

SIB EX3835R

VB64650 SM

Traduzione in lingua italiana del Brevetto Europeo  
domanda n° 12726374.7, pubblicazione n° 2714911  
a nome di GlaxoSmithKline Biologicals S.A.  
di 89 rue de l'Institut  
1330 Rixensart  
Belgio

“COMPOSIZIONE IMMUNOGENA”

Jacopo de Benedetti  
USBM-043R B

## DESCRIZIONE

### CAMPO TECNICO

La presente invenzione riguarda antigeni da *Clostridium difficile*. In particolare l'invenzione riguarda antigeni proteici ricombinanti comprendenti frammenti di tossina A e/o tossina B. L'invenzione riguarda inoltre composizioni immunogene oppure vaccini comprendenti questi antigeni, e l'uso dei vaccini e delle composizioni immunogene dell'invenzione nella profilassi oppure nella terapia. L'invenzione riguarda anche metodi di immunizzazione utilizzando le composizioni dell'invenzione, e l'uso delle composizioni dell'invenzione nella produzione di un medicamento.

### FONDAMENTO

*C.difficile* è la più importante causa di infezioni intestinali nosocomiali ed è la maggiore causa di colite pseudomembranosa negli esseri umani (Bartlett et al *Am. J. Clin. Nutr.* 11 suppl:2521-6 (1980)). Il tasso di mortalità associato complessivo per individui infettati con *C.difficile* è stato calcolato che è 5.99% entro 3 mesi dalla diagnosi, con mortalità più alta associata con l'età avanzata, essendo 13.5% in pazienti oltre 80 anni (karas et al *Journal of Infection* 561:1-9 (2010)). L'attuale trattamento per infezione da *C. difficile* è la somministrazione di antibiotici (metronidazolo e vancomicina), tuttavia vi sono prove di ceppi che sono resistenti a questi antibiotici (Shah et al., *Expert Rev. Anti Infect. Ther.* 8(5), 555-564 (2010)). Di conseguenza vi è una necessità di composizioni immunogene capaci di indurre anticorpi, e/o una risposta immunitaria protettiva, contro *C.difficile*.

### BREVE SOMMARIO

L'enterotossicità di *C. difficile* è principalmente dovuta all'azione di due tossine, tossina A e tossina B. Queste sono entrambe potenti citotossine (Lyerly et al *Current Microbiology* 21:29-32 (1990)). I domini C-terminali di tossina A e tossina B comprendono unità ripetitive, per esempio il dominio C-terminale di tossina A è costituito da unità ripetitive contigue (Dove et al *Infect. Immun.* 58:480-499 (1990)), per questa ragione il dominio C-terminale può essere riferito come il "dominio ripetitivo". Queste porzioni di ripetizioni possono essere separate inoltre in ripetizioni corte (SR) e ripetizioni lunghe (LR) come descritto in Ho et al (PNAS 102:18373-

18378 (2005)).

La struttura di un frammento di 127-aa dal C terminale del dominio ripetitivo di tossina A è stata determinata (Ho *et al* PNAS 102:18373-18378 (2005)). Questo frammento formava un ripiegamento tipo  $\beta$ -solenioide, composto prevalentemente da  $\beta$  filamenti con una bassa percentuale di  $\alpha$  eliche.

E' stato dimostrato che frammenti di tossina A, in particolare frammenti del dominio C-terminale, possono portare a una risposta immunitaria protettiva nei criceti (Lyerly *et al* Current Microbiology 21:29-32 (1990)), WO96/12802 e WO00/61762.

E' nota la difficoltà insita nella progettazione di proteine di fusione che si ripieghino correttamente durante l'espressione. I polipeptidi della presente invenzione sono proteine di fusione in cui è mantenuta la struttura tipo  $\beta$  solenoide nativa, e che si osserva che forniscono una risposta immunitaria contro sia tossina A sia tossina B nei topi.

In un primo aspetto dell'invenzione vi è fornito un polipeptide comprendente un primo frammento e un secondo frammento, in cui

- (i) il primo frammento è un frammento di dominio ripetitivo di tossina A;
- (ii) il secondo frammento è un frammento di dominio ripetitivo di tossina B;
- (iii) il primo frammento ha una prima estremità prossimale;
- (iv) il secondo frammento ha una seconda estremità prossimale; e

in cui il primo frammento e il secondo frammento sono separati da meno di oppure esattamente 5 amminoacidi nella struttura primaria, in cui il polipeptide suscita anticorpi che neutralizzano sia tossina A sia tossina B, in cui la prima estremità prossimale è all'interno di una ripetizione corta, la seconda estremità prossimale è all'interno di una ripetizione corta e la prima estremità prossimale e la seconda estremità prossimale non interrompono porzioni di ripetizione corta-ripetizione lunga-ripetizione corta.

In un secondo aspetto dell'invenzione vi è fornito un polinucleotide codificante per un polipeptide dell'invenzione.

In un terzo aspetto dell'invenzione vi è fornito un vettore comprendente un polinucleotide dell'invenzione legato a un promotore inducibile.

In un quarto aspetto dell'invenzione vi è fornita una cellula ospite comprendente un vettore dell'invenzione oppure un polinucleotide dell'invenzione.

In un quinto aspetto dell'invenzione vi è fornita una composizione immunogena comprendente un polipeptide dell'invenzione e un eccipiente farmaceuticamente accettabile.

In un sesto aspetto dell'invenzione vi è fornito un vaccino comprendente una composizione immunogena dell'invenzione.

In un settimo aspetto dell'invenzione vi è fornito un uso di una composizione immunogena dell'invenzione oppure di un vaccino dell'invenzione per uso nel trattamento oppure nella prevenzione di malattia da *C.difficile*.

#### BREVE DESCRIZIONE DEI DISEGNI

FIG. 1- Liste di sequenze di polipeptidi dell'invenzione.

FIG. 2 - Rappresentazione figurata dei domini C-terminali di ToxA e ToxB, con le ripetizioni SR rappresentate come box bianche e le box di LR rappresentate come box nere.

FIG. 3 - Rappresentazione figurata di una giunzione tra la terza SR VIII di ToxA e la quarta SR II di ToxB usata nella Fusione 1.

FIG. 4 - Rappresentazione figurata di una giunzione tra la seconda SR VIII di ToxA e la terza SR II di ToxB usata nella Fusione 2.

FIG.5 - Rappresentazione figurata di una giunzione tra LRVII di ToxA e LR II di ToxB usata nella Fusione 3 (contenente soltanto parte di LRVII di ToxA e parte di LR II di ToxB).

FIG.6 - Rappresentazione figurata di una giunzione tra la seconda SR VIII di ToxA e la terza SR I di ToxB usate nella Fusione 4.

FIG.7 - Rappresentazione figurata di una giunzione comprendente un linker di glicine tra l'ultimo residuo della sequenza di proteina ToxA e l'inizio della quarta SR II di ToxB usato nella Fusione 5.

FIG.8 - Grafici che descrivono la distribuzione di fusioni 1-5 ToxA-ToxB di *C.difficile* come determinata

secondo la velocità di sedimentazione mediante ultracentrifugazione analitica. Riquadro a) descrive la distribuzione di Fusione 1, riquadro b) descrive la distribuzione di Fusione 2, riquadro c) descrive la distribuzione di Fusione 3, riquadro d) descrive la distribuzione di Fusione 4 e riquadro e) descrive la distribuzione di Fusione 5.

FIG.9 - Grafico che descrive lo spettro Lontano-UV di Fusioni, 2, 3, 4, e 5 misurato usando dicroismo circolare. Lo spettro per fusione 2 è rappresentato da una linea con i punti rappresentati come piccoli quadrati, lo spettro per fusione 3 è rappresentato da una linea con i punti rappresentati come piccole forme a rombo, fusione 4 è rappresentata da una linea con i punti rappresentati come cerchi, e fusione 5 è rappresentata da una linea con i punti rappresentati come forme a croce.

