.

- 5 La estructura terciaria de una muestra de proteína puede evaluarse mediante la evaluación de la inmovilización de los aminoácidos aromáticos. La observación de una señal de CD en la región UV cercana (250-50 nm) puede ser atribuible a la polarización de restos de fenilalanina, tirosina y triptófano y es un buen indicativo de que la proteína se pliega en una estructura bien definida.

Se usó el siguiente protocolo:

- 10 1. Los espectros de UV lejano se miden usando un trazado óptico de 0,01 cm desde 178 hasta 250 nm, con una resolución de 1 nm y un ancho de banda en un espectropolarímetro Jasco J-720. La temperatura de la celda se mantiene a 23 °C mediante un bloque de células RTE-111 termostático Peltier. Se mantiene un flujo de nitrógeno de 10 l/min durante las mediciones.
- 15 2. Los espectros de UV cercano se miden usando un trazado óptico de 0,01 cm desde 250 hasta 300 nm, con una resolución de 1 nm y un ancho de banda en un espectropolarímetro Jasco J-720. La temperatura de la celda se mantiene a 23 °C mediante un bloque de células RTE-111 termostático Peltier. Se mantiene un flujo de nitrógeno de 6 l/min durante las mediciones.

- 20 La observación de los espectros de UV lejano (Figura 9) para las cinco proteínas de fusión ToxA-ToxB sugiere un contenido débil de estructuras de hélice alfa y un alto contenido de estructuras de lámina beta. Además, todas las proteínas mostraron un máximo a 230 nm, que es inusual para las proteínas globulares solubles. Esta particularidad ha sido bien caracterizada en la literatura y se asocia con un pequeño grupo de proteínas conocidas por su ausencia de alfa hélice y su alto contenido en lámina beta y aminoácidos aromáticos (Zsila, Analytical Biochemistry, 391 (2009) 154-156). Esas particularidades son coherentes con la estructura que se espera para las proteínas de fusión ToxA-ToxB. Se compararon las estructuras cristalinas de 13 proteínas que muestran los espectros de CD
- 25 característicos con una señal positiva a 230 nm (Banco de Datos de Proteínas). El contenido medio de la estructura secundaria de esas proteínas es de 42 % lámina beta  $\pm$  9 % y 7 % hélice alfa  $\pm$  6 %. Esto indica fuertemente que la firma espectral de las proteínas de fusión ToxA-ToxB es diagnóstico de una lámina beta alta y una proteína que contiene poca hélice alfa.

- 30 La observación de la forma de los espectros de UV cercano (Figura 10) para las cinco proteínas de fusión indica que al menos algunos de los aminoácidos aromáticos están inmovilizados, lo que es un fuerte indicativo de una estructura terciaria compacta y específica. Además, el tratamiento de la proteína con una concentración desnaturante de urea provocó la desaparición de la señal UV cercana, que es un indicativo adicional de que este espectro característico se debe al plegamiento de la proteína.

#### **Ejemplo 6- Inmunización de ratones con fragmentos ToxA o ToxB y Fusiones ToxA-ToxB**

- 35 Se inmunizaron ratones Balb/C con las construcciones descritas en los Ejemplos 2 y 3.

##### Inmunización de ratones

Se inmunizaron grupos de 15 ratones hembra Balb/c en los días 0, 14 y 28 con 3  $\mu$ g o 10  $\mu$ g de los fragmentos separados de la toxA y la toxB (véase el Ejemplo 2) así como con las proteínas de fusión ToxA-ToxB (véase el Ejemplo 3) adyuvante con AS03B. Un grupo de control de 10 ratones se vacunó con AS03B únicamente.

- 40 Los títulos de ELISA anti-ToxA y anti-ToxB se determinaron en sueros individuales recogidos el día 42 (después de III).

Los títulos de inhibición de la hemaglutinación se determinaron en sueros combinados Post III.

##### Respuesta ELISA Anti-ToxA y Anti-ToxB: Protocolo

- 45 Las muestras de los fragmentos de la toxA o toxB se revistieron a 1  $\mu$ g/ml en solución salina tamponada con fosfato (STF) en placas de microtitulación de alta unión (Nunc MAXISORP™), durante una noche a 4 °C. Las placas se bloquearon con STF-ASB 1 % durante 30 min a TA con agitación. Los antisueros de los ratones se diluyeron previamente 1/500 en STF-ASB 0,2 %-TWEEN™ 0,05 % y luego, se realizaron diluciones con factor de dilución dos adicionales en microplacas y se incubaron a TA durante 30 min con agitación. Después del lavado, se detectó el anticuerpo murino unido usando Jackson ImmunoLaboratories Inc. conjugado con peroxidasa IgG anti-ratón de
- 50 cabra (H+L) (ref: 115-035-003) diluido 1: 5000 en STF-ASB 0, 2 %-tween 0,05 %. Los anticuerpos de detección se incubaron durante 30 min a temperatura ambiente (TA) con agitación. El color se desarrolló usando 4 mg de O-fenilendiamina (OFD) +5  $\mu$ l de H<sub>2</sub>O<sub>2</sub> por 10 ml pH 4,5, tampón de citrato 0,1 M durante 15 minutos en la oscuridad a temperatura ambiente. La reacción se detuvo con 50  $\mu$ l de HCl, y la densidad óptica (DO) se leyó a 490 nm relativa a 620 nm.

El nivel de anticuerpos anti-ToxA o anti-ToxB presentes en los sueros se expresó en títulos de punto medio. Se calculó un GMT para las 15 muestras en cada grupo de tratamiento (10 para el grupo de control).

#### Ensayo de inhibición de la hemaglutinación: Protocolo

- 5 Se realizaron diluciones en serie con factor de dilución dos de antisueros combinados en ratones (25 µl) en solución salina tamponada con fosfato (STF) en microplacas con fondo en U de 96 pocillos.

Luego se añadieron 25 µl de Toxina A nativa (0,2 µg/pocillo) y las placas se incubaron a temperatura ambiente durante 30 minutos.

Después de la incubación, se añadieron 50 µl de eritrocitos de conejo purificados diluidos al 2 % en cada pocillo. Las placas se incubaron a 37 °C durante 2 horas.

- 10 Las placas se analizaron visualmente, presentando la hemaglutinación como eritrocitos difusos en el pocillo y observando la inhibición de la hemaglutinación como un punto rojo asentado en el pocillo.

Los títulos de inhibición se definieron como el recíproco de la dilución más alta del suero que inhibe la hemaglutinación.

#### Ensayo de citotoxicidad

- 15 Las células de fibroblastos IMR90 se cultivaron a 37 °C con 5 % de CO<sub>2</sub>, en EMEM + 10 % de suero bovino fetal + 1 % de glutamina + 1 % de antibióticos (penicilina-estreptomicina-anfotericina) y se sembraron en placas de cultivo de tejidos de 96 pocillos a una densidad de 5.10<sup>4</sup> células/pocillo.

Después de 24 h, el medio celular se retiró de los pocillos.

- 20 Se realizaron diluciones en serie con factor de dilución dos de antisueros combinados en ratones (50 µl) en medios celulares.

A continuación, se añaden 50 µl de Toxina B nativa (0,5 ng/ml) y las placas se incuban a 37 °C con 5 % de CO<sub>2</sub>, durante 24 horas.

Las células se observaron después de 24 horas, y se determinó la proporción de células redondeadas.

- 25 Los títulos de inhibición se definieron como el recíproco de la dilución más alta del suero que inhibe el 50 % del redondeo celular.

#### Resultados:

Los resultados de Elisa, que usan anticuerpos ToxA se describen en la Figura 11. Los anticuerpos anti-ToxA se indujeron después de la inmunización con ToxA únicamente, pero también con cada una de las 5 fusiones.

- 30 Las propiedades funcionales de estos anticuerpos se ensayaron en el ensayo de hemaglutinación. Este ensayo solo está adaptado para la evaluación de Tox A ya que no se observa hemaglutinación con Tox B.

Los títulos de inhibición de la hemaglutinación se describen en la Figura 12. Se observó inhibición de la hemaglutinación con los sueros o sueros del fragmento anti-ToxA dirigidos contra cada una de las fusiones ToxA-ToxB.

- 35 También se realizó un ELISA usando anticuerpos ToxB; los resultados de esto se ilustran en la Figura 13. Los anticuerpos anti-ToxB se indujeron después de la inmunización con el fragmento ToxB únicamente, pero también con las fusiones F2, F3 y F4.

Los títulos de inhibición de la citotoxicidad se describen en la Figura 14. Los títulos de inhibición obtenidos usando sueros de ratones inmunizados con el fragmento ToxB o las fusiones ToxA-ToxB fueron mayores que los obtenidos usando sueros de control.

- 40 **Ejemplo 7: Diseño, clonación, expresión y purificación de 4 proteínas de fusión**

Se diseñaron cuatro proteínas de fusión adicionales usando los principios de diseño descritos en el Ejemplo 1, estas se denominaron F54 Gli (SEQ ID NO: 11), F54 Nueva (SEQ ID NO: 13), F5 ToxB (SEQ ID NO: 15) y F52Nueva (SEQ ID NO: 17).

Estas proteínas de fusión se expresaron de acuerdo con el procedimiento descrito en el Ejemplo 2.

45

**Ejemplo 8- Evaluación del peso molecular de las fusiones ToxA-ToxB *C.difficile* descritas en SEQ ID NO:11, SEQ ID NO: 13. SEQ ID NO: 15 y SEQ ID NO: 17**

El peso molecular de las fusiones descritas en SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15 y SEQ ID NO: 17 se determinó como se describe en el Ejemplo 4.

- 5 La Figura 15 describe la distribución de estas cuatro proteínas de fusión adicionales como se determinó por ultracentrifugación analítica de la velocidad de sedimentación.

El peso molecular de la especie principal determinada a partir de la distribución C(S) de las cuatro fusiones de proteínas descritas en SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15 y SEQ ID NO: 17 corresponde a su forma monomérica y todas las proteínas muestran propiedades de sedimentación similares a las fusiones F1 a F5.

10 **Ejemplo 9- Evaluación de las estructuras secundaria y terciaria de las fusiones ToxA-ToxB *C.difficile* descritas en SEQ ID NO:11. SEQ ID NO: 13. SEQ ID NO: 15 y SEQ ID NO: 17**

Las estructuras secundarias y terciarias de las fusiones descritas en SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15 y SEQ ID NO: 17 se evaluaron de acuerdo con el procedimiento descrito en el Ejemplo 5. La CD UV lejana para estas proteínas de fusión se puede encontrar en la Figura 16, y los espectros UV cercanos para estas fusiones se pueden encontrar en la Figura 17.

- 15

El análisis de los espectros CD UV cercanos y lejanos de las proteínas descritas en SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, y SEQ ID NO: 17 muestra que los cuatro tienen la misma estructura de lámina beta alta que las fusiones F1 a F5. Además, la observación de los espectros UV cercanos no muestra una diferencia significativa en la posición de los aminoácidos aromáticos en la estructura terciaria en comparación con las fusiones F1 a F5.

20 **Ejemplo 10- Inmunización de ratones con fusiones ToxA-ToxB**

Se inmunizaron ratones Balb/c con las cuatro construcciones de proteína de fusión F54Gli (SEQ ID NO: 11), F54 Nueva (SEQ ID NO: 13), F5 ToxB (SEQ ID NO: 15) y F52 Nueva (SEQ ID NO: 17) como se describe en el Ejemplo 6.

- 25 Se llevó a cabo un ELISA usando la respuesta ELISA anti-ToxA y anti-ToxB: protocolo descrito en el Ejemplo 6 excepto que en este caso las muestras de los fragmentos de la toxA o toxB se revistieron a 2 µg/ml en solución salina tamponada con fosfato en placas de microtitulación de alta unión. Se realizó un ensayo de inhibición de la hemaglutinación como se describe en el Ejemplo 6. Se realizó un ensayo de la citotoxicidad de la toxB como se describe en el Ejemplo 6. Se realizó un ensayo de la citotoxicidad de la toxA adicional como se describe a continuación.

30 Ensayo de la citotoxicidad de ToxA

Las células HT29 se cultivaron a 37 °C con 5 % de CO<sub>2</sub> en DMEM + 10 % de suero bovino fetal + 1 % de glutamina + 1 % de antibióticos (penicilina-estreptomicina-anfotericina) y se sembraron en placas de cultivo de tejidos de 96 pocillos a una densidad de 5.10<sup>4</sup> células/pocillo.

Después de 24 h, el medio celular se retiró de los pocillos.

- 35 Se realizaron diluciones en serie con factor de dilución dos de antisueros combinados en ratones (50 µl) en medios celulares.

A continuación, se añaden 50 µl de Toxina B nativa (0,15ng/ml) y las placas se incuban a 37 °C con 5 % de CO<sub>2</sub>, durante 48 horas.

Las células se observaron después de 48 horas, y se determinó la proporción de células redondeadas.

- 40 Los resultados de los ensayos ELISA anti-toxA, Elisa anti-toxB, inhibición de la hemaglutinación y citotoxicidad se describen en las Figuras 18, 19, 20, 21 y 22 respectivamente.

## REIVINDICACIONES

1. Un polipéptido que comprende un primer fragmento y un segundo fragmento, en el que
  - (i) el primer fragmento es un fragmento del dominio de repetición de la toxina A y comprende al menos 100 aminoácidos;
  - (ii) el segundo fragmento es un fragmento del dominio de repetición de la toxina B y comprende al menos 100 aminoácidos;
  - (iii) el extremo proximal del primer fragmento está situado dentro de una primera porción de repetición;
  - (iv) el extremo proximal del segundo fragmento está situado dentro de una segunda porción de repetición y
 en el que el primer fragmento y el segundo fragmento están separados por menos de o exactamente 5 aminoácidos en la estructura primaria, en el que el polipéptido provoca anticuerpos que neutralizan la toxina A y la toxina B y en el que la primera porción de repetición y la segunda porción de repetición tienen una identidad de secuencia entre sí superior al 50 %.
2. El polipéptido de la reivindicación 1, en el que la primera porción de repetición y la segunda porción de repetición tienen una alta similitud estructural entre sí, en el que las dos secuencias tienen una alta similitud estructural cuando su porcentaje de identidad es superior al 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % o es el 100 %.
3. El polipéptido de una cualquiera de las reivindicaciones 1-2, en el que el polipéptido provoca una respuesta inmune protectora en un huésped mamífero contra cepas de *C.difficile*.
4. El polipéptido de una cualquiera de las reivindicaciones 1-3, en el que el primer fragmento y el segundo fragmento comprenden menos del 25 %, 20 %, 18 % o 15 % de estructura alfa helicoidal.
5. El polipéptido de una cualquiera de las reivindicaciones 1-4, en el que el primer fragmento y el segundo fragmento comprenden más del 25 %, 30 %, 35 %, 38 % o 40 % de estructura de lámina beta.
6. El polipéptido de una cualquiera de las reivindicaciones 1-5, en el que el primer extremo proximal está dentro de la porción de repetición VII o la porción de repetición VIII de la toxina A.
7. El polipéptido de una cualquiera de las reivindicaciones 1-6, en el que el segundo extremo proximal está dentro de la porción de repetición II o la porción de repetición I de la toxina B.
8. El polipéptido de la reivindicación 6, en el que el primer extremo proximal está dentro de la porción de repetición VII de la toxina A.
9. El polipéptido de la reivindicación 7, en el que el segundo extremo proximal está dentro de la porción de repetición II de la toxina B.
10. El polipéptido de una cualquiera de las reivindicaciones 1-9, en el que el primer extremo proximal está dentro de una repetición larga.
11. El polipéptido de una cualquiera de las reivindicaciones 1-10, en el que el segundo extremo proximal está dentro de una repetición larga.
12. El polipéptido de una cualquiera de las reivindicaciones 1-11, en el que el primer extremo proximal está dentro de la repetición larga VII de la toxina A (aminoácidos 2614-2644).
13. El polipéptido de la reivindicación 12, en el que el primer extremo proximal está dentro de los aminoácidos 2620 y 2660 de la toxina A.
14. El polipéptido de una cualquiera de las reivindicaciones 1-13, en el que el segundo extremo proximal está dentro de la repetición larga II de la toxina B (aminoácidos 2028-2057).
15. El polipéptido de la reivindicación 14, en el que el primer extremo proximal está dentro de los aminoácidos 2030 y 2050 de la toxina B.
16. El polipéptido de cualquiera de las reivindicaciones anteriores, en el que el polipéptido es parte de una proteína de fusión más grande.
17. Un polinucleótido que codifica el polipéptido de una cualquiera de las reivindicaciones 1-16.
18. Un vector que comprende el polinucleótido de la reivindicación 17.
19. El vector de la reivindicación 18 que comprende además un promotor inducible.
20. El vector de la reivindicación 19, en el que el promotor inducible se activa mediante la adición de una cantidad suficiente de IPTG.