FIG.10 - Grafico che descrive lo spettro vicino-UV di Fusioni 2, 3, 4, e 5 misurato usando dicroismo circolare. Lo spettro per fusione 2 è rappresentato da una linea con i punti rappresentati come forme a croce, lo spettro per fusione 3 è rappresentato da una linea con i punti rappresentati come cerchi, lo spettro per fusione 4 è rappresentato da una linea con i punti rappresentati come triangoli, e lo spettro per fusione 5 è rappresentato da una linea con i punti rappresentati come forme a piccoli rombi.

FIG.11 - Grafico che mostra immunogenicità anti-ToxA in topi immunizzati con un frammento del C-terminale di tossina A (aa 2387-2706), un frammento del C-terminale di tossina B (aa 1750-2360), oppure fusioni 1, 2, 3, 4 oppure 5.

FIG.12 - Grafico che mostra inibizione di emoagglutinazione in topi immunizzati con un frammento del C-terminale di tossina A (aa 2387-2706), un frammento del C-terminale di tossina B (aa 1750-2360), oppure fusioni 1, 2, 3, 4 oppure 5.

FIG.13 - Grafico che mostra immunogenicità anti-ToxB in topi immunizzati con un frammento del C-terminale di tossina A (aa 2387-2706), un frammento del C-terminale di tossina B (aa 1750-2360), oppure fusioni 1, 2, 3, 4 oppure 5.

FIG.14 – Titoli di inibizione di citotossicità da topi immunizzati con un frammento del C-terminale di tossina A (aa 2387-2706), un frammento del C-terminale di tossina B (aa 1750-2360), oppure fusioni 1, 2, 3, 4 oppure 5.

FIG.15 - Grafici che descrivono la distribuzione di fusioni ToxA-ToxB di *C. difficile* F52New, F54Gly, F54New e F5ToxB come determinata secondo la velocità di sedimentazione mediante ultracentrifugazione analitica. Riquadro a) descrive la distribuzione di F52New, riquadro b) descrive la distribuzione di F54Gly, riquadro c) descrive la distribuzione di F54New e riquadro d) descrive la distribuzione di F5ToxB.

FIG.16 - Grafico che descrive lo spettro Lontano-UV di fusioni F52New, F54Gly, F54New e F5ToxB misurato usando dicroismo circolare. Lo spettro per F52New è rappresentato da una linea con i punti rappresentati come croci doppie, lo spettro per F54Gly è rappresentato da una linea con i punti rappresentati come triangoli, F54New è rappresentata da una linea con i punti rappresentati come quadrati, e F5ToxB è rappresentata da una linea con i punti rappresentati come forme a croce.

FIG.17 - Grafico che descrive lo spettro Vicino-UV di fusioni F52New, F54Gly, F54New e F5ToxB misurato usando dicroismo circolare. Lo spettro per F52New è rappresentato da una linea con i punti rappresentati come croci doppie, lo spettro per F54Gly è rappresentato da una linea con i punti rappresentati come triangoli, F54New è rappresentata da una linea con i punti rappresentati come quadrati, e F5ToxB è rappresentata da una linea con i punti rappresentati come forme a croce.

FIG.18 - Grafico che mostra risultati ELISA anti-ToxA per topi immunizzati con le fusioni F2, F52New, F54Gly, F54New oppure F5ToxB.

FIG.19 - Grafico che mostra risultati ELISA anti-ToxB per topi immunizzati con le fusioni F2, F52New, F54Gly, F54New oppure F5ToxB.

FIG.20 - Grafico che mostra inibizione di emoagglutinazione in topi immunizzati con le fusioni F2, F52New, F54Gly, F54New oppure F5ToxB.

FIG.21 - Grafico che mostra titoli di citotossicità in cellule HT29 da topi immunizzati con le fusioni F2, F52New, F54Gly, F54New oppure F5ToxB.

FIG.22 - Grafico che mostra titoli di citotossicità in cellule IMR90 da topi immunizzati con le fusioni F2, F52New, F54Gly, F54New oppure F5ToxB.

## DESCRIZIONE DETTAGLIATA

### POLIPEPTIDI

L'invenzione riguarda un polipeptide comprendente un primo frammento e un secondo frammento, in cui

- (i) il primo frammento è un frammento di dominio ripetitivo di tossina A;
- (ii) il secondo frammento è un frammento di dominio ripetitivo di tossina B;
- (iii) il primo frammento ha una prima estremità prossimale;
- (iv) il secondo frammento ha una seconda estremità prossimale; e

in cui il primo frammento e il secondo frammento sono separati da meno di oppure esattamente 5 amminoacidi nella struttura primaria, in cui il polipeptide suscita anticorpi che neutralizzano sia tossina A sia tossina B, in cui la prima estremità prossimale è all'interno di una ripetizione corta, la seconda estremità prossimale è all'interno di una ripetizione corta e la prima estremità prossimale e la seconda estremità prossimale non interrompono porzioni di ripetizione corta-ripetizione lunga-ripetizione corta.

Il termine polipeptide si riferisce a una sequenza contigua di amminoacidi.

Il termine "dominio ripetitivo di tossina A" si riferisce al dominio C-terminale della proteina tossina A da *C.difficile*, comprendente sequenze ripetute. Questo dominio si riferisce a amminoacidi 1832-2710 di tossina A da ceppo VPI10463 (ATCC43255) e loro equivalenti in un ceppo differente, la sequenza di amminoacidi 1832-2710 da ceppo VPI10463 (ATCC43255) corrisponde a amminoacidi 1832-2710 di SEQ ID NO:1.

Il termine "dominio ripetitivo di tossina B" si riferisce al dominio C-terminale della proteina tossina B da *C.difficile*. Questo dominio si riferisce a amminoacidi 1834-2366 da ceppo VPI10463 (ATCC43255) e loro equivalenti in un ceppo differente, la sequenza di amminoacidi 1834-2366 da ceppo VPI10463 (ATCC43255) corrisponde a amminoacidi 1834-2366 di SEQ ID NO:2.

Le tossine A e B da *C.difficile* sono proteine conservate, tuttavia la sequenza differisce di una piccola quantità tra i ceppi, inoltre la sequenza amminoacidica per tossine A e B in ceppi differenti può differire nel numero di amminoacidi.

L'invenzione perciò il termine dominio ripetitivo di tossina A e/o dominio ripetitivo di tossina B per riferirsi a

una sequenza che è una variante con 90%, 95%, 98%, 99% o 100% di identità di sequenza con amminoacidi 1832-2710 di SEQ ID NO:1 oppure una variante con 90%, 95%, 98%, 99% o 100% di identità di sequenza con amminoacidi 1834-2366 di SEQ ID NO:2. In una forma di realizzazione una “variante” è un polipeptide che varia dai polipeptidi di riferimento per sostituzioni amminoacidiche conservative, per cui un residuo è sostituito con un altro con le stesse proprietà fisico-chimiche. Tipicamente tali sostituzioni sono tra Ala, Val, Leu e Ile; tra Ser e Thr; tra i residui acidi Asp e Glu; tra Asn e Gln, e tra i residui basici Lys e Arg; oppure residui aromatici Phe e Tyr. In una forma di realizzazione un “frammento” è un polipeptide che comprende una porzione contigua di almeno 250 amminoacidi di un polipeptide.

Inoltre la numerazione di amminoacidi può differire tra i domini C-terminali di tossina A (oppure tossina B) da un ceppo e tossina A (oppure tossina B) da un altro ceppo. Per questa ragione il termine “equivalenti in un ceppo” differente si riferisce a amminoacidi che corrispondono a quelli di un ceppo di riferimento (ad esempio, *C. difficile* VP110463), ma che si trovano in una tossina da un ceppo differente e che possono così essere numerati differentemente. Una regione di amminoacidi “equivalenti” può essere determinata allineando le sequenze delle tossine dai differenti ceppi. I numeri di amminoacidi forniti per tutta la presente si riferiscono a quelli di ceppo VP110463.

Il termine “frammento” di un polipeptide oppure di una proteina si riferisce a una porzione contigua di almeno 100, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 o 600 amminoacidi da quel polipeptide oppure da quella proteina. Il termine “primo frammento” si riferisce a una porzione contigua di almeno 100, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 o 600 amminoacidi del dominio ripetitivo di tossina A. Il termine “secondo frammento” si riferisce a una porzione contigua di almeno 100, 200, 230, 250, 280, 300, 350, 400, 450 o 500 amminoacidi del dominio ripetitivo di tossina B.

Il termine “prima estremità prossimale” si riferisce alla estremità del primo frammento (frammento ToxA) che è covalentemente legata al secondo frammento (frammento ToxB) oppure covalentemente legata a una sequenza linker tra il primo e il secondo frammento ed è più vicina al secondo frammento nella struttura primaria. Il termine “seconda estremità prossimale” si riferisce all’estremità del secondo frammento che è covalentemente

legata al primo frammento ( frammento ToxA) oppure covalentemente legata a una sequenza linker tra il primo e il secondo frammento ed è più vicina al primo frammento nella struttura primaria.

Il polipeptide può essere parte di una proteina più grande quale un precursore oppure una proteina di fusione. E' spesso vantaggioso includere una sequenza amminoacidica aggiuntiva che contiene sequenze che favoriscono la purificazione quali residui di istidina multipli, oppure una sequenza aggiuntiva per la stabilità durante la produzione ricombinante. Inoltre, si considera anche l'aggiunta di polipeptide esogeno oppure di coda lipidica oppure di sequenze polinucleotidiche per aumentare il potenziale immunogeno della molecola finale.

I frammenti possono essere posizionati in modo tale che l'N-terminale del primo frammento sia adiacente al C-terminale del secondo frammento, in alternativa il C-terminale del primo frammento può essere adiacente all'N-terminale del secondo frammento, oppure il C-terminale del primo frammento può essere adiacente al C-terminale del secondo frammento, oppure l'N-terminale del primo frammento può essere adiacente all'N-terminale del secondo frammento.

La parola "adiacente" significa separato da meno di oppure esattamente 5, 2, 1 o 0 amminoacidi nella struttura primaria.

Il polipeptide dell'invenzione suscita anticorpi che neutralizzano tossina A oppure tossina B oppure ambedue. In una forma di realizzazione il polipeptide suscita anticorpi che neutralizzano tossina A. In un'ulteriore forma di realizzazione il polipeptide suscita anticorpi che neutralizzano tossina B. In un'ulteriore forma di realizzazione il polipeptide suscita anticorpi che neutralizzano tossina A e tossina B.

Se un polipeptide suscita anticorpi contro una tossina può essere misurato immunizzando topi con una composizione immunogena comprendente il polipeptide, raccogliendo sieri e analizzando i titoli anti-tossina dei sieri usando ELISA. I sieri dovrebbero essere confrontati con un campione di riferimento ottenuto da topi che non sono stati immunizzati. Un esempio di questa tecnica si può trovare nell'esempio 6. Il polipeptide dell'invenzione suscita anticorpi che neutralizzano tossina A se i sieri contro il polipeptide danno una lettura ELISA più di 10%, 20%, 30%, 50%, 70%, 80%, 90%, o 100% più alta del campione di riferimento.

In un'ulteriore forma di realizzazione il polipeptide dell'invenzione suscita una risposta immunitaria protettiva in

un ospite mammifero contro ceppi di *C.difficile*. In una forma di realizzazione l'ospite mammifero è selezionato dal gruppo composto da topo, coniglio, porcellino d'India, primate non umano, scimmia e essere umano. In una forma di realizzazione l'ospite mammifero è un topo. In un'ulteriore forma di realizzazione l'ospite mammifero è un essere umano.

Se un polipeptide suscita una risposta immunitaria protettiva in un ospite mammifero contro ceppi di *C. difficile* può essere determinato usando un saggio di provocazione. In un tale saggio l'ospite mammifero è vaccinato con il polipeptide e provocato mediante esposizione a *C.difficile*, il tempo che il mammifero sopravvive dopo provocazione è confrontato con il tempo che sopravvive un mammifero di riferimento che non è stato immunizzato con il polipeptide. Un polipeptide suscita una risposta immunitaria protettiva se un mammifero immunizzato con il polipeptide sopravvive almeno 10%, 20%, 30%, 50%, 80%, 80%, 90%, o 100% più a lungo di un mammifero di riferimento che non è stato immunizzato dopo provocazione con *C.difficile*.

La struttura nativa dei domini C-terminali da tossine A e B è costituita da una estesa struttura tipo  $\beta$  solenoide. Questa struttura è costituita principalmente da strutture a foglietto  $\beta$ , con una minoranza di strutture a  $\alpha$  elica come osservata in Ho et al (PNAS 102:18373-18378 (2005)). Le strutture secondarie presenti possono essere determinate usando dicroismo circolare. Per esempio misurando la forma e l'entità degli spettri CD nella regione del lontano-UV (190-250nm) e confrontando i risultati con quelli di strutture note. Questo può essere effettuato usando un cammino ottico di 0.01cm da 178 a 250nm, con una risoluzione di 1nm e un'ampiezza di banda su uno spettropolarimetro Jasco J-720, per esempio come osservato nell'esempio 5 sotto.

In una forma di realizzazione il primo frammento comprende meno di 25%, 23%, 20%, 18%, 15%, 10%, o 7% di struttura secondaria a alfa elica. In una forma di realizzazione il secondo frammento comprende meno di 28%, 25%, 23%, 20%, 18%, 15%, 10%, o 7% di struttura secondaria a alfa elica. In un'ulteriore forma di realizzazione ambedue il primo frammento e il secondo frammento comprendono meno di 28%, 25%, 23%, 20%, 18%, 15%, 10%, o 7% di struttura secondaria a alfa elica.

In una forma di realizzazione il primo frammento comprende più di 20%, 25%, 28%, 30%, 33%, 35%, 38%, 40%, o 42% di struttura a foglietto beta. In una forma di realizzazione il secondo frammento comprende più di

20%, 25%, 28%, 30%, 33%, 35%, 38%, 40%, o 42% di struttura a foglietto beta. In un'ulteriore forma di realizzazione sia il primo frammento sia il secondo frammento comprende più di 20%, 25%, 28%, 30%, 33%, 35%, 38%, 40%, o 42% di struttura a foglietto beta.

La Figura 2 rappresenta l'organizzazione dei domini C-terminali di ToxA e ToxB. Il dominio C-terminale di tossina A è costituito da 8 porzioni di ripetizioni (indicate come porzione di ripetizioni I, porzione di ripetizioni II, porzione di ripetizioni III, porzione di ripetizioni IV, porzione di ripetizioni V, porzione di ripetizioni VI, porzione di ripetizioni VII e porzione di ripetizioni VIII) ognuna di queste porzioni di ripetizioni può essere ulteriormente divisa in ripetizioni corte (SR) che sono rappresentate come box bianche nella figura 2 e ripetizioni lunghe (LR) che sono rappresentate come box nere nella figura 2 (eccetto per la porzione Tox A di ripetizione VIII che non ha una ripetizione lunga). Ciascuna delle ripetizioni lunghe ha qualche similarità strutturale e di sequenza con le altre ripetizioni lunghe. Analogamente le ripetizioni corte hanno qualche similarità di sequenza e strutturale l'una con l'altra. Il dominio C-terminale di tossina B è costituito da 5 porzioni di ripetizioni suddivise in SR e LR. Ciascuna porzione di ripetizioni contiene una LR e tra 2 e 5 SR (eccetto per porzione di ripetizioni V di ToxB che non ha una ripetizione lunga). Ai fini della descrizione la frase "una porzione di ripetizioni" si riferisce a una delle otto porzioni di ripetizioni di ToxA (indicate come I, II, III, IV, V, VI, VII, e VIII) oppure una delle cinque porzioni di ripetizioni di ToxB (indicate come I, II, III, IV oppure VI). Come usato nella presente il termine "prima porzione di ripetizioni" si riferisce a una porzione di ripetizioni (oppure una porzione di ripetizioni parziale) dal dominio ripetitivo di tossina A. Il termine "seconda porzione di ripetizioni" si riferisce a una porzione di ripetizioni (oppure una porzione di ripetizioni parziale) dal dominio ripetitivo di tossina B. Per gli scopi della descrizione il termine "ripetizione lunga" si riferisce a uno dei domini LR rappresentati come box nere nella Figura 2. Ai fini della descrizione il termine "ripetizione corta" si riferisce a uno dei domini SR rappresentati come box bianche nella Figura 2.