21. Una célula huésped que comprende el vector de una cualquiera de las reivindicaciones 18-20 o el polinucleótido de la reivindicación 17.
22. La célula huésped de la reivindicación 21, en la que la célula huésped es una bacteria gram negativa.
23. La célula huésped de la reivindicación 22, en la que la célula huésped es *E.coli*.
- 5 24. Una composición inmunogénica, que comprende el polipéptido de una cualquiera de las reivindicaciones 1-16 y un excipiente farmacéuticamente aceptable.
25. La composición inmunogénica de la reivindicación 24, que comprende además un adyuvante.
26. La composición inmunogénica de una cualquiera de las reivindicaciones 24-25 que comprende además antígenos adicionales.
- 10 27. La composición inmunogénica de la reivindicación 26, en la que los antígenos adicionales son antígenos procedentes de una bacteria seleccionada de del grupo que consiste en *S.pneumoniae*, *H.influenzae*, *N.meningitidis*, *E.coli*, *M.cattarhalis*, tétanos, difteria, tos ferina, *S.epidermidis*, enterococos, *S.aureus*, y *Pseudomonas aeruginosa*.
28. La composición inmunogénica de una cualquiera de las reivindicaciones 24-27, que comprende además un sacárido de *C. difficile*.
- 15 29. Una vacuna que comprende la composición inmunogénica de una cualquiera de las reivindicaciones 24-28 y un excipiente farmacéuticamente aceptable.
30. La composición inmunogénica de una cualquiera de las reivindicaciones 24-28 o la vacuna de la reivindicación 29, para su uso en el tratamiento o prevención de la enfermedad de *C.difficile*.

**FIGURA 1**  
**SEQ ID NO: 1 - secuencia de la toxina A**

MSLISKEELIKLAYSIRPRENEYKTILTNLDEYNKLTNNNNENKYLQKKLNESIDVFMN  
KYKTSSRNRLSNLKKDILKEVILIKNSNTSPVEKNLHFVWIGGEVSDIALEYIKQWADI  
NAEYNIKLWYDSEAFVNTLKKAIVESSTTEALQLLEEEIQNPQFDNMKFYKKRMEFIYD  
RQKRFINYYSQINKPTVPTIDDIKSHLVSEYNRDETVELSYRTNSLRKINSNHGIDIR  
ANSLFTEQELLNIYSQELLNRGNLAAASDIVRLLALKNFGGVYLDVDMPLPGIHSDFKTI  
SRPSSIGLDRWEMIKLEAIMKYKYYINNYTSENFDDKLDQQLKDNFKLIIESKSEKSEIFS  
KLENLNVSDLEIKIAFALGSVINQALISKQGSYLTNLVIEQVKNRYQFLNQHLNPAIESD  
NNFTDTTKIFHDSLFNSATAENSMFLTKIAPYLQVGFMPPEARSTISLSGPGAYASAYYDF  
INLQENTIEKTLKASDLIEFKFPENNLSQLTEQEINSLWSFDQASAKYQFEKYVRDYTG  
SLSEDNGVDFNKNLTALDKNYLLNNKIPSNNVVEEAGSKNYVHYIIQLQGDDISYEATCNLF  
SKNPKNISIIQRNMNESAKSYFLSDDGESILELNKYRIPERLKNKEKVVTFIGHGKDEF  
NTSEFARLSVDSLSNEISSFLDTIKLDISPKNVEVNLLGCNMFSDYDFNVEETYPGKLLLS  
IMDKITSTLPDVNKNISITIGANQYEVRIINSEGRKELLAHSGKWINKEEAIMSDLSKEYI  
FFDSIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIESSIGDYIYY  
EKLEPVKNIIHNSIDDLIDFNLENVSDLEYELKLLNNLDEKYLISFIDENSTYSV  
RFINKSNGESVYVETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNT  
LNAAFFTQSLIDYSSNKDVIINDISTSVKQLYAQIFSTGINTIYDSIQIVNLTISNAVNDT  
INVLPITITEGIPVSTILDGINLCAAIKELLDEHDPLLKKELEAKVGVLAINMSLSIAAT  
VASIVGIGAETIFLLPIAGISAGIPSLVNNELIHDKATSVVNYFNHLSSESKYGLPKT  
EDDKILVPIDDLVISEIDFNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIP  
SLSIYSAIGIETENLDFSKKIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLY  
PGKEYWRFYAFFDYATTLKPVYEDTNIKIKLDDKTRNFIMPTITTTNEIRNKLSSYFSDGA  
GGTYSLLLSSYPISNTINLSKDDLWIFNIDNEVREISIEINGTIKKGLIKDVLSKIDINK  
NKLIIGNQTIIDFGSDIDNKDRIIFLTCELDDKISLIIIEINLVAKSYSLLSGDKNYLISN  
LSNTIEKINTLGLDSKNIAINYTDESNNKYFGAISKTSQKSIHYKKDSKNILEFYNDST  
LEFNSKDFIAEDINVMKDDINTITGKYVDNNTDKSIDFSISLVSKNQVKNGLYLNES  
VYSSYLDFVKNSDGHHNTSNEMNLFLDNISFWKLFGFENINFVIDKYFTLVGKTNLGYVE  
FICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTGEDISTSLDFSIEPLYG  
IDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINLDSSSFYKWS  
TEGSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSF  
NSENELDRDHLGFKIIDNKTYYYDEDSKLVKGLININNSLFYFDPIEFNLVTGWQTINGK  
KYYFDINTGAALTSYKIIINGKHFFYNNDGVMQLGVFKGPDGFYFAPANTQNNNIEGQAI  
VYQSKFLTTLNGKKYYFDNNSKAVTGWRIINNEKYYFNPNNAAIAAVGLQVIDNNKYFNP  
TAIISKGWQTVNGSRYYFDTDTAIAFNGYKTIDGKHFFYFSDCVVKIGVFESTSNGFEYFA  
PANTYNNNIEGQAIYVQSKFLTTLNGKKYYFDNNSKAVTGLQTIDSKKYYFNTNTAEAAATG  
WQTDGKKYYFNTNTAEAAATGWQTDGKKYYFNTNTAIASTGYTIINGKHFFYNTDGIMQ  
IGVFKGPNGFYFAPANTDANNIEGQAILYQNEFLTTLNGKKYYFGSDSKAVTGWRIINNK  
KYYFNPNNAAIAAHLCTINNDKYYFSYDGIHQNGYITIERNNFYFDANNESKMVTGVFKG  
PNGFEYFAPANTHNNNIEGQAIYVQNKFLTTLNGKKYYFDNNSKAVTGWQTDGKKYYFNL  
NTAEAAATGWQTDGKKYYFNLNTAEAAATGWQTDGKKYYFNTNTFIASSTGYTSINGKHFF  
FNTDGIMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNKFLTTLNGKKYYFGSDSKAVT  
GLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASSTGYTIISGKHFFYNTDGIM  
QIGVFKGPDGFYFAPANTDANNIEGQAIYQNRFLYLHDNIYYFGNNSKAATGWVTIDG  
NRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAI  
RYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVVYFMPDTAMAAAGGLFEIDGVIYFFGV  
DGVKAPGIYG

**SEQ ID NO: 2 - secuencia de la toxina B**

MSLVNRKQLEKMANVRFRTQDEYVAIILDALEEYHNMSENTVVEKYLKLDINSITDIYI  
DTYKKSGRNKALKKFKEYLVTEVLELKNNNLTPVEKNLHFVWIGGQINDTAINYINQWKD  
VNSDYNVNVFYDSNAFLINTLKKTVVESAINDTLESFRENLDPRFDYNKFFRKRMEIIY  
DKQKNFINYYKAQREENPELIIDDIVKTYLSNEYSKEIDELNTYIEESLNKITQNSGNDV  
RNFEFVKNGESFNLYEQELVERWNLAAASDILRISALKEIGGMYLDVDMPLPGIQPDLFES

IEKPSSVTVDWFEMTKLEAIMKYKEYIPEYTSSEHFDMLDEEVQSSSFESVLASKSDKSEIF  
 SSLGDMEASPLEVKIAFNKSGIINQGLISVKDSYCSNLIVKQIENRYKILNNSLNPaise  
 DNDFTNTNTFIDSIMAEANADNGRFMMELGKYLRVGFFPDVKTTINLSGPEAYAAAYQD  
 LLMFKEGSMNIHLIEADLRNFEISKTNISQSTEQEMASLWSFDDARAKAQFEYKRNIFE  
 GSLGEDDNLDFSQNIIVVDKEYLLEKISSLARSSERYIHYIVQLQGDKISYEAACNLFAK  
 TPYDSVLFQKNIEDSEIAYYYPNCDGEIQEIDKYKIPSIISDRPKIKLTFICHGKDEPNT  
 DIFAGFDVDSLSTEIEAAIDLAKEDISPKEIEINLLGCNMFSSYSINVEETYPGKLLLKVK  
 DKISELMPSISQDSIIIVSANQYEVRIINSEGRRELLDHSGEWINKEESIICKDISKEYTSF  
 NPKENKITVKSKNLPELSTLLQEIARNNSNSSDIELEEKVMLTECEINVISNIDTQIVEER  
 IEZAKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSSHIFISFEDISSETDEGFSIRF  
 INKETGESIFVETEKTFSEYANHITTEEISKIKGTIFDVTNGKLVKKVNLDTHEVNTLN  
 AAFETIQSLIEYNSSKESLSNLSVAMKVQVYAQLFSTGLNTITDAKVVVELVSTALDETID  
 LLPTLSEGLPIIATIIDGVSLGAAIKELSETSDPLLRQEIEAKIGIMAVNLTATTALIT  
 SSLGIASGFSILLVPLAGISAGIPSIIVNNELVLRCKATKVVDYFKHVSLETEGVFTLLD  
 DKIMMPQDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDIDHFFSAPSITYREPHL  
 SIYDVLEVQKEELDLSKDLMLVLPNAPNRVFAWETGWTPLRLSLENDGTLKLLDRIRDNYEG  
 EFWRYFAFIADALITTLKPRYEDTNIRINLDSNTRSFIVPIITTEYIREKLSYSFYGGG  
 GYALSLSQYNMGINIELSESVDWIIDVDNVVRDVTIESDKIKKGDLEIGILSTLSIEEN  
 KIIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIEVDLLSKSYKLLISGELKILMLNS  
 NHIQQKIDYIGFNSQLKNIPYSFVDSGKENGFINSTKEGLFVSELDPVVLISKVYMD  
 DSKPSFGYYSNNLKDVKVITKDNVNILTGYLLKDKIKISLSLTQDEKTIKLSVHLDSE  
 GVAEILKFMNRKGNNTSDSLMSFLESMNIKSIFVNFLQSNIKFILDANFIISGTTISIGQ  
 FEFICDENDNIQPYFIKFNLTETNYTLYVGNRQNMIVEPNYDLDDSGDISSTVINFSQKY  
 LYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYINEKINVNINDLSIRY  
 VWSNDGNDFILMSTSEENKVSQVKIRFVNVEFKDKTLANKLSFNFSQKQDVPVSEIILSFT  
 PSYYEDGLIGYDLGLVSLYNEKFIYINNFGMMVSGLIYINDSLYYFKPPVNNLTIGFVTVG  
 DDKYYFNPINGGAASIGETIIDDKNYFNFQSGVLQTVGFSTEDGFKYFAPANTLDENLEG  
 EALDFTGKLIIDENIYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFN  
 SDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFA  
 HHNEDLGNEEGEEISYSIGILNFNKIIYFDDSFATAVVGWKLDEGSKYYFDEDTAEAYIG  
 LSLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGIIESGVQNIIDNIFYIDNNGIVQIG  
 VFDTSDCYKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESD  
 KYFNPETKKACKGINLIDDIKYFDEKGIMRTGLISFENNYYFNENGEMQFGYINIEDKMFYF  
 GEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYI  
 AATGSVIIDGEEYFDPDTAQLVISE

**SEQ ID NO: 3 - secuencia de Fusión 1**

MGWQITDGKKYFNTNTAIASTGYTIINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNEF  
 LTLNGKKYFSGDSKAVTGWRIINNKKYFNPNNIAAIIHLCTINNDKYYFSYDGLQNGYITIERNNFYFDANN  
 ESKMVTGVFKGPNGFYFAPANTHNNNIEGQAIYQNKFLTLNGKKYFNDNSKAVTGWQITDGKKYFNLNTAE  
 AATGWQITDGKKYFNLNTAEATGWQITDGKKYFNTNTFIASGTYSINGKHFFNTDGMQIGVFKGPNGFY  
 YFAPANTDANNIEGQAILYQNKFLTLNGKKYFSGDSKAVTGLRTIDGKKYFNTNTAVAVTGWQITNGKKYFN  
 TNTSIASGTYTIISGKHFFNTDGMQIGVFKGPDGFYFAPANTDANNIEGQAIYQNRFLYLHDNIYFNGNS  
 KAATGWVTIDGNRYFEPNTAMGANGYKTIDNKNFYFRNLGPQIGVFKGSNGFEYFAPANTDANNIEGQAIYQN  
 RFLHLLGKIYFNGNSKAVTGWQITNGKVYFMPDAMAAAGGLFEIDGVIYFFGVDGVKAPGFVSINDNKHYFD  
 DSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSIGILNFNKIIYFDDSF  
 AVVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGIIESGVQNIIDNIFY  
 IDNNGIVQIGVFDTSDGYKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESDKYFN  
 PETKKACKGINLIDDIKYFDEKGIMRTGLISFENNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPD  
 GFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDTAQLVISE

**SEQ ID NO: 4 - secuencia de Fusión 2**

MGWQITDGKKYFNTNTAIASTGYTIINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNEF  
 LTLNGKKYFSGDSKAVTGWRIINNKKYFNPNNIAAIIHLCTINNDKYYFSYDGLQNGYITIERNNFYFDANN  
 ESKMVTGVFKGPNGFYFAPANTHNNNIEGQAIYQNKFLTLNGKKYFNDNSKAVTGWQITDGKKYFNLNTAE

AATGWQTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNTNTFIASGTYSINGKHFYFNTDGMQIGVFKGPNNGFE  
YFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFN  
TNTSIASGTYTIIISGKHFFYFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAIYQNRFLYLHDNIYYFGNNS  
KAATGWVTIDGNRYFFEPNTAMGANGYKTIIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIYQNR  
RFLHLLGKIYYFGNNSKAVTGWQTINGKVVYFMPDTAMAAAGGLNQIGDYKYYFNSDGMQKGFVSINDNKHFFD  
DSGVMKVGYTEIDGKHFFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNKIIYYFDDSF  
VVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFFYFSDSGIIESGVQNIIDN  
IDNNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVVRVGEDVYYFGETYTIETGWIYDMENESDKYYFN  
PETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPD  
GFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDTAQLVISE

#### SEQ ID NO: 5 - secuencia de Fusión 3

MGWQTIDGKKYYFNTNTAIASTGYTIIINGKHFYFNTDGMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNIAAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNNGFEYFAPANTHNNNIEGQAIYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAE  
AATGWQTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNTNTFIASGTYSINGKHFYFNTDGMQIGVFKGPNNGFE  
YFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFN  
TNTSIASGTYTIIISGKHFFYFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAIYQNRFLYLHDNIYYFGNNS  
KAATGWVTIDGNRYFFEPNTAMGANGYKTIIDNKNFYFRNGLPQIGVFKGSNGFEYFAHHNEDLGNEEGEEISYSG  
ILNFNKIIYYFDDSFVVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFFYFSD  
SGIIESGVQNIIDNNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVVRVGEDVYYFGETYTIET  
GWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNYYFNENGEMQFGYINIEDKMF  
YFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDT  
AQLVISE

#### SEQ ID NO: 6 - secuencia de Fusión 4

MGWQTIDGKKYYFNTNTAIASTGYTIIINGKHFYFNTDGMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNIAAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNNGFEYFAPANTHNNNIEGQAIYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAE  
AATGWQTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNTNTFIASGTYSINGKHFYFNTDGMQIGVFKGPNNGFE  
YFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFN  
TNTSIASGTYTIIISGKHFFYFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAIYQNRFLYLHDNIYYFGNNS  
KAATGWVTIDGNRYFFEPNTAMGANGYKTIIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIYQNR  
RFLHLLGKIYYFGNNSKAVTGWQTINGKVVYFMPDTAMAAAGGETIIDDKNYYFNQSGVLQTVGFSTEDGFKYFA  
PANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGV  
MQKGFVSINDNKHFFDDSGVMKVGYTEIDGKHFFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGI  
LNFNKIIYYFDDSFVVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFFYFSDS  
GIIESGVQNIIDNNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVVRVGEDVYYFGETYTIET  
GWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNYYFNENGEMQFGYINIEDKMF  
YFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDT  
AQLVISE

#### SEQ ID NO: 7 - secuencia de Fusión 5

MGWQTIDGKKYYFNTNTAIASTGYTIIINGKHFYFNTDGMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNIAAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNNGFEYFAPANTHNNNIEGQAIYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAE  
AATGWQTIDGKKYYFNLNTAEAAATGWQTIDGKKYYFNTNTFIASGTYSINGKHFYFNTDGMQIGVFKGPNNGFE  
YFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFN  
TNTSIASGTYTIIISGKHFFYFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAIYQNRFLYLHDNIYYFGNNS  
KAATGWVTIDGNRYFFEPNTAMGANGYKTIIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIYQNR  
RFLHLLGKIYYFGNNSKAVTGWQTINGKVVYFMPDTAMAAAGGLFEIDGVIYFFGVDGKAPGIYGGGFVSINDN

KHYFDDSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFD  
DSFTAVVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFFSDSGIIESGVQNI  
DNYFYIDDNGIVQIGVFDTSBGYKYFAPANTVNDNIYQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENES  
KYFNPETKKACKGINLIDDIKYFDEKIMRTGLISFENNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGV  
FNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIIDGEEYFDPDTAQLVISE

**SEQ ID NO: 8 - secuencia del fragmento de toxina A individual**

MASTGYTSINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNKFLTNGKKYYFGSDSKAVT  
GLRTIDGKKYYFNTTAVAVTGWQTINGKKYYFNTNTSIASGTIISGKHFFNTDGMQIGVFKGPDGFYFA  
PANTDANNIEGQAIYQNRFLYLHDNIYFGNNSKAATGWVTIDGNRYFEPNTAMGANGYKTI DNKNFYFRNGL  
PQIGVFKGSNGFEYFAPANTDANNIEGQAIYQNRFLHLLGKIYFNGNNSKAVTGWQTINGKVVYFMPDTAMAAA  
GGLFEIDGVIYFFGVDGVKAP

**SEQ ID NO: 9 - secuencia del fragmento de toxina B individual**

MILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSKQDVPVSEIILSFTPSYYEDGLIGYDLGLVSLYNEKF  
YINNFMMVSGLIYINDSLYFKPPVNNLITGFVTVGDDKYFNPINGGAASIGETIIDDKNYYFNQSGVLQTV  
FSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIG  
DYKYFNSDGMQKGFVSINDNKHFFDDSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHHNEDLGN  
EEGEEISYSGILNFNNKIYYFDDSTAVVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVT  
INDKVFFSDSGIIESGVQNIIDNIFYIDDNGIVQIGVFDTSBGYKYFAPANTVNDNIYQAVEYSGLVRVGEDV  
YYFGETYTIETGWIYDMENESDKYFNPETKKACKGINLIDDIKYFDEKIMRTGLISFENNYYFNENGEMQF  
GYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVII  
DGEEYFDPDTA

**SEQ ID NO: 10 - secuencia de nucleótidos de F54Gli**

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTTAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTTATCAACGGCA  
AACACTTTTATTTTAAACACCGACGGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACGGCTTTGAATACTTTGCACCGGCAAAATACCGATGC  
CAATAAATATTGAAGGCCAGGCCATTCTGTATCAGAAATGAATTTCTGACCCCTGAACGGCAAAAAATATACTACTTTGGCAGCGATAGCAAAAGCAGTT  
ACCGGTTGGCGCATCATCAACAATAAGAAATATTACTTCAACCCGAATAATGCAATTGCAGCAATTCATCTGTGCACCATTAACAACGACAAAT  
ATTATTTTCAGCTATGACCGTATTCTGCAGAAATGGCTACATTCACATCGAACGCAACAATTTTATTTCGATGCCAACACGAAAGCAAAATGGT  
GACCGGTGTTTCAAAGGCCCTAAATGGTTTGGAGTATTTCGCTCCGGCAACACCCATAATAACAACATTGAAGGTGAGCGATCGTTTATCAG  
AACAAATTCCTGACCGTGAATGGTAAGAAATATACTATTTCGATAATGACAGCAAAAGCCGTGACCGGCTGGCAGACAAATTGACGGGAAGAAATATT  
ACTTTAATCTGAATACCGCAGAACGACGCAACCGGTTGGCAACGATCGACGGTAAAAAGTACTACTTCAACCTGAACACAGCCGAAGCAGCCAC  
AGGATGGCAGACTATTGATGGAAAAAATACTATTTCACACCAACACCTTTATTGCATCTACCGGTTATACAGCATTAAACGGTAAACATTTC  
TACTTCAACACCGATGGTATCATGCAGATCGCGGTTTTCAAAGGTCCAAATGGTTTCGAATACTTTGCCCTTCGCAATACAGATGCAAAATAACA  
TCGAGGCTCAGCAATCTCTACCAAAACAAATTTCTGACCGTGAATGGGAAAAAATATTACTTTGGTAGCGATTCTAAAGCCGTTACCGGCTCT  
CGCTACCATTTGATGGTAAAAAATACTACTTTAATACGAATACAGCCGTTGCGGTTACAGGCTGGCAGACCAATTAACGGGAAAAAATACTATT  
AACCAAAATACAGCATTGCTCTAACCGGTTATACCATTTTCGGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAAATCGGAGTCT  
TTAAAGGACCTGATGGGTTGCAATATTTCGCGCTGCGAACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAAATCGCTTTCT  
GTATCTGCACGACAACATCTATTATTTTGGCAACAATTCAAAGCAGCCACCGGCTGGGTTACAATTGATGGCAACCGCTACTATTTCGAACCG  
AATACCGCAATGGGTGCAAAATGGCTACAAACCATCGATAATAAAATTTCTATTTCGCAACGGTCTGCCGAGATCGGGGTATTTAAAGGTA  
GCAACCGGCTTCAATACTTCGCTCCAGCGAATACGGACGCGAACAATATTGAGGGTCAAGCGATTTCGTTATCAAAACCGTTTTCGCTCTGCT  
GGGCAAAATCTACTACTTTGGCAATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGGATACCGCC  
ATGGCAGCAGCGGTTGCTGTTTGAATGATGGCGTGTATCTATTTTTCGTTGGTGGATGGTGTAAAGCACCGGGAATATACGGTGGTACCG  
GCTTTGTGACCGTGGGTGATGATAAATACTATTTCATCCGATTAAACGGTGGTGCAGCGAGCATTGGCGAAACCATCATCGATGACAAAACTA  
TTATTTCAACAGAGCGGTGTGCTGCAGACCGGTTGTTTAGCACCGAAGATGGCTTTAAATATTTCGCGCCAGCGAACCCCTGGATGAAAAAC  
CTGGAAGGCGAAGCGATTGATTTTACCGGCAAACTGATCATCGATGAAACATCTATTACTTCGATGATAACTATCGTGGTGGGTGGAATGGA  
AAGAACTGGATGGCGAAATGCATTTTTCCTCCGGAACCGGTAAGCGGTTTAAAGGCTGAACAGATCGCGGATTACAAATFACTACTTCAA  
CAGCGATGGCGTATGCAGAAAGGCTTTGTGAGCATCAACGATAACAAACACTATTTCGATGATAGCGGTGTGATGAAAGTGGGCTATACCGAA  
ATTGATGGCAAACTTTCTACTTCGCGGAAACCGGCAATGCGAGATTGGCGGTTCATACCGGAAGATGGTTTCAATATCTTCGCGACCATTA  
ACGAAGATCTGGGTAAACGAAGGCGAAGAAATTAGCTATAGCGGCATCTGAACCTTCAACAAACAAATCTACTACTTTGATGATAGCTTTAC  
CGCGGTGTTGGCTGGAAGATCTGGAAGATGCGAGCAAAATATTATTTCGATGAAGATACCGCGAAGCGTATATTGGCTGAGCGCTGATTAA  
GATGCCAGTACTATTTTAAACGATGATGGCATTATGCAGGTGGGTTTCGTGACCATTAATGATAAAGTGTCTATTTCAGCGATAGCGGCATT  
TTGAAGCGCGGTGCAAGACATTGATGATAACTACTTCTACATCGATGATAACGGCATTGTGCAGATCGCGGTTTTTGATACAGCGATGGCTA  
CAAAATATTTCGACCGGCCAATACCGTGAACGATAAATTTATGGCCAGCGGTGGAATATAGCGGCTGGTGGCGAAGATGTGTAT  
TATTTCCGGCGAAACCTATACCATCGAAACCGGCTGGATTATGATATGGAAGCGAAGCGATAAATATTACTTTAATCCGGAACGAAAAAG  
CGTGCAGAGGCTTAACCTGATGATGATGATCAAAATCTATTTGATGAAAGGCAATATGCGTACCGGCTGATTAGCTTCGAAAAACAA  
CTATTACTTCAACGAAACCGTGAAATGCAGTTTCGGCTACATCAACATCGAAGATAAATGTTCTACTTCGCGAAGATGGTGTATGAGATT

GGTGTTTTAAACACCCCGGATGGCTTCAAATACTTTGCCCATCAGAATACCCCTGGATGAAAATTTCCAAGGTGAAAGCATTAACTATACCGGCT  
GGCTGGATCTGGATGAAAAACGCTACTACTTCACCGATGAATACATTGCGGCGACCGGACGGCTGATTATGATGGCGAAGAATACTACTTCGA  
TCCGGATACCGCGCAGCTGGTGATTAGCGAACATCATCATCACCAT

**SEQ ID NO: 11- secuencia de aminoácidos de F54Gli**

MATGWQITIDGKKYYFNTNTAIASSTGYTIINGKHFFNTDGMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEFLTLN  
GKKYYFGSDSKAVTGWRIINNKKYYFNPNAIAAHLCTINNDKYYFSYDGLQNGYITIERNNFYDANNESKMVTGVF  
KGFNGFEYFAPANTHNNNIEGQAIYVQNKFLTLNGKKYYFDNDSKAVTGWQITIDGKKYYFNLNTAEAAATGWQITIDGKKYY  
FNLNTAEAAATGWQITIDGKKYYFNTNTFIASSTGYTSINGKHFFNTDGMQIGVFKGPNGFEYFAPANTDANNIEGQAILY  
QNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQITINGKKYYFNTNTSIASSTGYTIISGKHFFNTDG  
IMQIGVFKGPDGFEYFAPANTDANNIEGQAIYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYFEPNTAMGANGYKT  
IDNKNFYFRNLGPQIGVFKGSNGFEYFAPANTDANNIEGQAIYQNRFLHLLGKIYYFGNNSKAVTGWQITINGKVVYFMP  
DTAMAAAGGLFEIDGVIYFFGVDGVKAPGIYGGTGFTVVGDDKYYFNPNINGGAASIGETIIDDKNYYFNQSGVLQTCGVFS  
TEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEHMYFSPETGKAFKGLNQIGDYKYYFN  
SDGVMQKGFVSINDNKHFFDSDGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHNNEDLNEEGEEISYSGIL  
NFNKNIYYFDDSFYAVVGWDLLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDGIMQVGFVTINDKVFYFSDSGIIESG  
VQNIDNYYFYIDDNGIVQIGVFDTSBGYKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESD  
KYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPD  
GFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIDGEEYFDPDTAQLVISEHHHHH

**SEQ ID NO: 12 - secuencia de nucleótidos de F54Nueva**

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTTAATACCAACACCGCAATTGCAAGCACCGGCTATACCATATCAACGGCA  
AACACTTTTATTTAAACACCGACGGCATTATGCGAGATTGGTGTGTTTAAAGGTCCGAACGGCTTTGAATACCTTTGCACCGGCAAAATACCGATGC  
CAATAATATTGAAGGCCAGGCCATTCTGTATCAGAAATGAATTTCTGACCCCTGAACGGCAAAAAATACCTACTTTGGCAGCGATAGCAAGCAGTT  
ACCGGTTGGCGCATCATCAACAATAAGAAATATTACTTCAACCCGAATAATGCAATTGCGAGCAATTCACTCTGTGCACCATTAACAACGACAAAT  
ATTATTTGAGCTATGACGGTATTCTGCGAGAAATGGCTACATTACCATCGAAGCAACAACTTTTATTTCGATGCCAACCAAGCAAAATGGT  
GACCGGTGTTTTCAAAGGCCCTAATGGTTTTGAGTATTTCGCTCCGGCAACACCCATAATAACAACATTGAAGGTGAGCGCATCGTTTATCAG  
AACAAATCTCTGACCGTGAATGGTAAGAAATACTATTTCGATAAATGACAGCAAGCCGTGACCGGCTGGCAGCAATGACGGGAAGAAATATT  
ACTTTAATCTGAATACCGCAGAAGCAGCAACCGGTTGGCAACCGATCGACGGTAAAAAGTACTACTTCAACCTGAACACAGCCGAAGCAGCCAC  
AGGATGGCAGACTATTGATGCAAAAAATACTATTTCACACCAACACCTTTATTGCGATCTACCGGTTATACAGCATTAAACGGTAAACATTTCT  
TACTTCAACACCGATGGTATCATGCGAGATCGGCGTTTTCAAAGGTCCAATGGTTTCGAATACTTTGCCCTGCCAATACAGATGCAAAATAACA  
TCGAGGGTCAGGCAATCTCTGACCAAAACAAATTTCTGACCCCTGAATGGGAAAAAATATTACTTTGGTAGCGATTCTAAAGCCGTTACCGGTCT  
GCGTACCATTGATGGTAAAAAATACTACTTTAATACGAATACAGCCGTTGCGGTTACAGGCTGGCAGACCATTAACGGGAAAAAATACTATTTT  
AACACAAATACCGCATTCGCTCAACGGGTTATACCATTTTCGGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAAATCGGAGTCT  
TTAAAGGACCTGATGGTTTGAATATTTCGCGCTGCGAACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAAATCGCTTTCT  
GTATCTGACAGCAACATCTATTATTTTGGCAACAATTCAAAAGCAGCCACCGGCTGGGTTACAATTGATGGCAACCGCTACTATTTCGAACCG  
AATACCGCAATGGGTGCAAAATGGCTACAAAACCATCGATAATAAAAAATTTCTATTTCGCAACCGGCTGCCCGCAGATCGGGTATTAAAGGTA  
GCAACGGCTTCGAATACTTCGCTCCAGCGAATACGGACGGCAACATATTGAGGGTCAAGCGATTCTGTTATCAAAACCGTTTTCTGCACTGCT  
GGGCAAAATCTACTACTTTGGCAATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGACTATTTATGCGCGATACCGGC  
ATGGCAGCAGCCGTTGCTGTTTGAATTTGATGGCGTGATCTATTTTTTTGGTGGTGGTGGTTAAAGCAGTTACCGGCTTTGTGACCGTGG  
GTGATGATAAAATCTATTTCATCCGATTAAACGGTGGTGCAGCGAGCATTTGGCGAAACCATCATCGATGACAAAAATCTATTATTCAACCGAG  
CGGTGTGCTGCGACCGGTTGTTAGCACCGAAGATGGCTTTAAATATTTCGCGCAGCGAACACCCCTGGATGAAACCTGGAAGCGAAGCG  
ATTGATTTTACCGGCAAACTGATCATCGATGAAACATCTATTACTTCGATGATAACTATCGTGGTGGTGGTGAATGGAAGAACTGGATGGCG  
AAATGCAATATTTTTTTCGGAACCGGTAAGCGTTTAAAGGCCCTGAACCGATCGGCGATTACAAATACTACTTCAACAGCGATGGCGTGAT  
GCAGAAGCGCTTTCTGAGCATCAACGATAACAAACACTATTTCGATGATAGCGGTGCTGATGAAAGTGGCTATACCGAAATGATGGCAACAT  
TTCTACTTCGCGGAAACCGCGAAATGAGATTGGCGTGTTCATACCGAAGATGGTTTCAAATCTTCGCGCACCATAACGAAGATCTGGGTA  
ACGAAGAAGCGCAAGAAATAGCTATAGCGGCATCTGAACCTTCAACAACAAAACTACTACTTTGATGATAGCTTTACCGCGGTGGTGGGCTG  
GAAAGATCTGGAAGATGGCAGCAAAATATTATTTCGATGAAGATACCGCGAAGCGTATATTGGCCTGAGCCTGATTAAACGATGGCCAGTACTAT  
TTTAACGATGATGGCATTATGCGAGTGGGTTTCGTGACCATTAATGATAAAGTGTCTATTTCAGCGATAGCGGCATTATTGAAGCGCGGTGC  
AGAACATTGATGATAACTACTTCTACATCGATGATAACGGCATTTGCGCAGATCGGCGTTTTTGATACCGCGATGGCTACAAATATTTCGACCC  
GGCAATACCGTGAACGATAACATTATGGCCAGCGGTTGAATATAGCGGTCTGGTGGCTGTGGCGAAGATGTGATTATTTCGGCGAAACC