Pertanto per esempio, la porzione di ripetizioni I di ToxA contiene tre SR e una LR, che possono essere riferite come la prima SRI di ToxA, la seconda SRI di ToxA, la terza SRI di ToxA e la LRI di ToxA, rispettivamente.

La prima estremità prossimale si ritiene che sia all'interno di una "porzione di ripetizioni" se il primo

frammento termina in un amminoacido che è all'interno di quella porzione di ripetizioni (cioè, la prima estremità prossimale contiene soltanto parte della sequenza della porzione di ripetizioni). Analogamente la seconda estremità prossimale si ritiene che sia all'interno di una "porzione di ripetizioni" se il secondo frammento termina in un amminoacido che è all'interno di quella porzione di ripetizioni. Per esempio la prima estremità prossimale è all'interno di "una porzione di ripetizioni I di ToxA" se il primo frammento termina con uno qualunque di amminoacidi 1832-1924 (incluso) di VPI10463 oppure loro equivalenti in un altro ceppo. La prima estremità prossimale non è all'interno di una porzione ripetizione corta - ripetizione lunga-ripetizione corta se il primo frammento termina con un amminoacido che non è all'interno di quella porzione ripetizione corta-ripetizione lunga-ripetizione corta.

Le posizioni amminoacidiche di ogni dominio sono state definite per tossina A e tossina B da ceppo VPI10463 (ATCC43255). Queste sono come segue

Nome		posizione iniziale	posizione finale
ToxA_I	SR1	1832	1852
	SR2	1853	1873
	SR3	1874	1893
	LR	1894	1924
ToxA_II	SR1	1925	1944
	SR2	1945	1965
	SR3	1966	1986
	SR4	1987	2007
	SR5	2008	2027
	LR	2028	2058
ToxA_III	SR1	2059	2078

	SR2	2079	2099
	SR3	2100	2120
	SR4	2121	2141
	SR5	2142	2161
	LR	2162	2192
ToxA_IV	SR1	2193	2212
	SR2	2213	2233
	SR3	2234	2253
	SR4	2254	2275
	LR	2276	2306
ToxA_V	SR1	2307	2326
	SR2	2327	2347
	SR3	2348	2368
	SR4	2369	2389
	SR5	2390	2409
	LR	2410	2440
ToxA_VI	SR1	2441	2460
	SR2	2461	2481
	SR3	2482	2502
	SR4	2503	2522
	LR	2523	2553
ToxA_VII	SR1	2554	2573
	SR2	2574	2594
	SR3	2595	2613
	LR	2614	2644
ToxA_VIII	SR1	2645	2664
	SR2	2665	2686

	SR3	2687	2710
ToxB_I	SR1	1834	1854
	SR2	1855	1876
	SR3	1877	1896
	LR	1897	1926
ToxB_II	SR1	1927	1946
	SR2	1947	1967
	SR3	1968	1987
	SR4	1988	2007
	SR5	2008	2027
	LR	2028	2057
ToxB_III	SR1	2058	2078
	SR2	2079	2099
	SR3	2100	2119
	SR4	2120	2139
	SR5	2140	2159
	LR	2160	2189
ToxB_IV	SR1	2190	2212
	SR2	2213	2233
	SR3	2234	2253
	SR4	2254	2273
	SR5	2274	2293
	LR	2294	2323
ToxB_V	SR1	2324	2343
	SR2	2344	2366

Tabella 1

Per questa ragione il termine “porzione di ripetizioni” può riferirsi a amminoacidi 1832-1924, 1925-2058, 2059-2192, 2193-2306, 2307-2440, 2441-2553, 2554-2644 oppure 2645-2710 di tossina A (SEQ ID

NO:1), oppure amminoacidi 1834-1926, 1927-2057, 2058-2189, 2190-2323 oppure 2324-2366 di tossina B (SEQ ID NO:2) oppure loro equivalenti in un ceppo differente di *C.difficile*.

Per questa ragione il termine 'ripetizione corta' può riferirsi a amminoacidi 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 oppure 2687-2710 di tossina A (SEQ ID NO:1) oppure amminoacidi 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 oppure 2344-2366 di tossina B (SEQ ID NO:2) oppure loro equivalenti in un ceppo differente di *C.difficile*.

Analogamente il termine “ripetizione lunga” può riferirsi a amminoacidi 1894-1924, 2028-2058, 2162-2192, 2276-2306, 2410-2440, 2523-2553 oppure 2614-2644 di tossina A (SEQ ID NO:1) oppure amminoacidi 1897-1926, 2028-2057, 2160-2189 oppure 2294-2323 di tossina B (SEQ ID NO:2) oppure loro equivalenti in un ceppo differente di *C.difficile*.

Analogamente il termine “porzione ripetizione corta-ripetizione lunga-ripetizione corta” può riferirsi a amminoacidi 1874-1944, 2008-2078, 2142-2212, 2254-2326, 2390-2460, 2503-2573, oppure 2595-2664 di tossina A (SEQ ID NO:1) oppure amminoacidi 1877-1946, 2008-2078, 2140-2212 oppure 2274-2343 di tossina B (SEQ ID NO:2) oppure loro equivalenti in un ceppo differente di *C.difficile*. Il termine “non interrompe una porzione ripetizione corta-ripetizione lunga-ripetizione corta” significa che l'estremità prossimale è in una regione che non interrompe la struttura della porzione ripetizione corta-ripetizione lunga-ripetizione corta, in generale questo significa che l'estremità prossimale non è all'interno di una ripetizione lunga e non all'interno delle ripetizioni corte che costituiscono una porzione ripetizione corta-ripetizione lunga-ripetizione corta, eccetto che l'estremità prossimale può essere nella regione di 1, 2, 3, 4, 5 o 6 amminoacidi della ripetizione corta che sono più distanti dalla ripetizione lunga nella sequenza. In una forma di realizzazione il termine “non interrompe una porzione ripetizione corta-ripetizione lunga-ripetizione corta” significa che l'estremità prossimale non è

all'interno della porzione ripetizione corta-ripetizione lunga-ripetizione corta.

In una forma di realizzazione la prima estremità prossimale non è all'interno di amminoacidi 1878-1940, 2146-2208, 2012-2074, 2258-2322, 2394-2456, 2507-2569, 2599-2660 oppure 2593-2660 di tossina A (SEQ ID NO:1) oppure loro equivalenti in un ceppo differente di *C.difficile*. In una seconda forma di realizzazione la seconda estremità prossimale non è all'interno di amminoacidi 1881-1942, 2012-2074, 2144-2208, oppure 2278-2339 di tossina B (SEQ ID NO:2) oppure loro equivalenti in un ceppo differente di *C.difficile*. In un'ulteriore forma di realizzazione la prima estremità prossimale non è all'interno di amminoacidi 1878-1940, 2146-2208, 2012-2074, 2258-2322, 2394-2456, 2507-2569, 2599-2660 oppure 2593-2660 di tossina A (SEQ ID NO:1) oppure loro equivalenti in un ceppo differente di *C. difficile* e la seconda estremità prossimale non è all'interno di amminoacidi 1881-1942, 2012-2074, 2144-2208, oppure 2278-2339 di tossina B (SEQ ID NO:2) oppure loro equivalenti in un ceppo differente di *C.difficile*.

In una forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni V (amminoacidi 2307-2440 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente), VI (amminoacidi 2441-2553 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente), VII (amminoacidi 2554-2644 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) oppure VIII (amminoacidi 2645-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) di tossina A. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni VII (amminoacidi 2554-2644 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) di tossina A. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni VIII (amminoacidi 2645-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) di tossina A.

In una forma di realizzazione la seconda estremità prossimale è all'interno di porzione di ripetizioni I (amminoacidi 1834-1926 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente), II (amminoacidi 1927-2057 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente), oppure iii (amminoacidi 2058-2189 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la seconda estremità prossimale è all'interno di porzione di ripetizioni II (amminoacidi 1927-2057

di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la seconda estremità prossimale è all'interno di porzione di ripetizioni I (amminoacidi 1834-1926 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B.