TATACCATCGAAACCGGCTGGATTTATGATATGGAACGAAAGCGATAAATATTACTTTAATCCGGAACGAAAAAGCGTGCAAAGGCATTA  
ACCTGATCGATGATATCAAAATACATATTTGATGAAAAAGGCATTTATGCGTACCGGTCTGATTAGCTTCGAAAAACAACACTATTACTTCAACGA  
AAACGGTGAAATGCAGTTCGGCTACATCAACATCGAAGATAAAATGTTCTACTTCGGCGAAGATGGTGTATGCAGATTGGTGTTTTTAACACC  
CCGGATGCGCTTCAAATACCTTGGCCATCAGAATACCTGGATGAAAATTTGGAAGCTGAAAGCATTAACATACCGGCTGGCTGGATCTGGATG  
AAAAACGCTACTACTTCACCGATGAATACATTGCGGCGACCGCGAGCTGATTATTGATGGCGAAGAATACTACTTCGATCCGGATACCGCGCA  
GCTGGTGATTAGCGAACATCATCATCATCACCAT

**SEQ ID NO: 13 - secuencia de aminoácidos de F54Nueva**

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFFNTDGMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQN  
EFLTLNGKKYYFGSDSKAVTWRIINNKKYYFNPNNIAIAHLCTINNDKYYFSYDGLQNGYITIERNNFYFDANNESK  
MVTGVFKGPNNGFEYFAPANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDNDSKAVTWQTIDGKKYYFNLNTAEAAATGWQTI  
DGKKYYFNLNTAEAAATGWQTIIDGKKYYFNTNTFIASTGYTSINGKHFFNTDGMQIGVFKGPNNGFEYFAPANTDANNIE  
GQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTWQTINGKKYYFNTNTSIASTGYTIISGKHF  
YFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAIYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYFEPNTAMG  
ANGYKTIDNKNFYFRNGLPQIGVFKGPNNGFEYFAPANTDANNIEGQAIYQNRFLHLLGKIYYFGNNSKAVTWQTINGK  
VYVFMEDTAMAAAGGLFEIDGVIYFFGVGVKAVTGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTVGF  
STEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYF  
NSDGVMQKGFVSINDNKHFFDDSGVMKVGYTEIDCKHFFYAENGEMQIGVFNTEDGFKYFAHNEDLGNEEGEEISYSGI  
LNFNNKIYYFDDSFYAVVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGMQVGVFVTINDKVYFSDSGIIES  
GVQNIIDNYYIIDDNGIVQIGVFEDTSDGYKYFAPANTVNDNIYQGAVEYSGLVRVGEDVYFGETYTIETGWIYDMENES  
DKYYFNPETKKACKGINLIDDIKYYFDEKGMRTGLISFENNYYFNENGEMQFGYINIEDKMFYFCEGDVMQIGVFNTF  
DGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDTAQLVISEHHHHH

**SEQ ID NO: 14 - secuencia de nucleótidos de F5 ToxB**

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTTAAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTTATCAACGGCA  
AACACTTTTATTTTAAACACCGACCGCATTATGAGATTTGGTGTGTTTAAAGGTCCGAACCGGCTTTGAATACTTTGCACCGGCAAAATACCGATGC  
CAATAATATTGAAGGCCAGGCCATTCTGTATCAGAATGAATTTCTGACCCCTGAACGGCAAAAAATACTACTTTGGCAGCGATAGCAAGCAGATT  
ACCGGTTGGCGCATCATCAACAATAAGAAATATTACTTCAACCCGAATAATGCAATTGCGAGCAATTCATCTGTGCACCATTAACAACGACAAAT  
ATTATTTGACGTATGACGGTATTCTGCAGAATGGCTACATTACCATCGAAGCGCAACAACTTTTATTTCGATGCGCAACAAAGCAAAATGGT  
GACCGGTGTTTTCAAAGGCCCTAATGGTTTTGAGTATTTTCGCTCCGGCAACACCCATAATAACAACATTGAAGGTGACGGCGATCGTTTTATCAG  
AACAAATTCCTGACCGTGAATGGTAAGAAATACTATTTCGATAATGACAGCAAGCCGTGACCGGCTGGCAGACAATTGACGGGAAGAAATATT  
ACTTTAATCTGAATACCGCAGAAGCAGCAACCGGTTGGCAACGATCGACGGTAAAAAGTACTACTTCAACCTGAACACAGCCGAAGCAGCCAC  
AGGATGGCAGACTATTGATGGAAAAAATACTATTTCAACACCAACACCTTTTATTGCATCTACCGGTTATACCGAGCATTAAACGGTAAACATTTC  
TACTTCAACACCGATGGTATCATGCAGATCGGCGTTTTCAAAGGTCCAATATGGTTTTCGAATACTTTGCCCTGCCAATACAGATGCAAAATAACA  
TCGAGGGTCAGGCAATCCTGTACCAAAACAAATTTCTGACCCCTGAATGGGAAAAAATATTACTTTGGTAGCGATTCTAAAGCCGTTACCGGTCT  
CGGTACCATTTGATGTTAAAAATACTACTTTAATACGAATACAGCCGTTGCGGTTACAGGCTGGCAGACCATTAAACGGGAAAAAATACTATTTT  
AACACAAATACCGCATTTGCTCAACGGGTTATACCATTTTTCGGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAAATCGGAGTCT  
TTAAAGGACCTGATGGGTTTCAATATTTTTCGGCTGCGAACACTGATGCGAACAAATCGAAGGACAGGCAATCCGCTATCAGAAATCGCTTTCT  
GTATCTGCACGACAACATCTATTATTTTGCAACAATTCAAAAGCAGCCACCGGCTGGGTACAATTTGATGGCAACCGCTACTATTTCGAACCG  
AATACCGCAATGGGTGCAATGGCTACAAAACCATCGATAATAAAAAATTTCTATTTTCGCAACCGTCTGCCGAGATCGGGGTATTTAAAGGTA  
GCAACGGCTTCGAAATCTCGCTCCAGCGAATACGGACGCGAACAACTATTGAGGGTCAAGCGATTCTGTTATCAAAACCGTTTTCTGCATCTGCT  
GGGCAAAATCTACTACTTTTGGCAATAACAGTAAAGCAGTTACTCGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGATACCGCC  
ATGGCAGCAGCGGTTGGTCTGTTTGAATTTGATGGCGTGATCTATTTTTCGGTGTGGATGGTGTAAAGCAGTGAGCGGTCTGATTATATTA  
ACGATAGCTGTATTACTTTAAACACCGGTTGAATAACCTGATTACCGGCTTTGTGACCGTGGGTGATGATAAATACTATTTCATCCGATTAA  
CGGTGGTGCAGAGCATTTGGCGAAACCATCATCGATGACAAAAAATACTATTTCACACAGAGCGGTGTGCTGCAGACCGGTTGTTTAGCACC  
GAAGATGGCTTTAAATATTTTGGCCAGCGAACCCCTGGATGAAAACCTGGAAGGCGAAGCGATTGATTTTACCGGCAACATGATCATCGATG  
AAACATCTATTACTTCGATGATAACTATCTGGTGGCGTGGAATGGAAGAACTGGATGGCGAAATGCATTATTTTCTCGGAAACCGGTAA  
AGCGTTTAAAGGCTGAACAGATCGGCGATTACAAATACTACTTCAACAGCGATGGCGTGTGCGAAGAGGCTTTGTGAGCATCAACGATAAC  
AAACACTATTTTCGATGATAGCGGTGTGATGAAAGTGGGCTATACCGAAATTTGATGGCAACATTTCTACTTCGCGGAAACCGCGAAATGCAGA  
TTGGCGTGTTCATACCGAAGATGGTTTCAAATACTTCGCGCACCATACGAAGATCTGGGTAACGAAGAGCGAAGAAATAGCTATAGCGG  
CATCTTGAACCTTCAACAACAAATCTACTACTTTGATGATAGCTTTACCGCGGTGGTGGGCTGGAAAGATCTGGAAGATGGCAGCAAAATATTAT

TTGATGAAGATACCGCGAAGCGTATATTGGCCTGAGCCTGATTAACGATGGCCAGTACTATTTTAACGATGATGGCATTATGCAGGTGGGTT  
TCGTGACCATTAAATGATAAAGTGTCTATTTCAGCGATAGCGGCATTATTGAAAGCGGCGTGAGAACATTGATGATAACTACTTCTACATCGA  
TGATAACGGCATTGTGCAGATCGGCGTTTTTGATACCGAGCGATGGCTACAAATATTTTCGACCGGCCAATACCGTGAACGATAAATTTATGGC  
CAGCGGTGGAAATATAGCGGTCTGGTGGTGGGCGAAGATGTGTATTATTTCGGCGAAACCTATACCATCGAAACCGGCTGGATTTATGATA  
TGGAAAACGAAAGCGATAAATATTACTTTAATCGGAAACGAAAAAGCGTGCAAGGCATTAACTGATCGATGATATCAAAATACTATTTTGA  
TGAAAAGGCATTATGCGTACCGGTCTGATTAGCTTCGAAAACAACAATATTACTTCAACGAAACCGGTGAAATGCAGTTCGGCTACATCAAC  
ATCGAAGATAAAATGTTCTACTTTCGGCGAAGATGGTGTATGCGAGATTGGTGTTTTAAACACCCCGGATGGCTTCAAATACTTTGCCATCAGA  
ATACCGTGGATGAAAATTTCGAAGGTGAAAGCATTAACTATACCGGTGGCTGGATGATGAAAAACGCTACTTTCACCGATGAATACAT  
TGGCGCGACCGGCAGCGTGATTATTGATGGCGAAGAATACTACTTCGATCCGGATACCGCGCAGCTGGTATTAGCGAACATCATCATCATC  
CAT

**SEQ ID NO: 15 - secuencia de aminoácidos de F5 ToxB**

MATGWQITIDGKKYYFNTNTAIASTGYTIINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQN  
EFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNIAAIAHLCTINNDKYYFSYDGLQNGYITIERNNFYFDANNESEK  
MVTGVFKGPNGFYFAPANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDNDSKAVTGWQITIDGKKYYFNLNTAEAAATGWQTI  
DGKKYYFNLNTAEAAATGWQITIDGKKYYFNTNTFIASTGYTSINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIE  
GQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIIISGKHF  
YFNTDGMQIGVFKGPDGFYFAPANTDANNIEGQAIYQNRFLYLHDNIYYPGNNSKAATGWVTIDGNRYFEPNTAMG  
ANGYKTIIDNKNYFRNGLPQIGVFKGSGNFEYFAPANTDANNIEGQAIYQNRFLHLGLKIYYFGNNSKAVTGWQTINGK  
VYYFMPDTAMAAAGGLFEIDGVIYFFGVGVKAVSGLIYINDSLYYFKPPVNNLITGVFTVGDKKYFNPINGGAASIGE  
TIIDDKNYFQNSGVLTQGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYFPDNYRGAVEWKELDGEMHYF  
SPETGKAFKGLNQIGDYKYFNSDGVMMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFFPAENGEMQIGVFNTEDGPKY  
FAHHNEDLGNEEGEEISYSGILNFNKKIYFDDSFYAVVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQ  
VGFTVINDKVYFSDSGIIESGVQNIIDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDV  
YVFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYFDEKGIMRTGLISFENNYYFNPENGEMQFGYINI  
EDKMFYFGEDGVMQIGVFNTPDGPKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPD  
TAQLVISEHHHHHH

**SEQ ID NO: 16 - secuencia de nucleótidos de F52Nueva**

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTTAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTATCAACGGCA  
AACACTTTTATTTTAACACCGACGGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACGGCTTTGAATACTTTGCACCGGCAAAATACCGATGC  
CAATAATATTCAAGGCCAGGCCATTCTGTATCAGAATGAATTTCTGACCGTGAACGGCAAAAAATACTACTTTGGCAGCGATAGCAAGCAGTT  
ACCGGTTGGCGCATCATCAACAATAAGAAATATTACTTCAACCCGAATAATGCAATTGCAGCAATTTCATCTGTGCACCATTAACAACGACAAAT  
ATTATTTGAGCTATGACGGTATTCTGCAGAATGGCTACATTACCATCGAACGCAACAATTTTATTTTCGATGCCAACACGAAAGCAAAATGGT  
GACCGGTGTTTTCAAGGCCCTAATGGTTTTGAGTATTTCGCTCCGGCAACACCCATAATAACAACATTGAAGGTCAGGCGATCGTTTATCAG  
AACAAATTCCTGACCTGAATGGTAAGAAATACTATTTCGATAATGACAGCAAGCCGTGACCGGCTGGCAGACAATTGACGGGAAGAAATATT  
ACTTTAATCTGAATACCGCAGAACGCAACCGGTTGGCAACGATCGACGGTAAAGAGTACTACTTCAACCTGAACACAGCCGAAGCAGCCAC  
AGGATGGCAGACTATTGATGGAAAAATACTATTTCACACCAACACCTTTATTGCATCTACCGGTTATACCAGCATTAAACGGTAAACATTTCT  
TACTTCAACACCGATGGTATCATGCAGATCGGCGTTTTCAAGGTCCTCAATGGTTTCGAATACTTTGCCCTGCCAATACAGATGCAAAATAACA  
TCGAGGTCAGGCAATCCTGTACCAAAACAAATTTCTGACCGTGAATGGGAAAAAATATTACTTTGGTAGCGATTCTAAAGCCGTTACCGGTCT  
GCGTACCATTGATGGTAAAAATACTACTTTAATACGAATACAGCGGTTGCGGTTACAGGCTGGCAGACCATTAACGGGAAAAAATACTATT  
AACACAAATACCGCATTGCTCAACGGGTTATACCATTATTTCGGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAAATCGGAGTCT  
TTAAAGGACCTGATGGTTTCGAATATTTTGGCGCTGCAACACTGATGCGAACAAATATCGAAGGACAGGCAATCCGCTATCAGAATCGCTTTCT  
GTATCTGCACGACAACATCTATTATTTTGGCAACAATTCAAAGCAGCCACCGGCTGGGTACAATTGATGGCAACCGCTACTATTTCGAACCG  
AATACCGCAATGGGTGCAAAATGGCTACAAAACCATCGATAATAAAATTTCTATTTTTCGCAACGGCTCGCCGAGATCGGGGTATTAAAGGTA  
GCAACGGCTTCGAATACTTCGCTCCAGCGAATACGGACGCGAACAATATTGAGGGTCAAGCGATTCTGTTATCAAAACCGTTTTCTGCATCTGCT  
GGGCAAAATCTACTACTTTGGCAATAACAGTAAAGCAGTTACTTGGATGGCAGACAATCAATGGTAAAGTACTATTTTATGCGCGATACCGCC  
ATGGCAGCAGCCGGTGGCTGTTTGAATTTGATGGCGTGATCTATTTTTTGGTGTGGATGGTGTAAAGCAGTGAAGGCCCTGAACAGATCG  
GCGATTACAAATACTACTTCAACAGCGATGGCGTGATGCAGAAAGGCTTTGTGAGCATCAACGATAACAACACTATTTCGATGATAGCGGTGT  
GATGAAAGTGGGCTATACCGAAATGATGGCAACATTTCTACTTCGCGGAAAAACGGCGAAATGCAGATTGGCGTGTCAATACCGAAGATGGT

TTCAAATACTTCGCGCACCATAACGAAGATCTGGGTAACGAAGAAGCGGAAGAAATTAGCTATAGCGGCATCCTGAACTTCAACAACAAAAATCT  
 ACTACTTTTGATGATAGCTTTACCGCGGTGGTGGGCTGGAAAGATCTGGAAGATGGCAGCAAAATATTATTTCGATGAAGATACCGCGGAAGCGTA  
 TATTGGCCTGAGCCTGATTAACGATGGCCAGTACTATTTAACGATGATGGCATTATGCAGGTGGGTTTCGTGACCATTAAATGATAAAGTGTTTC  
 TATTTTCAGCGATAGCGGCATTATTGAAAGCGGCGTGCAGAACATTGATGATACTACTTCTACATCGATGATAACGGCATTGTGCAGATCGGCG  
 TTTTGTATACCAGCGATGGCTACAAATATTTGCAACCGGCCAATACCGTGAACGATAACATTTATGGCCAGGCGGTGGAATATAGCGGTCTGGT  
 GCGTGTGGGCGAAGATGTGTATTATTTCGGCGAAACCTATACCATCGAAACCGGCTGGATTTATGATATGGAAAACGAAAGCGATAAAATATTAC  
 TTTAATCCGAAAACGAAAAAGCGTGCAAAAGGCATTAACTGATCGATGATATCAAAATCTATTTTGATGAAAAAGGCATTATGCGTACCGGTC  
 TGATTAGCTTCGAAAACAACAATTTACTTCAACGAAAACGGTGAAATGCAGTTTCGGCTACATCAACATCGAAGATAAAATGTTCTACTTCGG  
 CGAAGATGGTGTATGCAGATTGGTGTTTTAAACCCCGGATGGCTTCAAAATCTTTGCCCATCAGAATACCGTGGATGAAAAATTCGAAGGT  
 GAAAGCATTAATATACCGGCTGGCTGGATCTGGATGAAAAACGCTACTACTTCACCGATGAATACATTGCGGCGACCGGCGAGCGTGATTATTG  
 ATGGCGAAGAATACTACTTCGATCCGGATACCGCGCAGCTGGTGATTAGCGAACATCATCATCATCACCAT