In una forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni VIII (amminoacidi 2645-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) di tossina A e la seconda estremità prossimale è all'interno di porzione di ripetizioni I (amminoacidi 1834-1926 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni VIII (amminoacidi 2645-2710 di SEQ ID NO:1), oppure loro equivalenti in un ceppo differente) di tossina A e la seconda estremità prossimale è all'interno di porzione di ripetizioni II (amminoacidi 1927-2057 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni VII (amminoacidi 2554-2644 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) di tossina A e la seconda estremità prossimale è all'interno di porzione di ripetizioni I (amminoacidi 1834-1926 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni VII (amminoacidi 2554-2644 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) di tossina A e la seconda estremità prossimale è all'interno di porzione di ripetizioni II (amminoacidi 1927-2057 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni VI (amminoacidi 2441-2553 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di porzione di ripetizioni I (amminoacidi 1834-1926 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni VI (amminoacidi 2441-2553 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di porzione di ripetizioni II (amminoacidi 1927-2057 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di

ripetizioni V (amminoacidi 2307-2440 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di porzione di ripetizioni I (amminoacidi 1834-1926 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di porzione di ripetizioni V (amminoacidi 2307-2440 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di porzione di ripetizioni II (amminoacidi 1927-2057 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente) di tossina B.

In una forma di realizzazione la prima estremità prossimale è all'interno di amminoacidi 2690-2710, oppure 2695-2710, oppure 2700-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente. In un'ulteriore forma di realizzazione la prima estremità prossimale è all'interno di amminoacidi 2670-2700, oppure 2675-2695, oppure 2680-2690 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente. In una forma di realizzazione la seconda estremità prossimale è all'interno di amminoacidi 1860-1878 di Tossina B oppure loro equivalenti in un ceppo differente. In una forma di realizzazione la seconda estremità prossimale è all'interno di amminoacidi 1950-1980, 1955-1975 oppure 1960-1970 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente. In un'ulteriore forma di realizzazione la seconda estremità prossimale è all'interno di amminoacidi 1978-2008, 1983-2003 oppure 1988-1998 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente. In un'ulteriore forma di realizzazione la seconda estremità prossimale è all'interno di amminoacidi 1860-1878, 1854-1876, 1857-1887, 1862-1882, oppure 1867-1877 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente.

In una forma di realizzazione il primo frammento è costituito da un dominio ripetitivo di tossina A intero (amminoacidi 1832-2710). In una forma di realizzazione il secondo frammento è costituito da un dominio ripetitivo di tossina B intero (amminoacidi 1833-2366).

In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2687-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 4 di porzione di ripetizioni II di tossina B (amminoacidi 1988-2007 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una

forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2687-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni II di tossina B (amminoacidi 1968-1987 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2687-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni II di tossina B (amminoacidi 1947-1967 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2687-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni I di tossina B (amminoacidi 1877-1896 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2687-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni I di tossina B (amminoacidi 1855-1876 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2687-2710 di SEQ ID NO:1), oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 1 di porzione di ripetizioni I di tossina B (amminoacidi 1834-1854 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VIII di tossina A (amminoacidi 2665-2686 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 4 di porzione di ripetizioni II di tossina B (amminoacidi 1988-2007 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VIII di tossina A (amminoacidi 2665-

2686 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni II di tossina B (amminoacidi 1968-1987 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VIII di tossina A (amminoacidi 2665-2686 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni II di tossina B (amminoacidi 1947-1967 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VIII di tossina A (amminoacidi 2665-2686 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni I di Tossina B (amminoacidi 1877-1896 di tossina B oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VIII di tossina A (amminoacidi 2665-2686 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni I di tossina B (amminoacidi 1855-1876 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VIII di tossina A (amminoacidi 2665-2686 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 1 di porzione di ripetizioni I di tossina B (amminoacidi 1834-1854 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni vii di tossina A (amminoacidi 2574-2594 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 4 di porzione di ripetizioni II di tossina B (amminoacidi 1988-2007 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni vii di tossina A (amminoacidi 2574-2594 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni II

di tossina B (amminoacidi 1668-1987 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni vii di tossina A (amminoacidi 2574-2594 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni II di tossina B (amminoacidi 1947-1967 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VII di tossina A (amminoacidi 2574-2594 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni I di tossina B (amminoacidi 1855-1876 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VII di tossina A (amminoacidi 2574-2594 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 1 di porzione di ripetizioni I di tossina B (amminoacidi 1834-1854 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VI di tossina A (amminoacidi 2482-2502 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 4 di porzione di ripetizioni II di tossina B (amminoacidi 1988-2007 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VI di tossina A (amminoacidi 2482-2502 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni II di tossina B (amminoacidi 1968-1987 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VI di tossina A (amminoacidi 2482-2502 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni II di tossina B (amminoacidi 1947-1967 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di

porzione di ripetizioni VI di tossina A (amminoacidi 2482-2502 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni I di tossina B (amminoacidi 1855-1876 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VI di tossina A (amminoacidi 2482-2502 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 1 di porzione di ripetizioni I di tossina B (amminoacidi 1834-1854 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VI di tossina A (amminoacidi 2461-2481 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 4 di porzione di ripetizioni II di tossina B (amminoacidi 1988-2007 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VI di tossina A (amminoacidi 2461-2481 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni II di tossina B (amminoacidi 1968-1987 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VI di tossina A (amminoacidi 2461-2481 di SEQ ID NO:1), oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni II di tossina B (amminoacidi 1947-1967 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VI di tossina A (amminoacidi 2461-2481 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni I di tossina B (amminoacidi 1855-1876 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente). In una forma di realizzazione la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VI di tossina A (amminoacidi 2461-2481 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente) e la seconda estremità prossimale è all'interno di ripetizione corta 1 di porzione di

ripetizioni I di tossina B (amminoacidi 1834-1854 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente).

In una forma di realizzazione la prima estremità prossimale è all'interno di amminoacidi 2690-2710, oppure 2695-2710, oppure 2700-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente e la seconda estremità prossimale è all'interno di amminoacidi 1950-1980, 1955-1975 oppure 1960-1970 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente. In una forma di realizzazione la prima estremità prossimale è all'interno di amminoacidi 2690-2710, oppure 2695-2710, oppure 2700-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente e la seconda estremità prossimale è all'interno di amminoacidi 1978-2008, 1983-2003 oppure 1988-1998 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente. In una forma di realizzazione la prima estremità prossimale è all'interno di amminoacidi 2690-2710, oppure 2695-2710, oppure 2700-2710 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente e la seconda estremità prossimale è all'interno di amminoacidi 1857-1887, 1862-1882, oppure 1867-1877 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente. In una forma di realizzazione la prima estremità prossimale è all'interno di amminoacidi 2670-2700, oppure 2675-2695, oppure 2680-2690 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente e la seconda estremità prossimale è all'interno di amminoacidi 1950-1980, 1955-1975 oppure 1960-1970 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente. In una forma di realizzazione la prima estremità prossimale è all'interno di amminoacidi 2670-2700, oppure 2675-2695, oppure 2680-2690 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente e la seconda estremità prossimale è all'interno di amminoacidi 1978-2008, 1983-2003 oppure 1988-1998 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente. In una forma di realizzazione la prima estremità prossimale è all'interno di amminoacidi 2670-2700, oppure 2675-2695, oppure 2680-2690 di SEQ ID NO:1 oppure loro equivalenti in un ceppo differente e la seconda estremità prossimale è all'interno di amminoacidi 1857-1887, 1862-1882, 1860-1878 oppure 1867-1877 di SEQ ID NO:2 oppure loro equivalenti in un ceppo differente.

In una forma di realizzazione il primo frammento comprende almeno 100, 200, 300, 400 o 450 amminoacidi. In una forma di realizzazione il secondo frammento comprende almeno 100, 200, 300 o 400 amminoacidi.

In una forma di realizzazione il polipeptide comprende inoltre un linker. Questo linker può essere tra la prima estremità prossimale e la seconda estremità prossimale, in alternativa il linker può legare le estremità distali del primo frammento e/o del secondo frammento a una ulteriore sequenza di amminoacidi.

Una sequenza di linker peptidico può essere impiegata per separare il primo frammento e il secondo frammento.

Una tale sequenza di linker peptidico è incorporata nella proteina di fusione usando tecniche standard ben note nella tecnica. Sequenze di linker peptidico adatte possono essere scelte sulla base dei seguenti fattori: (1) loro capacità di adottare una conformazione estesa flessibile; (2) loro incapacità di adottare una struttura secondaria che potrebbe interagire con epitopi funzionali sul primo frammento e/o sui secondi frammenti; e (3) la mancanza di residui idrofobi oppure carichi che potrebbero reagire con gli epitopi funzionali di Tox A e/o ToxB. Sequenze di linker peptidico possono contenere residui Gly, Asn e Ser. Altri amminoacidi quasi neutri, quali Thr e Ala possono anche essere usati nella sequenza linker. Sequenze amminoacidiche che possono essere utilmente impiegate come linker includono quelle descritte in Maratea et al., *Gene* 40:39-46 (1985); Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262 (1986); Brevetto U.S. No. 4,935,233 e Brevetto U.S. No. 4,751,180.

In una forma di realizzazione il linker comprende tra 1-19, 1-15, 1-10, 1-5, 1-2, 5-20, 5-15, 5-15, 10-20, oppure 10-15 amminoacidi. In una forma di realizzazione il linker è un linker di glicine, il linker può comprendere (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 18 oppure 19) residui di glicina contigui multipli, oppure in alternativa il linker può comprendere alcuni residui di glicina e alcuni residui di altri amminoacidi quale alanina. In un'ulteriore forma di realizzazione il linker comprende un singolo residuo di glicina.

In una forma di realizzazione il polipeptide dell'invenzione è parte di una proteina di fusione più grande. Le proteine di fusione possono inoltre comprendere amminoacidi codificanti per una porzione immunogena di un ulteriore antigene proteico. Per esempio la proteina di fusione può inoltre comprendere una porzione immunogena di un antigene proteico ottenuto oppure derivato da un batterio selezionato dal gruppo che è composto da *S.pneumoniae*, *H.influenzae*, *N.meningitidis*, *E.coli*, *M.cattarhalis*, *C.tentani*, *C.diphtheriae*, *B.pertussis*, *S.epidermidis*, enterococchi, *S.aureus*, e *Pseudomonas aeruginosa*. In questo caso il linker può essere tra il primo frammento oppure il secondo frammento e una ulteriore porzione immunogena di un antigene

proteico.

Il termine "porzione immunogena dello stesso" oppure "frammento immunogeno" si riferisce a un frammento di un polipeptide in cui il frammento comprende un epitopo che è riconosciuto da linfociti T citotossici, linfociti T helper oppure cellule B. Appropriatamente, la porzione immunogena comprenderà almeno 30%, appropriatamente almeno 50%, specialmente almeno 75% e in particolare almeno 90% (ad esempio 95% o 98%) degli amminoacidi nella sequenza di riferimento. La porzione immunogena comprenderà appropriatamente tutte delle regioni dell'epitopo della sequenza di riferimento.

In una forma di realizzazione il polipeptide comprende un frammento immunogeno di SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25 oppure SEQ ID NO:27. In una forma di realizzazione il polipeptide comprende un frammento immunogeno di almeno 500, 550, 600, 650, 700, 750, 780, 800, 830, 850, 880, 900, 920, o 950 amminoacidi di SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25 oppure SEQ ID NO:27. In un'ulteriore forma di realizzazione il polipeptide comprende una variante di SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25 oppure SEQ ID NO:27, in un'ulteriore forma di realizzazione il polipeptide comprende una variante avente almeno 80%, 85%, 90%, 92%, 95%, 98%, 99%, o 100% di identità di sequenza con SEQ ID NO:3-SEQ ID NO:7.

In una forma di realizzazione il polipeptide comprende più di 450, 475, 500, 525, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825 o 850 amminoacidi da tossina A. In una forma di realizzazione il polipeptide comprende meno di 850, 825, 800, 775, 750, 725, 700, 675, 650, 625, o 600 amminoacidi da tossina A. In una forma di realizzazione il polipeptide comprende più di 350, 375, 400, 425, 450, 475, 500 o 525 amminoacidi da tossina B. In una forma di realizzazione il polipeptide comprende meno di 525, 500, 475, o 450 amminoacidi da tossina B.

Il termine "identità" è noto nella tecnica, è una relazione tra due o più sequenze polipeptidiche oppure due o più sequenze polinucleotidiche, come è il caso, come determinata confrontando le sequenze. Nella tecnica, "identità" significa anche il grado di relazionalità tra sequenze tra sequenze polipeptidiche oppure polinucleotidiche, come

è il caso, come determinato dalla corrispondenza tra stringhe di tali sequenze. L'"identità" può essere facilmente calcolata mediante metodi noti, includenti ma non limitati a quelli descritti in (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heine, G., Academic Press, 1987; e Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; e Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Metodi di determinazione di identità sono progettati per dare la più grande corrispondenza tra le sequenze testate. Inoltre, metodi per determinare l'identità sono codificati in programmi informatici disponibili pubblicamente. Metodi basati su programmi informatici per determinare identità tra le due sequenze includono, ma non vi si limitano, il programma di Needle BLASTP, BLASTN (Altschul, S.F. et al., J. Molec. Biol. 215: 403-410 (1990), e FASTA (Pearson and Lipman Proc. Natl. Acad. Sci. USA 85; 2444-2448 (1988). La famiglia di programmi BLAST è pubblicamente disponibile da NCBI e altre fonti (Manuale BLAST, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215: 403-410 (1990). L'algoritmo di Smith Waterman ben noto può anche essere usato per determinare l'identità.

Parametri per confronto tra sequenze di polipeptidi includono i seguenti:

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

Matrice di confronto: BLOSSUM62 da Henikoff and Henikoff,

Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992)

Penalità per gap: 10

Penalità per estensione di gap:0.5

Un programma utile con questi parametri è pubblicamente disponibile come il programma 'needle' dal pacchetto EMBOSS (Rice P. et al, Trends in Genetics 2000 col.16(6):276-277). I suddetti parametri sono i parametri di default per confronti tra peptidi (con nessuna penalità per gap finali).

Al fine di determinare l'identità di una sequenza di riferimento con SEQ ID NO:1, in una forma di realizzazione

l'identità di sequenza è calcolata sull'intera lunghezza della sequenza di riferimento. In un'ulteriore forma di realizzazione l'identità di sequenza è calcolata sull'intera lunghezza della sequenza in SEQ ID NO:1. Al fine di determinare l'identità di una sequenza di riferimento con SEQ ID NO:2, in una forma di realizzazione l'identità di sequenza è calcolata sull'intera lunghezza della sequenza di riferimento. In un'ulteriore forma di realizzazione l'identità di sequenza è calcolata sull'intera lunghezza della sequenza in SEQ ID NO:2.

Anche descritti nella presente sono un polipeptide comprendente (i) SEQ IDNO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33 oppure SEQ ID NO:34 oppure SEQ ID NO:35, (ii) una variante avente almeno 80%, 85%, 88%, 90%, 92%, 95%, 98%, 99% o 100% di identità con le SEQ ID NO:10-19 oppure (iii) un frammento di almeno 100, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 o 600 amminoacidi di SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33 o SEQ ID NO:34 o SEQ ID NO:35. In un'ulteriore forma di realizzazione il polipeptide comprende SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33 oppure SEQ ID NO:34 oppure SEQ ID NO:35, ii) una variante avente almeno 80%, 85%, 88%, 90%, 92%, 95%, 98%, 99% o 100% di identità con SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33 oppure SEQ ID NO:34 oppure SEQ ID NO:35; oppure (iii) un frammento di almeno 100, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 o 600 amminoacidi di SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33 o SEQ ID NO:34 oppure SEQ ID NO:35. In un'ulteriore forma di realizzazione il

polipeptide comprende SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33 oppure SEQ ID NO:34 oppure SEQ ID NO:35, ii) una variante avente almeno 80%, 85%, 88%, 90%, 92%, 95%, 98%, 99% o 100% di identità con SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33 oppure SEQ ID NO:34 oppure SEQ ID NO:35; oppure (iii) un frammento di almeno 100, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 o 600 amminoacidi di SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33 oppure SEQ ID NO:34 oppure SEQ ID NO:35.