**SEQ ID NO: 17 - secuencia de aminoácidos de F52Nueva**

MATGWQTIDGKKYYFNTNTAIASGYTTIINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNEFLTIN  
 GKKYYFGSDSKAVTGWRIINNKKYYFNPNNIAIAIHLCTINNDKYYFSYDGLQNGYITIERNNFYFDANNESKMVTGVF  
 KGPNGFEYFAPANTHNNNIEGQAIIVYQNKFLTINLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAEAAATGWQTIDGKKYY  
 FNLNTAEAAATGWQTIDGKKYYFNTNTFIASGYTSINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILY  
 QNKFLTINLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASGYTTISGKHFFNTDGM  
 IMQIGVFKGPDGFYFAPANTDANNIEGQAIIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYFEPNTAMGANGYKT  
 IDNKNFYFRNLPLQIGVFKGSNGFEYFAPANTDANNIEGQAIIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVVYFMP  
 DTAMAAAGGLFEIDGVIIYFFGVDGVKAVKGLNQIGDYKYYFNSDGMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFF  
 FAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGLNFNNKIYYFDDSFYAVVGWKLDEGSKYYFDEDTAAY  
 IGLSLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGIIESGVQNIIDNIFYIDDNGIVQIGVFDTSDGYKYFAPANTVN  
 DNIYQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESDKYYFNPETKACKGINLIDDIKYYFDEKGIMRTGLISF  
 ENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDE  
 YIAATGSVIDGEEYFDPDTAQLVISEHHHHH

**FIGURA 2**

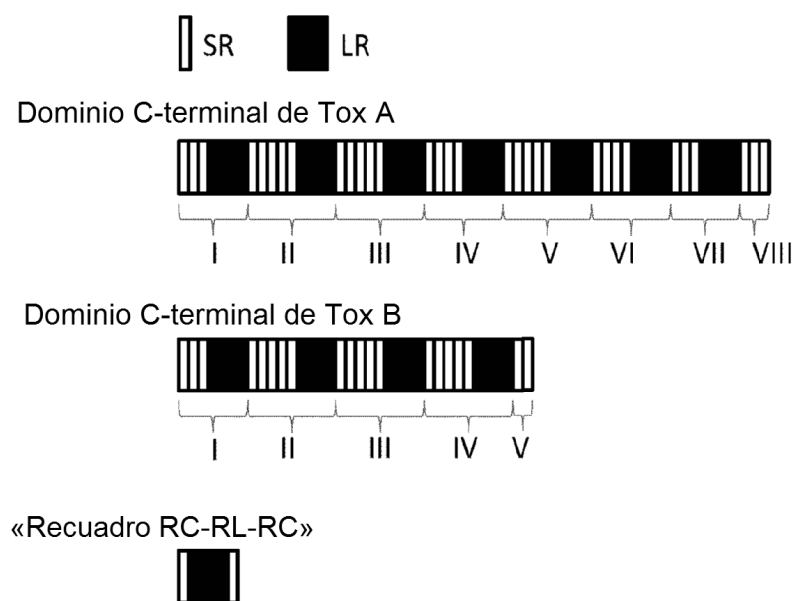


FIGURA 3

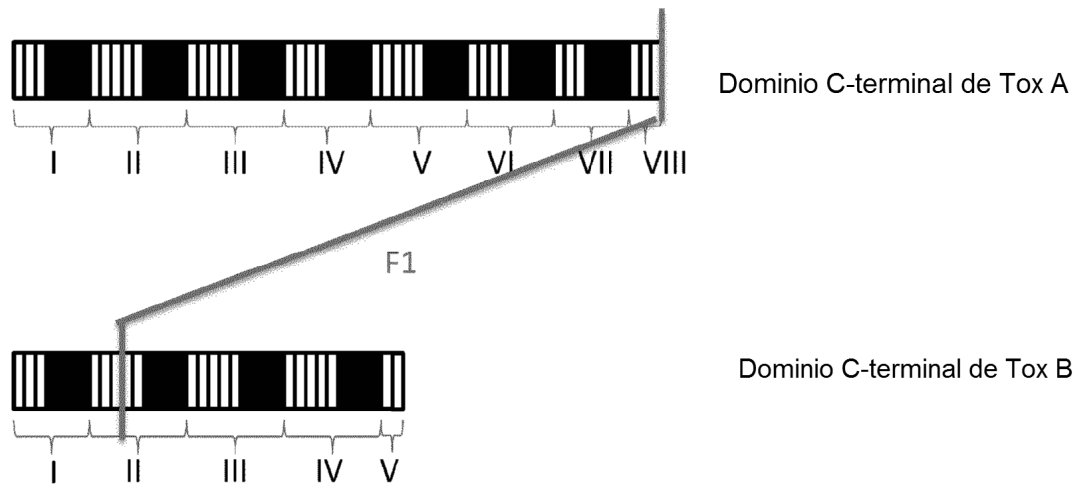
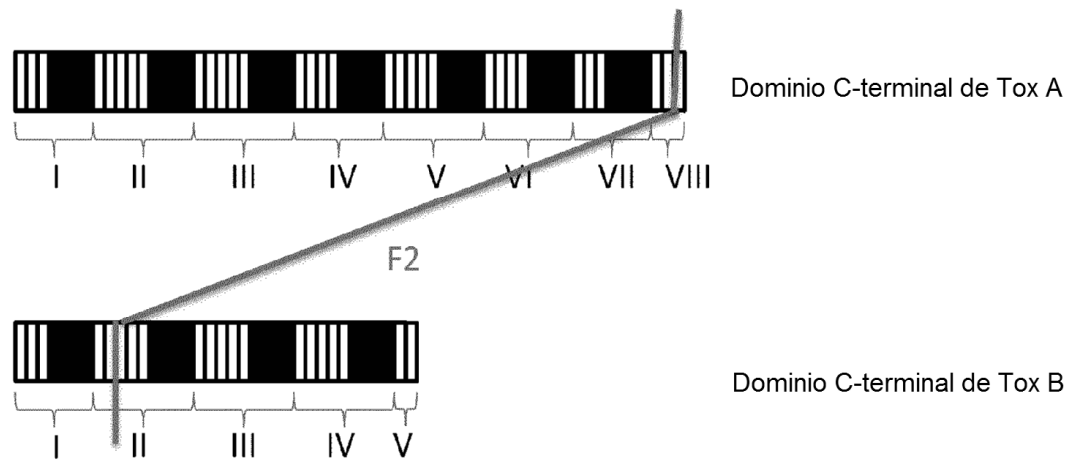


FIGURA 4



**FIGURA 5**

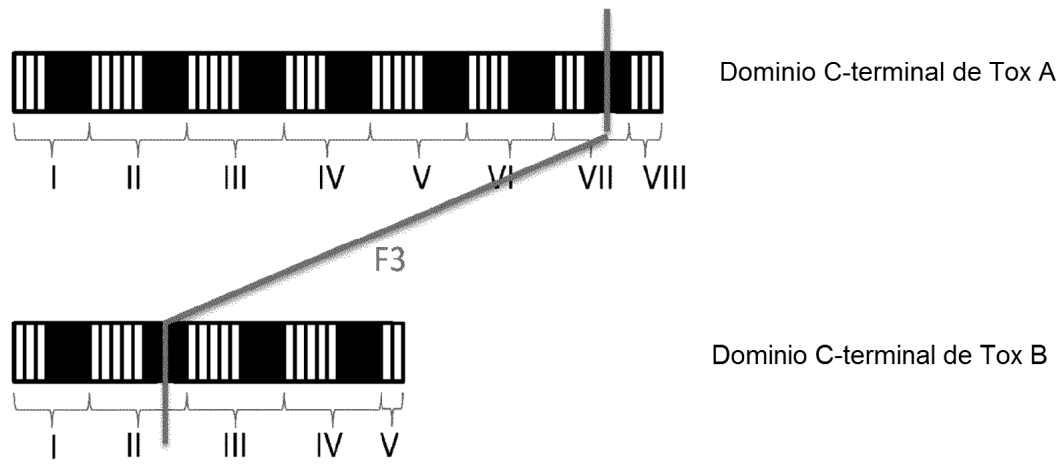


FIGURA 6

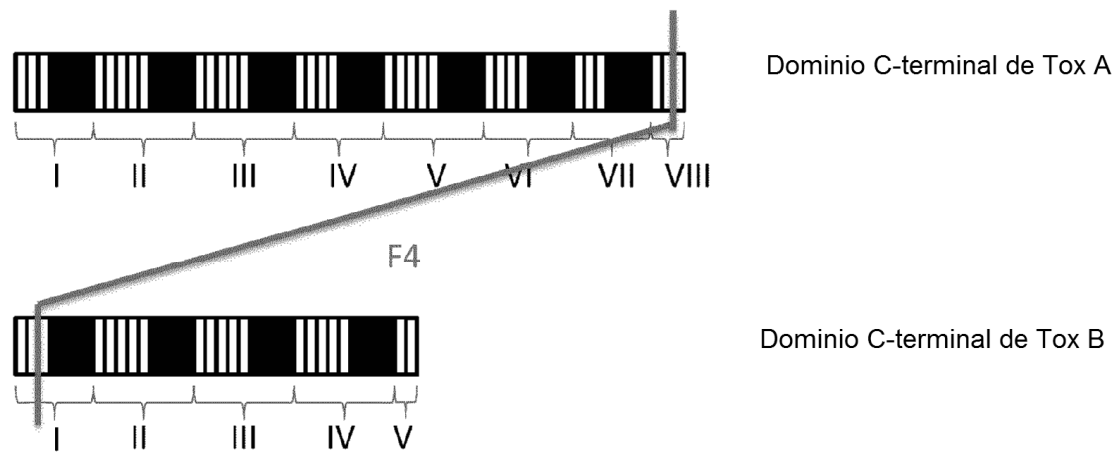
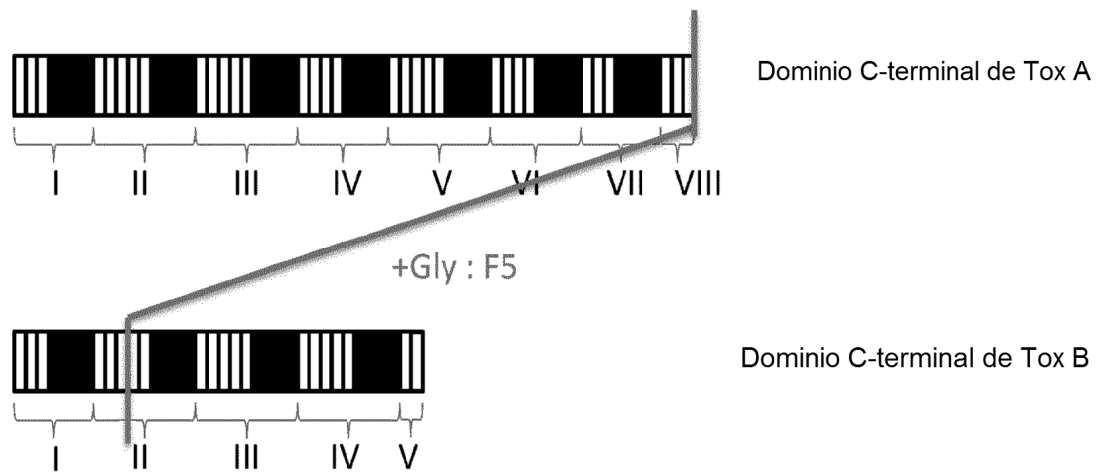
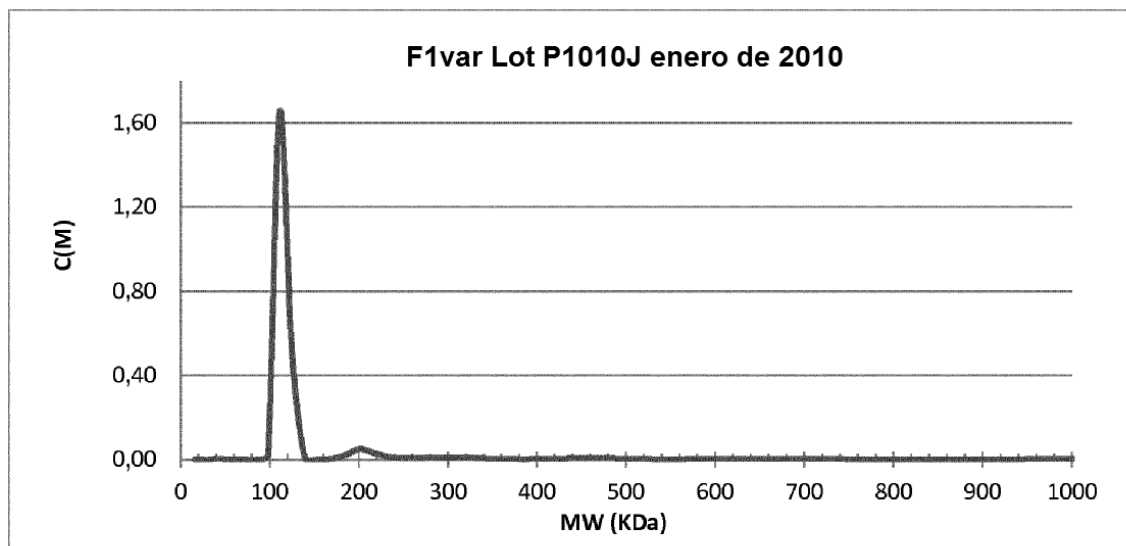


FIGURA 7

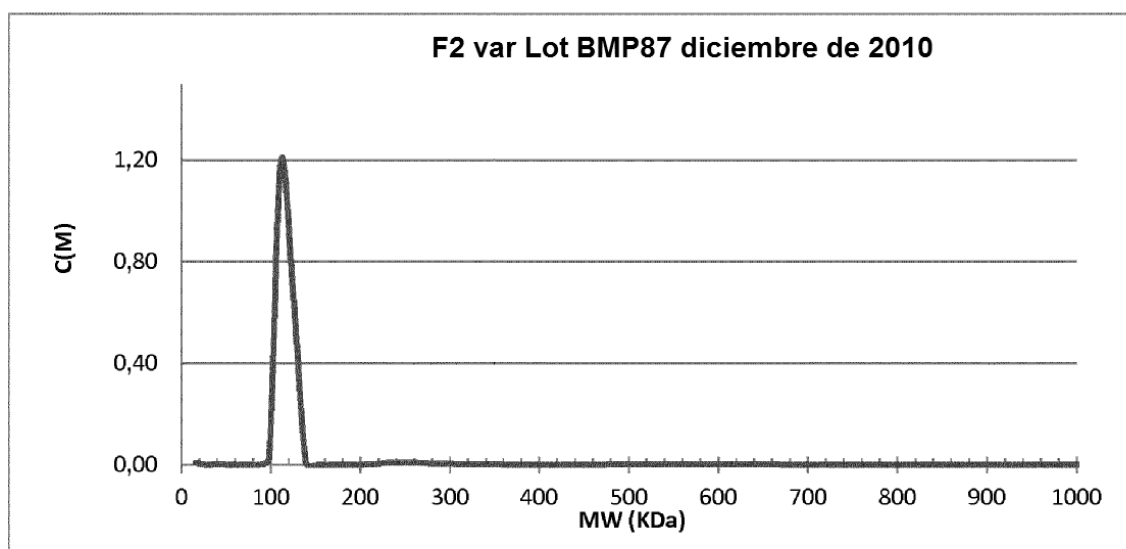


**FIGURA 8**

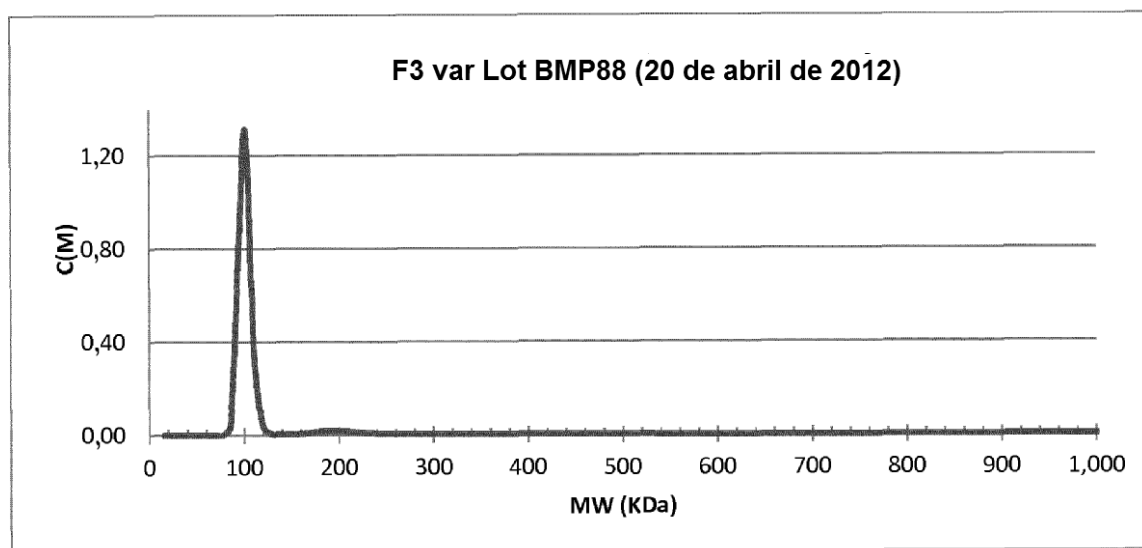
a)