In una forma di realizzazione il polipeptide comprende più di 450, 475, 500, 525, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825 o 850 amminoacidi da tossina A. In una forma di realizzazione il polipeptide comprende meno di 850, 825, 800, 775, 750, 725, 700, 675, 650, 625, o 600 amminoacidi da tossina A. In una forma di realizzazione il polipeptide comprende più di 350, 375, 400, 425, 450, 475, 500 o 525 amminoacidi da tossina B. In una forma di realizzazione il polipeptide comprende meno di 525, 500, 475, o 450 amminoacidi da tossina B. In un'ulteriore forma di realizzazione il polipeptide suscita anticorpi neutralizzanti che neutralizzano tossina A oppure tossina B oppure ambedue. In un'ulteriore forma di realizzazione il polipeptide suscita anticorpi che neutralizzano tossina A. In un'ulteriore forma di realizzazione il polipeptide suscita anticorpi che neutralizzano tossina B. In un'ulteriore forma di realizzazione il polipeptide suscita anticorpi che neutralizzano tossina A e tossina B. Il polipeptide dell'invenzione suscita anticorpi che neutralizzano tossina A se i sieri contro il polipeptide danno una readout ELISA maggiore di 10%, 20%, 30%, 50%, 70%, 80%, 90% o 100% più alta del campione di riferimento.

In un'ulteriore forma di realizzazione il polipeptide dell'invenzione suscita una risposta immunitaria protettiva in un mammifero ospite contro ceppi di *C.difficile*. In una forma di realizzazione l'ospite mammifero è selezionato

dal gruppo composto da topo, coniglio, porcellino d'India, scimmia, primate non umano e essere umano. In una forma di realizzazione l'ospite mammifero è un topo. In un'ulteriore forma di realizzazione l'ospite mammifero è un essere umano.

Se un polipeptide suscita una risposta immunitaria protettiva in un ospite mammifero contro ceppi di *C. difficile* può essere determinato usando un saggio di provocazione. In un tale saggio l'ospite mammifero è vaccinato con il polipeptide e provocato mediante esposizione a *C. difficile*, il tempo che il mammifero sopravvive dopo provocazione è confrontato con il tempo che un mammifero di riferimento che non è stato immunizzato con il polipeptide sopravvive. Un polipeptide suscita una risposta immunitaria protettiva se un mammifero immunizzato con il polipeptide sopravvive almeno 10%, 20%, 30%, 50%, 70%, 80%, 90%, o 100% più a lungo di un mammifero di riferimento che non è stato immunizzato dopo provocazione con *C. difficile*. In una forma di realizzazione il polipeptide dell'invenzione suscita una risposta immunitaria protettiva contro ceppi di *C. difficile* in un mammifero selezionato dal gruppo composto da topo, porcellino d'India, scimmia e essere umano. In una forma di realizzazione il mammifero è un topo, in un'ulteriore forma di realizzazione il mammifero è un essere umano.

La struttura nativa dei domini (ripetitivi) C-terminali da tossine A e B è costituita da una struttura tipo  $\beta$  solenoide estesa. Questa struttura è costituita principalmente da strutture a foglietto  $\beta$ , con una minoranza di strutture a  $\alpha$  elica come osservato in Ho et al (PNAS 102:18373-18378 (2005)). Le strutture secondarie presenti possono essere determinate usando dicroismo circolare. Per esempio misurando la forma e l'entità degli spettri CD nella regione del lontano-UV (190-250nm) e confrontando i risultati con quelli di strutture note. Questo può essere effettuato usando un cammino ottico di 0.01cm da 178 a 250nm, con una risoluzione di 1nm e una ampiezza di banda su uno spettropolarimetro Jasco J-720, per esempio come osservato nell'esempio 5 sotto.

In una forma di realizzazione il polipeptide comprende meno di 25%, 23%, 20%, 28%, 15%, 10%, o 7% di struttura secondaria a alfa elica. In un'ulteriore forma di realizzazione il polipeptide comprende più di 20%, 25%, 28%, 30%, 33%, 35%, 38%, 40%, o 42% di struttura a foglietto beta.

#### POLINUCLEOTIDI

L'invenzione fornisce inoltre un polinucleotide codificante per un polipeptide dell'invenzione. Ai fini dell'invenzione il termine "polinucleotide/i" generalmente si riferisce a qualunque poliribonucleotide oppure polideossiribonucleotide, che può essere RNA oppure DNA non modificato oppure RNA oppure DNA modificato includente regioni/forme a singolo e a doppio filamento.

Il termine "polinucleotide codificante per un peptide" come usato nella presente comprende polinucleotidi che includono una sequenza codificante per un peptide oppure un polipeptide dell'invenzione. Il termine comprende anche polinucleotidi che includono una singola regione continua oppure regioni discontinue codificanti per il peptide oppure per il polipeptide (per esempio, polinucleotidi interrotti da fago integrato, una sequenza di inserzione integrata, una sequenza di vettore integrata, una sequenza di trasposone integrata, oppure per via di un editing di RNA oppure una riorganizzazione di DNA genomico) insieme con regioni aggiuntive, che possono contenere anche sequenze codificanti e/o non codificanti.

Ci si renderà conto da quelli di ordinaria esperienza nella tecnica che, come risultato della degenerazione del codice genetico, vi sono molte sequenze nucleotidiche che codificano per un polipeptide come descritto nella presente. Alcuni di questi polinucleotidi portano minima similarità con la sequenza nucleotidica di qualunque gene nativo (cioè naturalmente esistente). Ciò nonostante, polinucleotidi che variano per via di differenze nel codon usage sono specificamente contemplate dalla presente invenzione, per esempio polinucleotidi che sono ottimizzati per selezione di codoni umani e/o di primate e/o di *E.coli*.

Sequenze codificanti per un desiderato polipeptide possono essere sintetizzate, per intero oppure in parte, usando metodi chimici ben noti nella tecnica (si veda Caruthers, M. H. *et al.*, *Nucl. Acids Res. Symp. Ser.* pp. 215-223 (1980), Horn *et al.*, *Nucl. Acids Res. Symp. Ser.* pp. 225-232 (1980)). In alternativa, la proteina stessa può essere prodotta usando metodi chimici per sintetizzare la sequenza amminoacidica di un polipeptide, oppure una porzione della stessa. Per esempio, la sintesi di peptidi può essere compiuta usando varie tecniche su fase solida (Roberge *et al.*, *Science* 269:202-204 (1995)) e una sintesi automatizzata può essere ottenuta, per esempio, usando il Sintetizzatore di Peptidi ASI 431 A (Perkin Elmer, Palo Alto, CA).

Inoltre, le sequenze polinucleotidiche della presente invenzione possono essere ingegnerizzate usando metodi

comunemente noti nella tecnica allo scopo di alterare sequenze codificanti per polipeptidi per una varietà di ragioni, includenti ma non limitate a, alterazioni che modificano il clonaggio, il processamento, e/o l'espressione del prodotto del gene. Per esempio, si può usare rimescolamento del DNA mediante frammentazione random e riassetto per PCR di frammenti genici e oligonucleotidi sintetici per ingegnerizzare le sequenze nucleotidiche. Inoltre, la mutagenesi sito-diretta può essere usata per inserire nuovi siti di restrizione, alterare patterns di glicosilazione, cambiare preferenza dei codoni, produrre varianti di splicing, oppure introdurre mutazioni, e così via.

### VETTORI

In un ulteriore aspetto dell'invenzione la presente invenzione riguarda un vettore comprendente un polinucleotide dell'invenzione legato a un promotore inducibile in modo tale che quando il promotore è indotto un polipeptide codificato dal polinucleotide sia espresso.