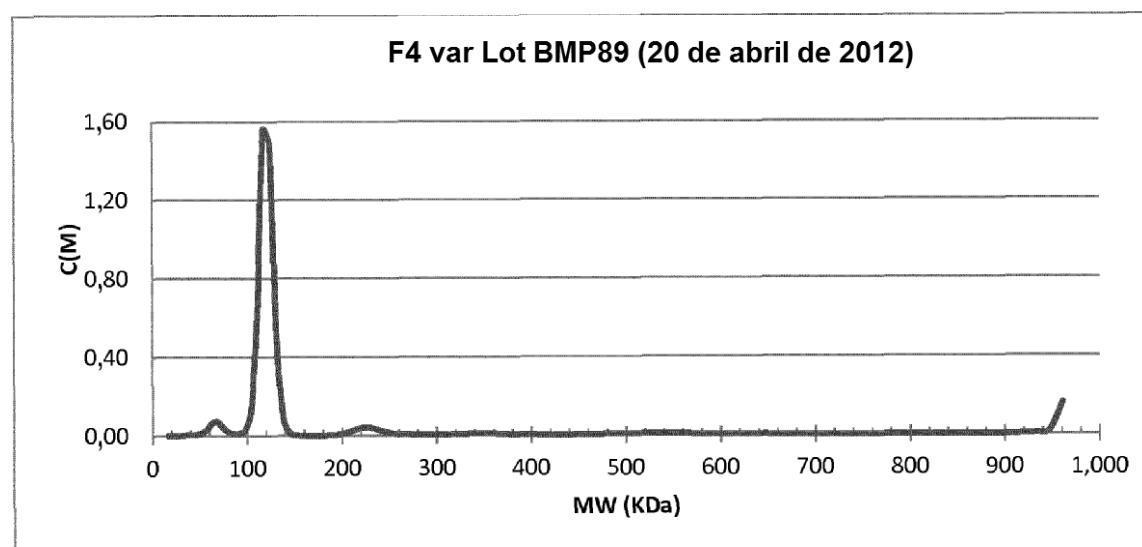
b)



c)



d)



e)

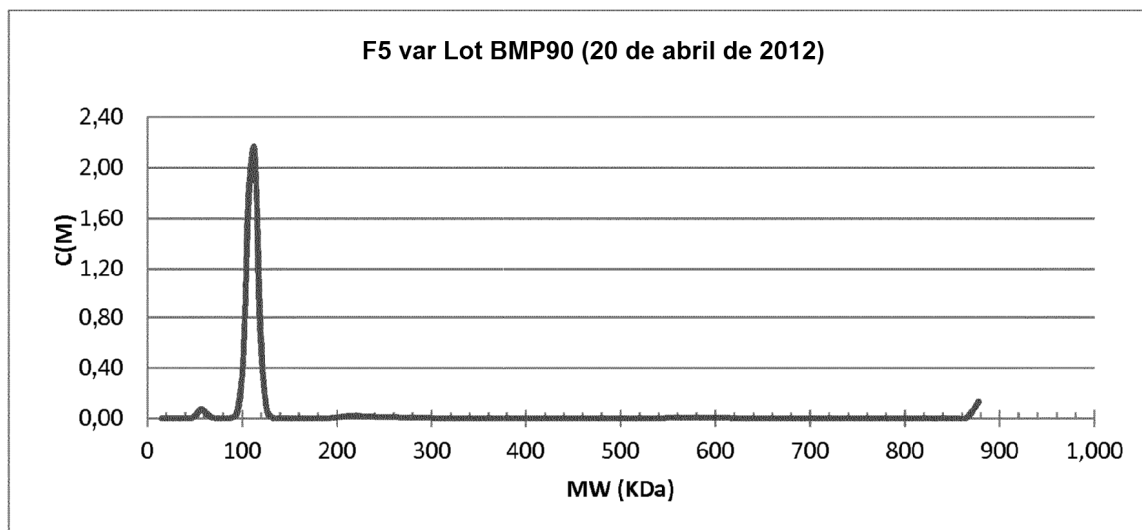


FIGURA 9

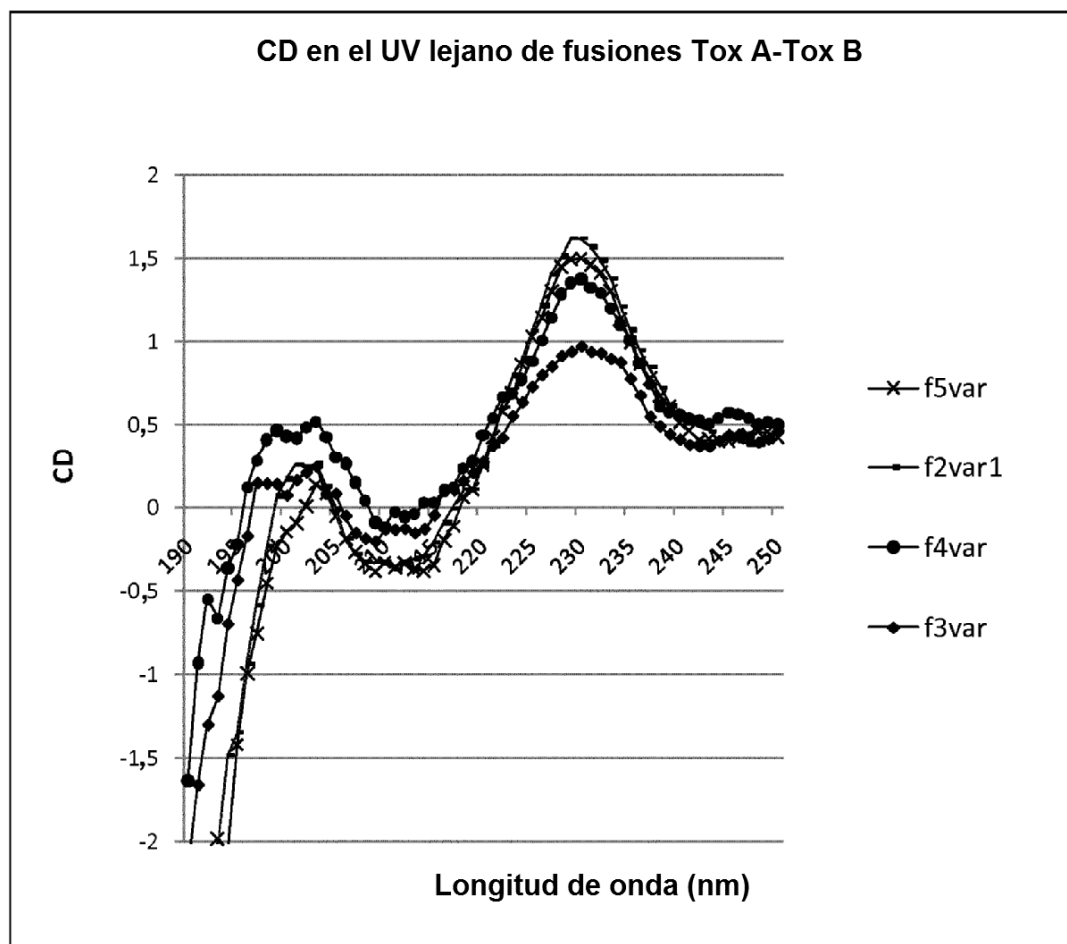


FIGURA 10

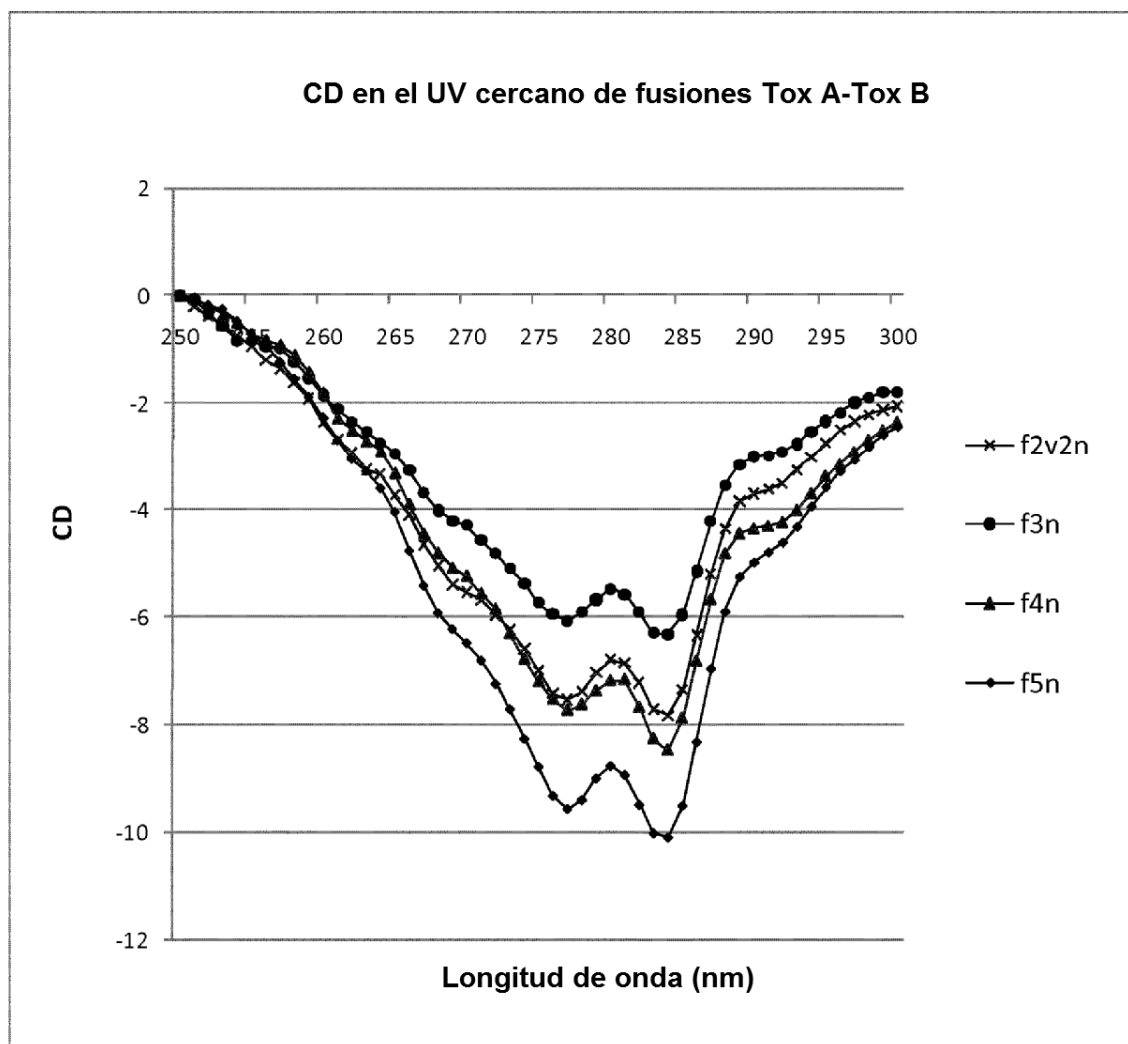


FIGURA 11

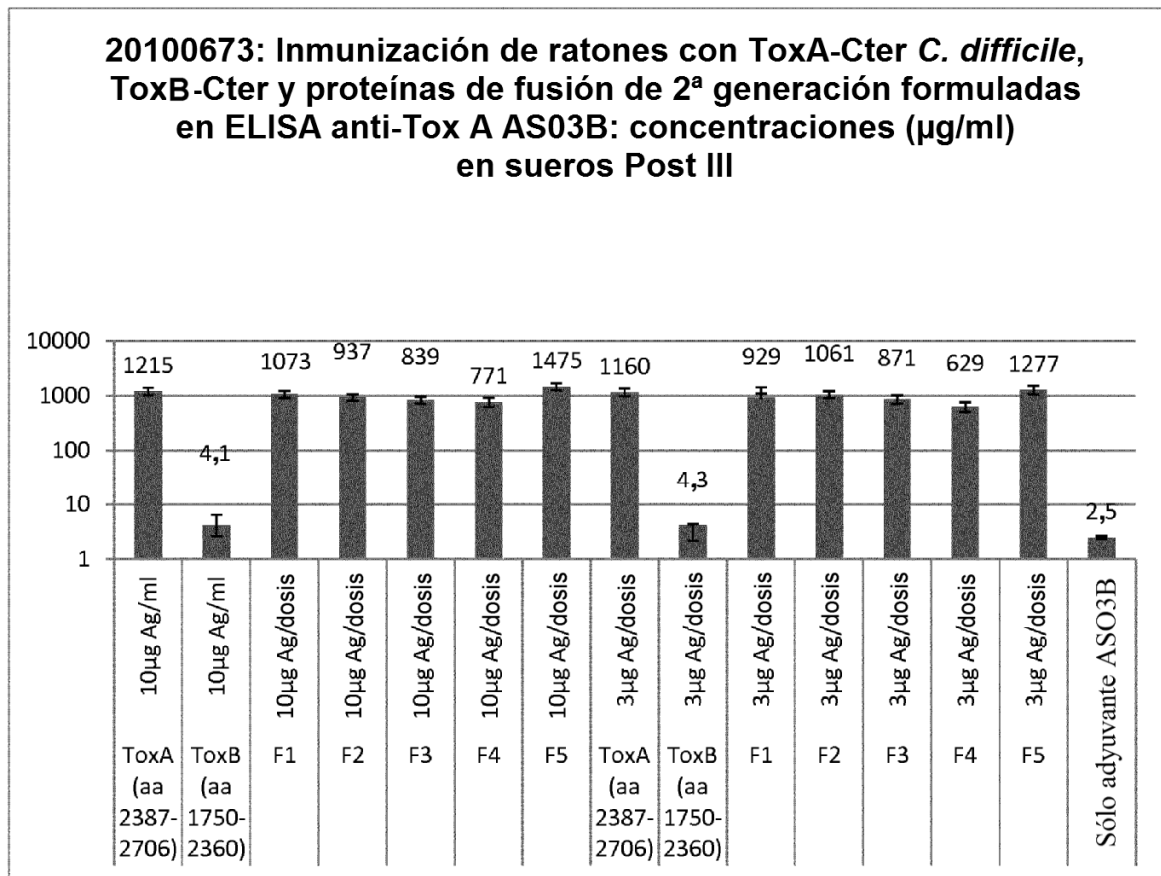


FIGURA 12

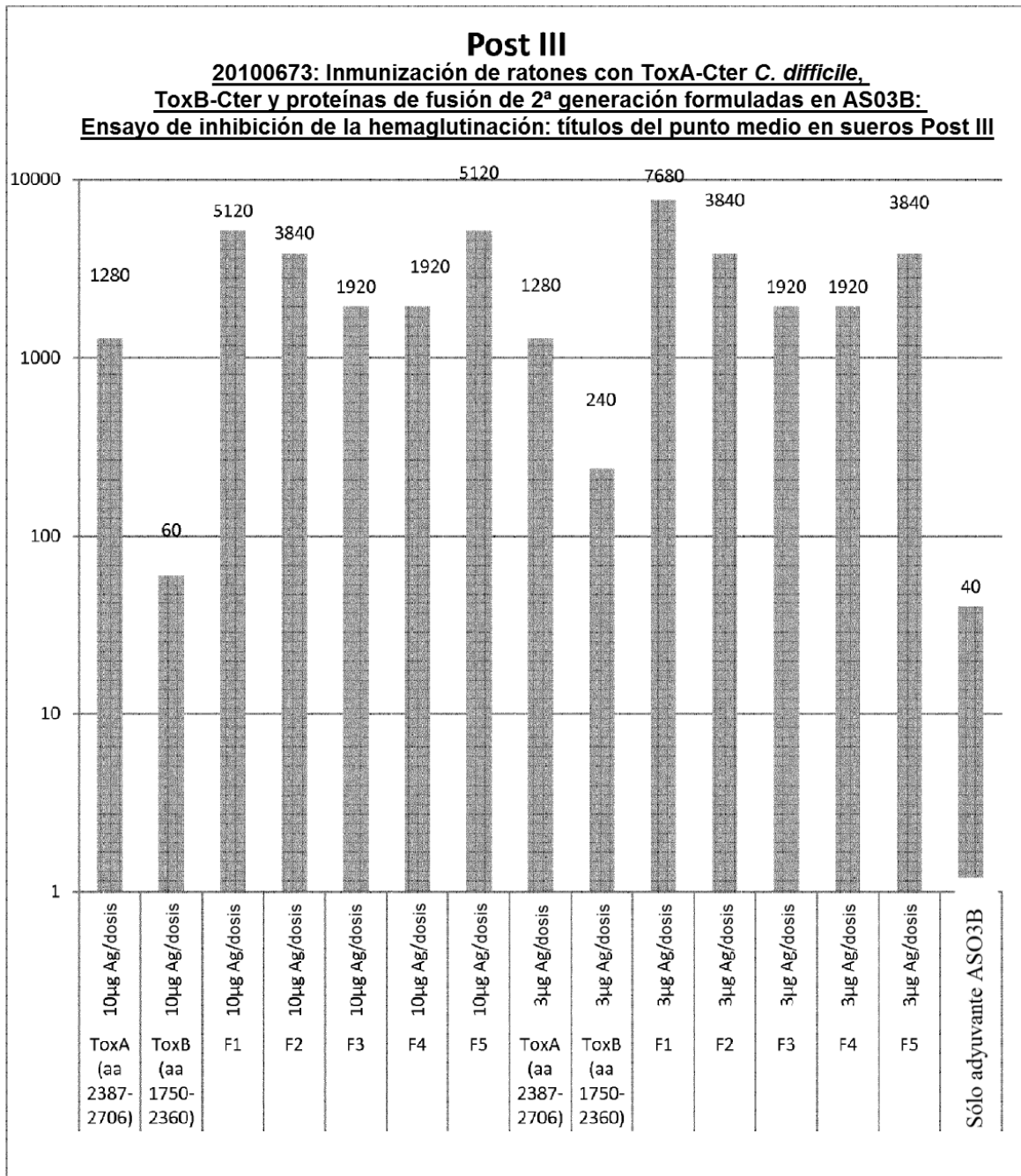


FIGURA 13

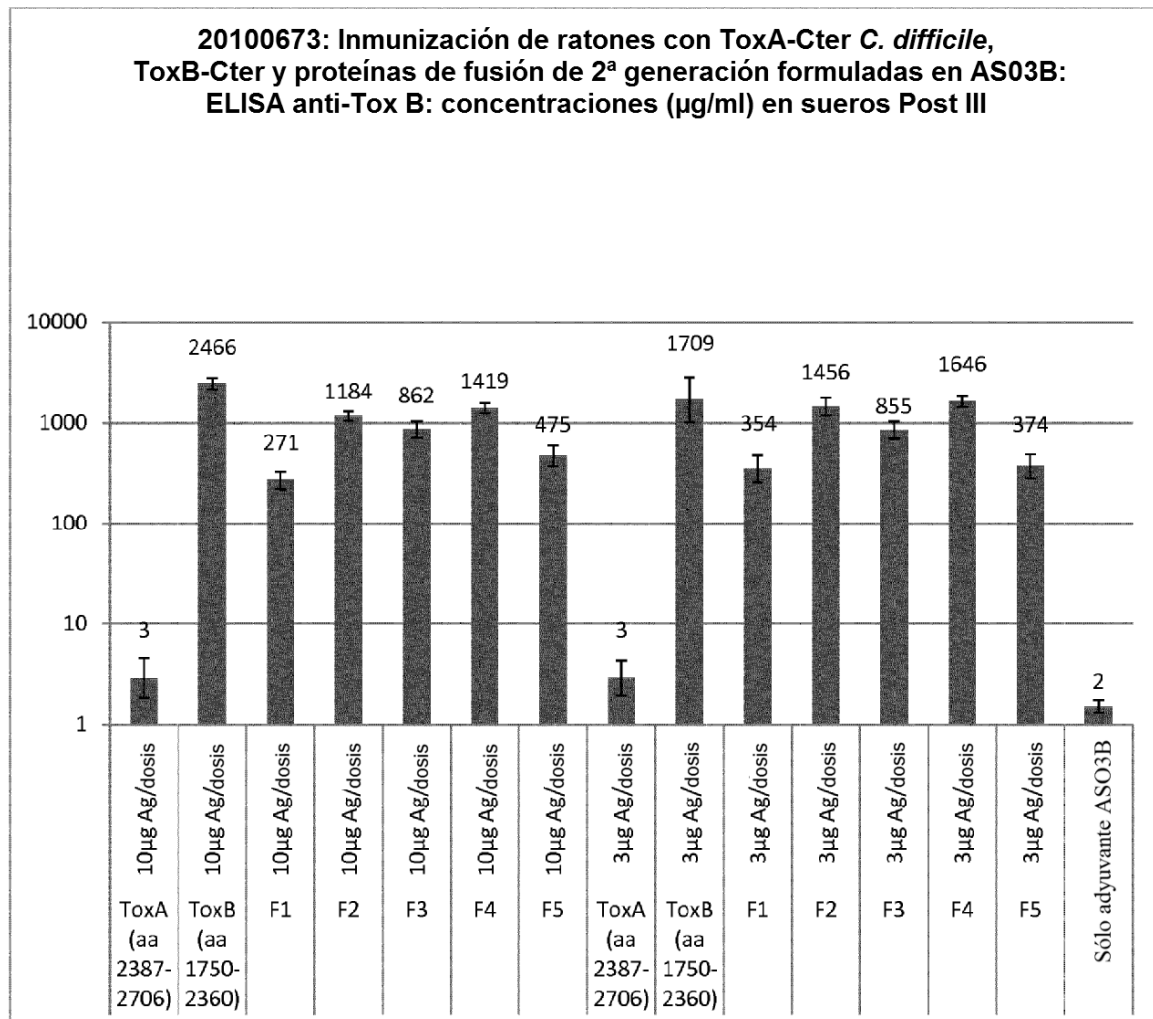
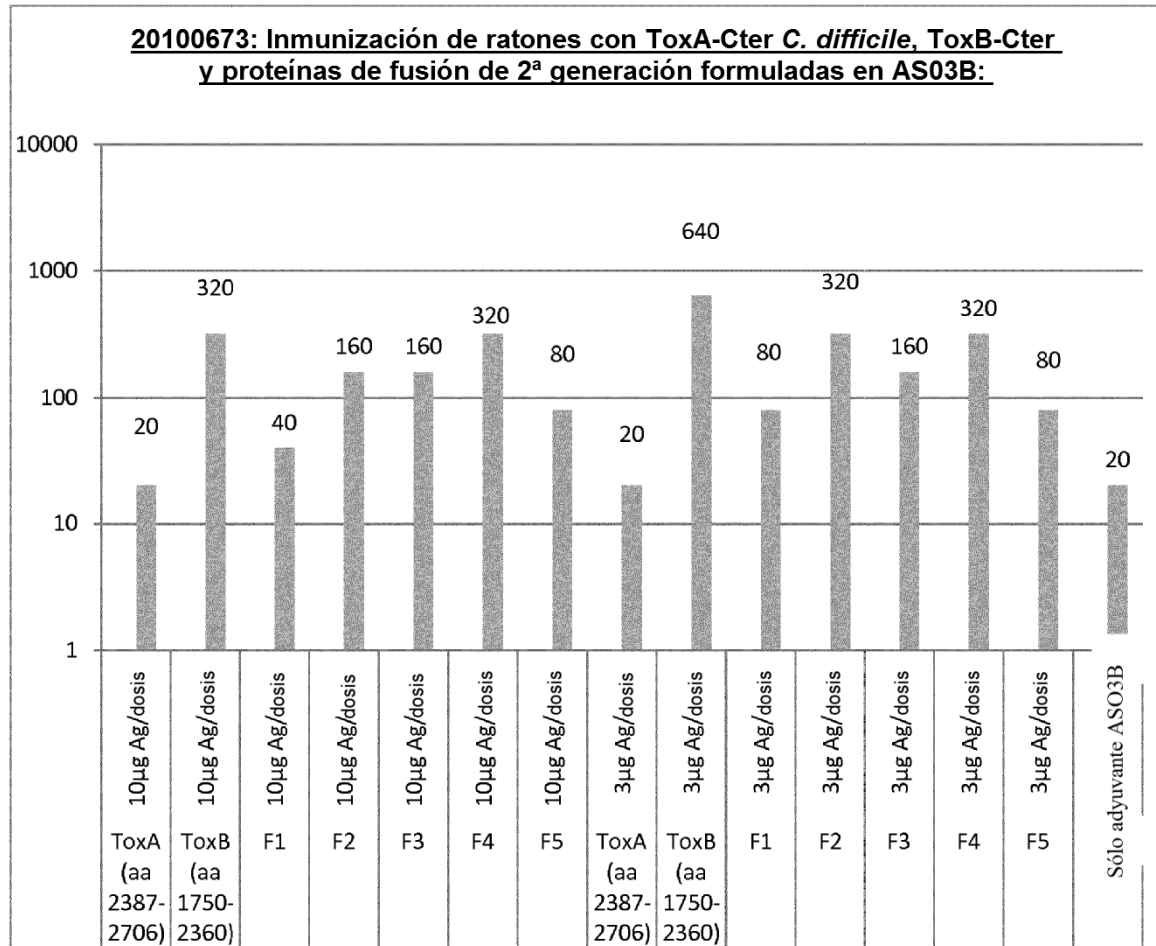
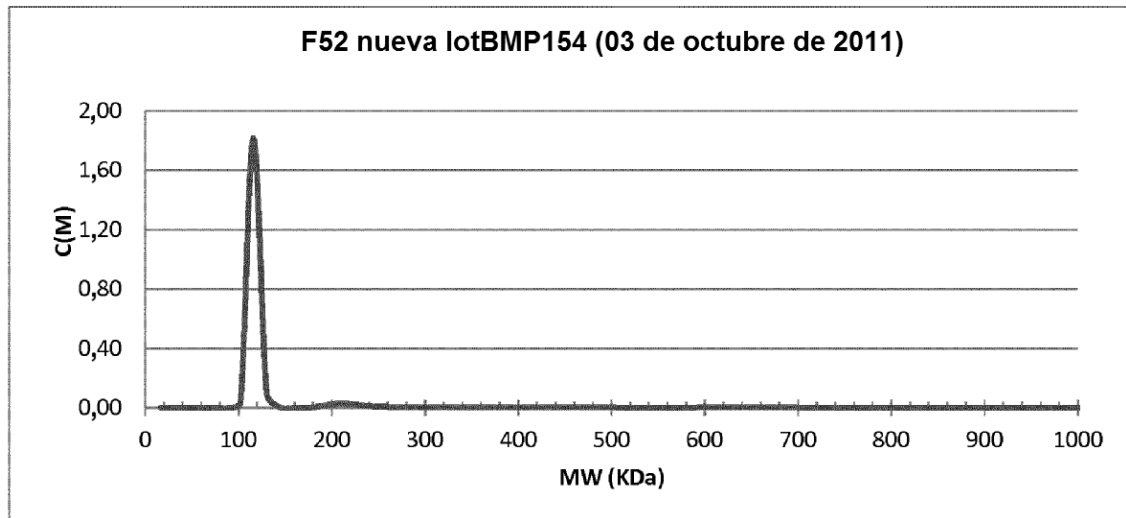


FIGURA 14

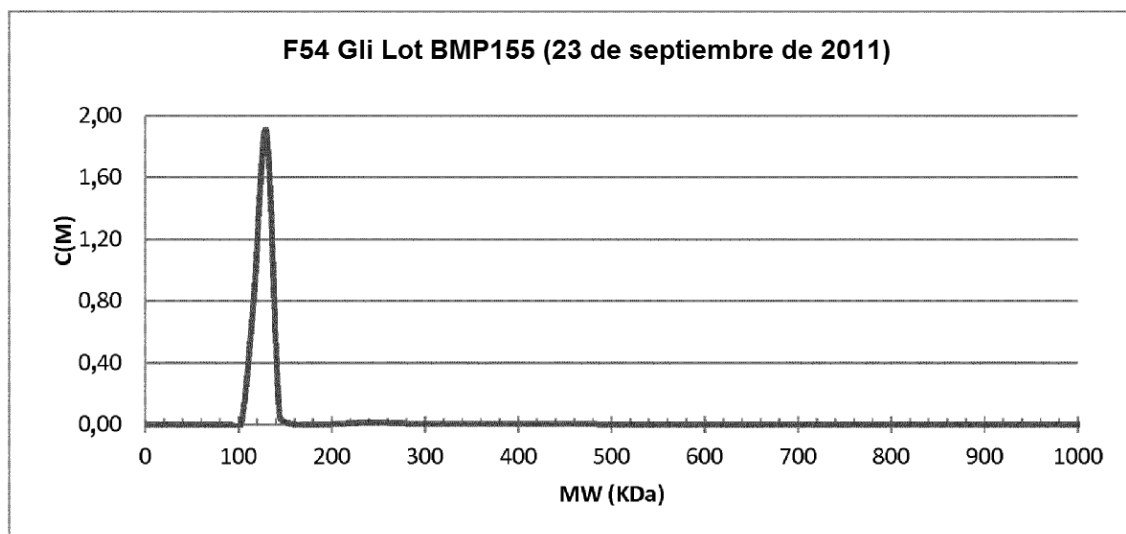


**FIGURA 15**

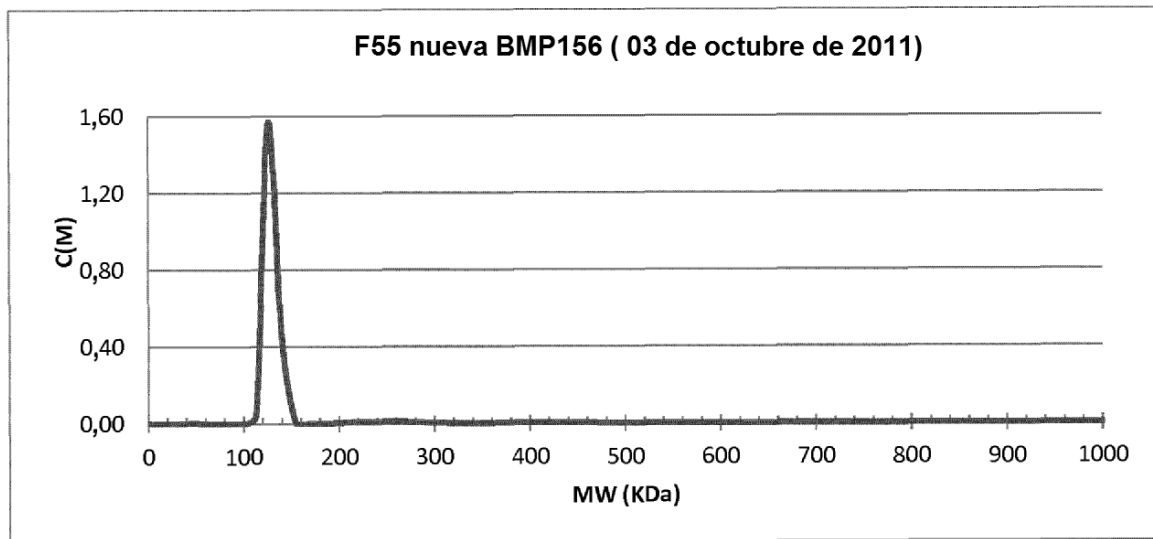
a)



b)



c)



d)

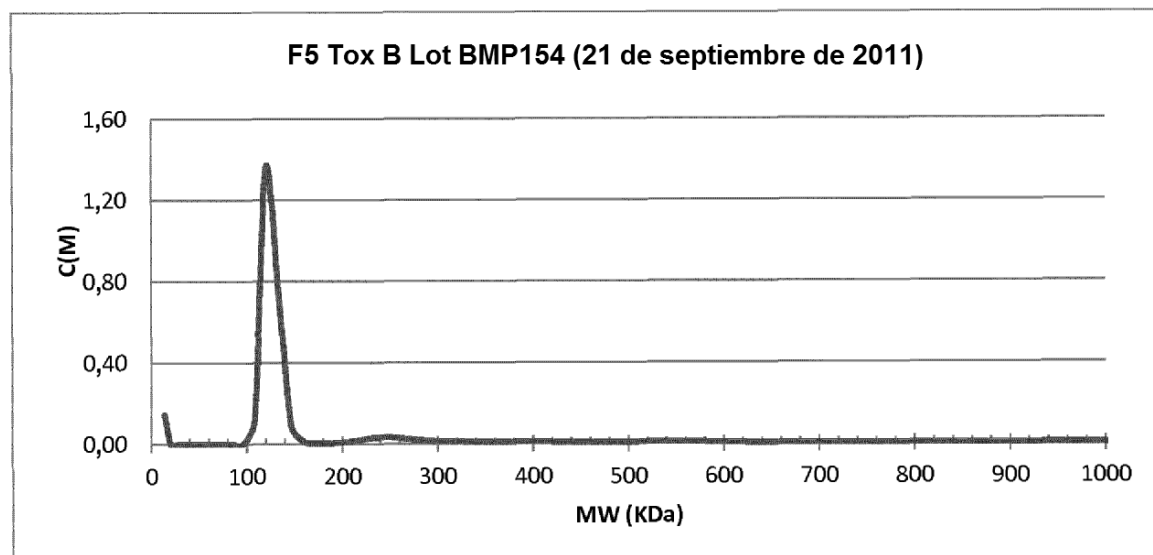


FIGURA 16

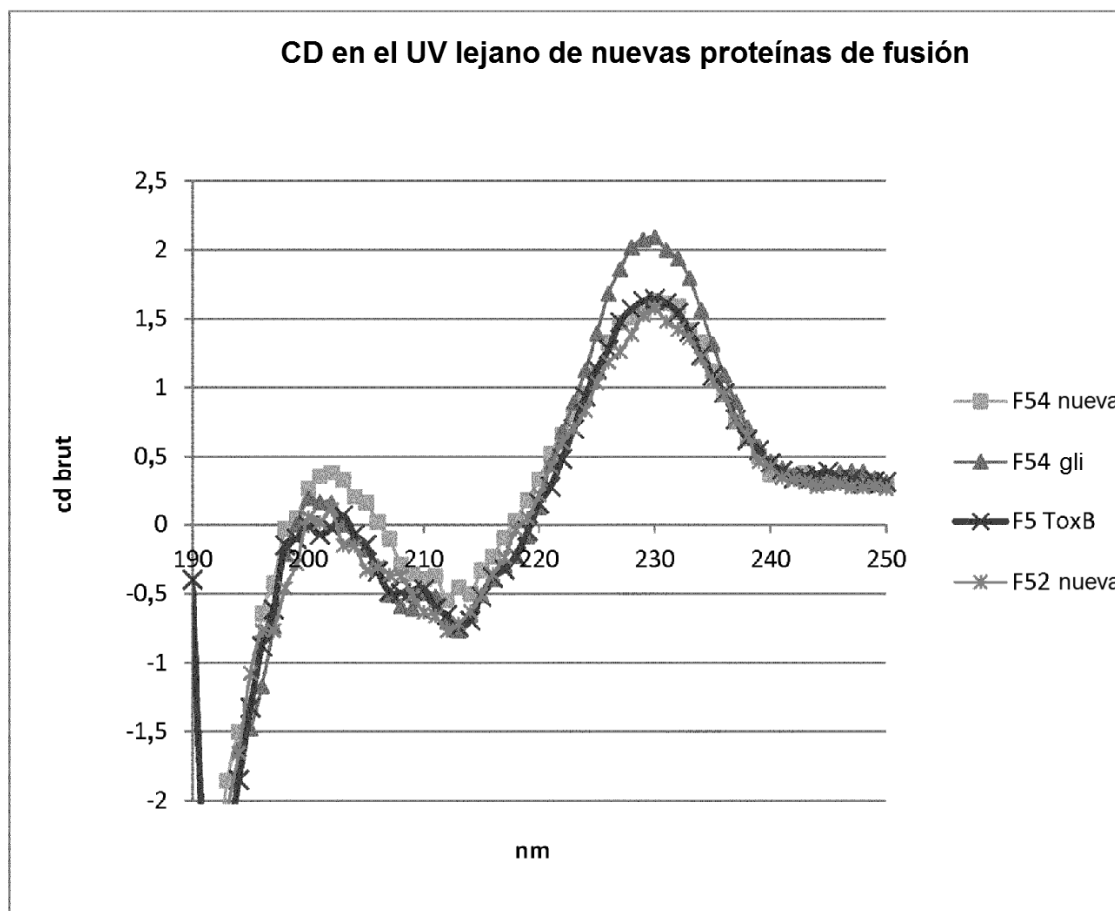


FIGURA 17

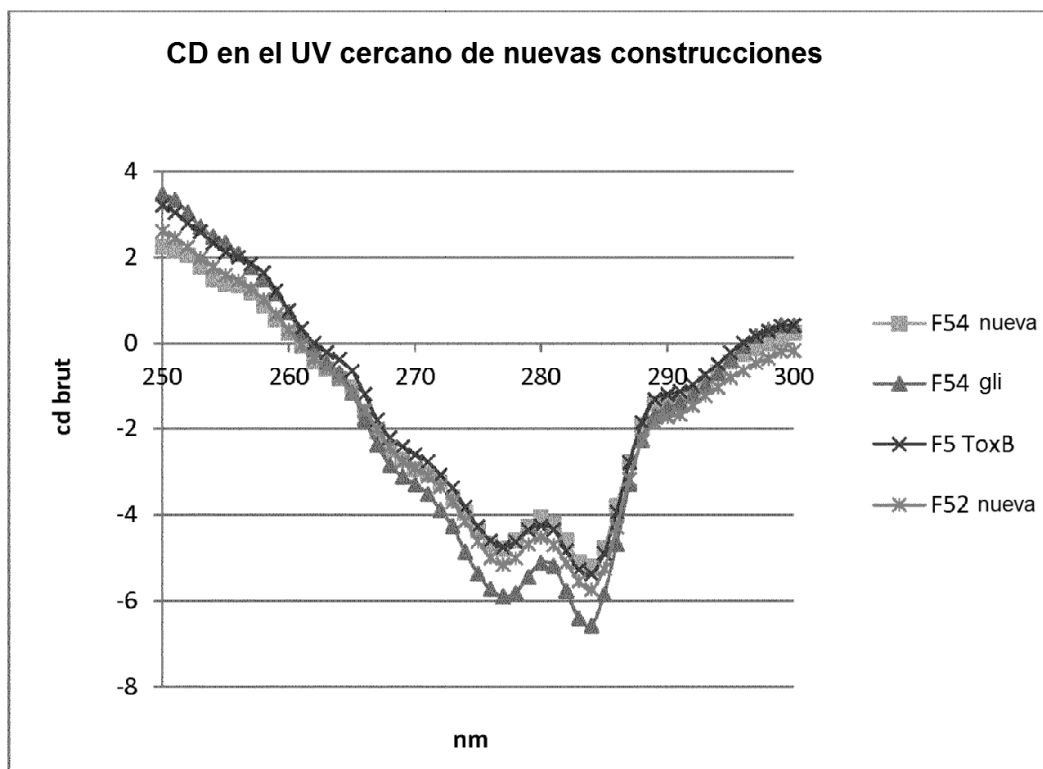


FIGURA 18

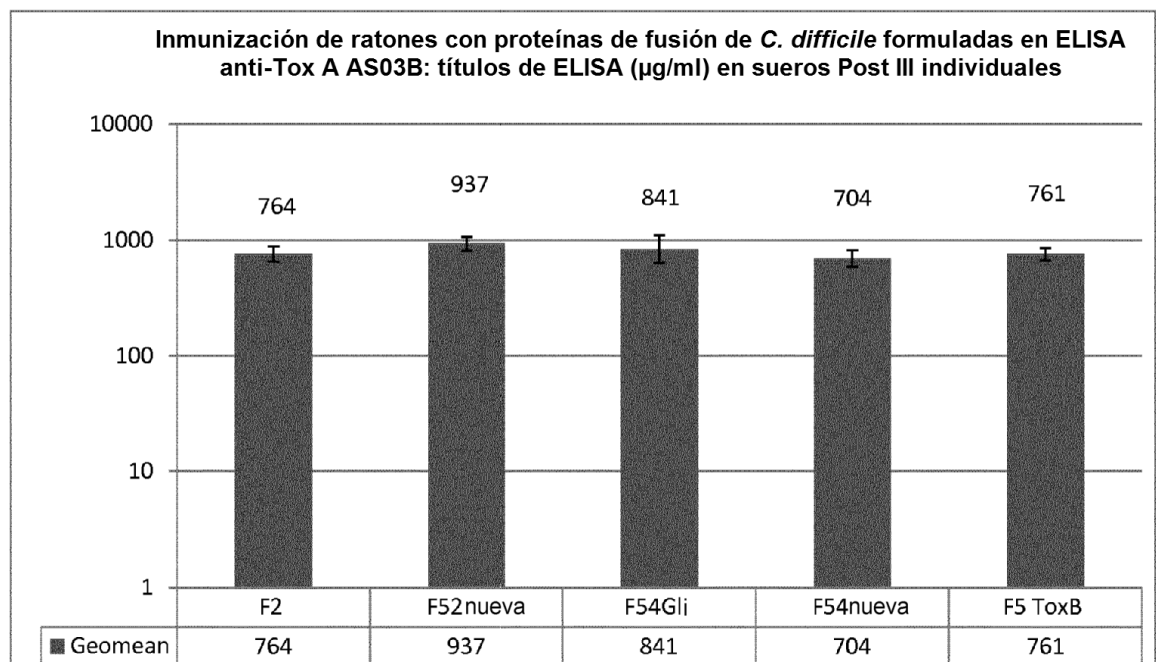


FIGURA 19

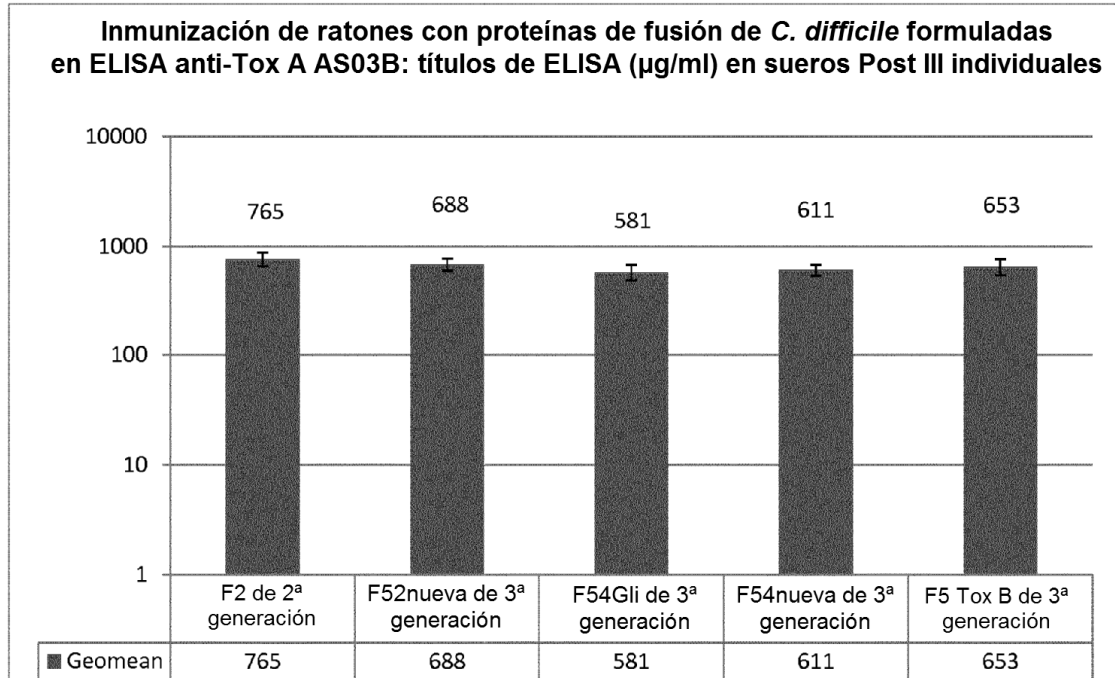


FIGURA 20

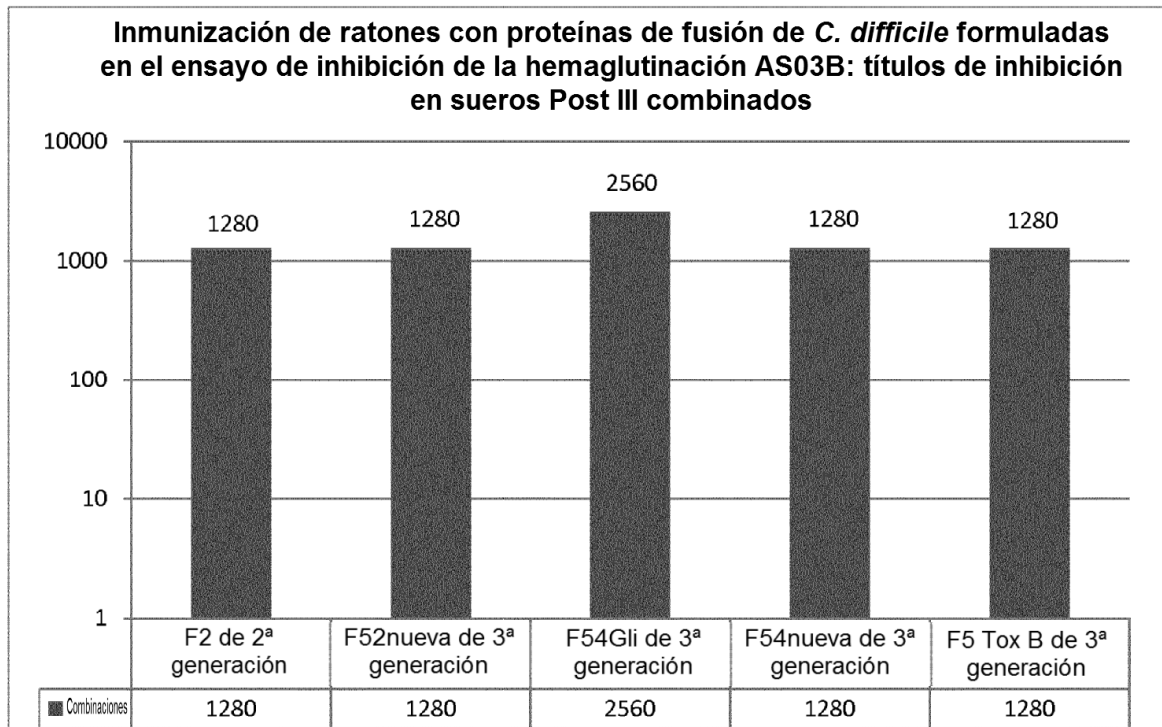


FIGURA 21

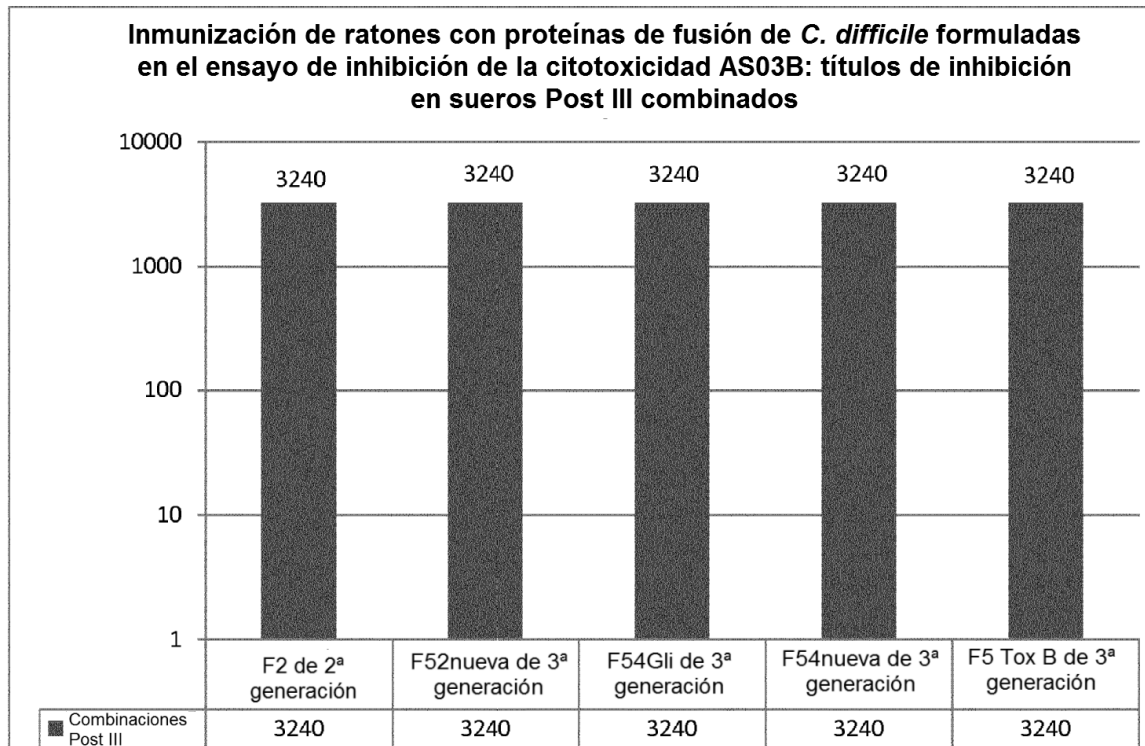


FIGURA 22