Un ulteriore aspetto dell'invenzione comprende detto vettore in cui il promotore inducibile è attivato mediante aggiunta di una quantità sufficiente di IPTG (Isopropil  $\beta$ -D-1-tiogalattopiranoside) preferibilmente al terreno di crescita. Facoltativamente questo è ad una concentrazione compresa tra 0.1 e 10mM, 0.1 e 5mM, 0.1 e 2.5mM, 0.2 e 10mM, 0.2 e 5mM, 0.2 e 2.5mM, 0.4 e 10mM, 1 e 10mM, 1 e 5mM, 2.5 e 10mM, 2.5 e 5mM, 5 e 10mM. In alternativa il promotore può essere indotto mediante una variazione in temperatura oppure in pH.

### CELLULE OSPITI

Per la produzione ricombinante dei polipeptidi dell'invenzione, le cellule ospiti possono essere geneticamente ingegnerizzate per incorporare sistemi di espressione oppure porzioni degli stessi oppure polinucleotidi dell'invenzione. L'introduzione di un polinucleotide nella cellula ospite può essere effettuata mediante metodi descritti in molti manuali di laboratorio standard, quali Davis, et al., BASIC METHODS IN MOLECULAR BIOLOGY, (1986) and Sambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), quali, trasfezione con fosfato di calcio, trasfezione mediata da DEAE-destrano, transvection, microiniezione, trasfezione mediata da lipidi cationici, elettroporazione, coniugazione, trasduzione, caricamento per raschiamento, introduzione balistica e infezione.

Esempi rappresentativi di ospiti appropriati includono cellule batteriche gram-negative, quali cellule di, E. coli, Acinetobacter, Actinobacillus, Bordetella, Brucella, Campylobacter, Cyanobacteria, Enterobacter, Erwinia, Francisella, Helicobacter, hemophilus, Klebsiella, Legionella, Moraxella, Neisseria, Pasteurella, Proteus, Pseudomonas, Salmonella, Serratia, Shigella, Treponema, Vibrio, Yersinia. In una forma di realizzazione la cellula ospite è una cellula di Escherichia coli. In alternativa cellule batteriche gram-positive possono anche essere usate. Una grande varietà di sistemi di espressione possono essere usati per produrre i polipeptidi dell'invenzione. In una forma di realizzazione il vettore è derivato da plasmidi batterici. Generalmente qualunque sistema oppure vettore adatto per mantenere, propagare oppure esprimere polinucleotidi e/o esprimere un polipeptide in un ospite può essere usato per espressione a questo proposito. La sequenza di DNA appropriata può essere inserita nel sistema di espressione mediante qualunque di una varietà di tecniche ben note e di routine, quali, per esempio, quelle espone in Sambrook et al., MOLECULAR CLONING, A LABORATORY MANUAL, (supra).

#### COMPOSIZIONI IMMUNOGENE E VACCINI

C'è inoltre fornita una composizione immunogena comprendente un polipeptide dell'invenzione e un eccipiente farmaceuticamente accettabile.

In una forma di realizzazione la composizione immunogena comprende inoltre un adiuvante. La scelta di un adiuvante adatto da miscelare con tossine batteriche oppure coniugati fatti usando i procedimenti dell'invenzione è nella conoscenza della persona esperta nella tecnica. Adjuvanti adatti includono un sale di alluminio quale gel di idrossido di alluminio oppure fosfato di alluminio oppure allume, ma possono anche essere altri sali di metallo quali quelli di calcio, magnesio, ferro oppure zinco, oppure possono essere una sospensione insolubile di tirosina acilata, oppure zuccheri acilati, saccaridi cationicamente oppure anionicamente derivatizzati, oppure polifosfazeni.

In una forma di realizzazione la composizione immunogena comprende inoltre antigeni aggiuntivi. In una forma di realizzazione gli antigeni aggiuntivi sono antigeni derivati da un batterio selezionato dal gruppo composto da S.pneumoniae, H.influenzae, N.meningitidis, E.coli, M.cattarrhalis, tetanus, diphtheria, pertussis, S.epidermidis,

enterococchi, *S.aureus*, e *Pseudomonas aeruginosa*. In un'ulteriore forma di realizzazione la composizione immunogena dell'invenzione può comprendere ulteriori antigeni da *C.difficile* per esempio le proteine del S-strato (WO01/73030).

C'è inoltre fornito un vaccino comprendente la composizione immunogena, questo vaccino può inoltre comprendere un eccipiente farmaceuticamente accettabile.

Le preparazioni vaccinali contenenti composizioni immunogene della presente invenzione possono essere usate per proteggere un mammifero suscettibile a infezione da *C.difficile* oppure per trattare un mammifero con una infezione da *C.difficile*, per mezzo di somministrazione di detto vaccino attraverso via sistemica oppure mucosale. Queste somministrazioni possono includere iniezione attraverso le vie intramuscolare, intraperitoneale, intradermica oppure sottocutanea; oppure attraverso somministrazione mucosale ai tratti orale/alimentare, respiratorio, genitourinario. Sebbene il vaccino dell'invenzione possa essere somministrato come una singola dose, componenti dello stesso possono anche essere co-somministrati insieme nello stesso tempo oppure in tempi differenti (per esempio coniugati di saccaridi da pneumococchi potrebbero essere somministrati separatamente, nello stesso tempo oppure 1-2 settimane dopo la somministrazione del qualunque componente proteico batterico del vaccino per la coordinazione delle risposte immunitarie l'una rispetto all'altra). Oltre a una singola via di somministrazione, si possono usare 2 differenti vie di somministrazione. Per esempio, saccaridi oppure coniugati di saccaridi possono essere somministrati per via intramuscolare (IM) oppure per via intradermica (ID) e proteine batteriche possono essere somistrate per via intranasale (IN) oppure per via intradermica (ID). Inoltre, i vaccini dell'invenzione possono essere somministrati IM per dosi di priming e IN per dosi di richiamo.

Il contenuto di tossine nel vaccino sarà tipicamente nel range 1-250µg, preferibilmente 5-50µg, il più tipicamente nel range 5 - 25µg. Dopo una vaccinazione iniziale, i soggetti possono ricevere una o diverse immunizzazioni di richiamo adeguatamente distanziate. La preparazione di vaccini è descritta in generale in Vaccine Design ("The subunit and adjuvant approach" (eds Powell M.F. & Newman M.J.) (1995) Plenum Press New York). L'incapsulamento all'interno di liposomi è descritto da Fullerton, Brevetto US 4,235,877.

Una applicazione dell'invenzione è un kit vaccinale, comprendente una fiala contenente una composizione immunogena dell'invenzione, facoltativamente in forma liofilizzata, e comprendente inoltre una fiala contenente un adiuvante come descritto nella presente. Si immagina che in questo aspetto dell'invenzione, l'adiuvante sarà usato per ricostituire la composizione immunogena liofilizzata.

Un'altra applicazione dell'invenzione è un metodo di prevenzione oppure di trattamento di infezione da *C.difficile* comprendente somministrare all'ospite una dose immunoprotettiva della composizione immunogena oppure del vaccino oppure del kit dell'invenzione. In una forma di realizzazione vi è fornito un metodo di prevenzione oppure di trattamento di episodi primari e/o ricorrenti di infezione da *C.difficile* comprendente somministrare all'ospite una dose immunoprotettiva della composizione immunogena oppure del vaccino oppure del kit dell'invenzione.

Un aspetto dell'invenzione è una composizione immunogena oppure un vaccino dell'invenzione per uso nel trattamento oppure nella prevenzione di malattia da *C.difficile*. In una forma di realizzazione vi è fornita una composizione immunogena dell'invenzione per uso nel trattamento oppure nella prevenzione di episodi primari e/o ricorrenti di malattia da *C.difficile*.

Una ulteriore applicazione dell'invenzione è uso della composizione immunogena oppure del vaccino oppure del kit dell'invenzione nella produzione di un medicamento per il trattamento oppure per la prevenzione di malattia da *C.difficile*. In una forma di realizzazione vi è fornita una composizione immunogena dell'invenzione per uso nella produzione di un medicamento per il trattamento oppure per la prevenzione di episodi primari e/o ricorrenti di malattia da *C. difficile*.

Intorno" oppure "approssimativamente" sono definiti come all'interno di 10% più o meno della cifra data per gli scopi dell'invenzione.

I termini "comprendente", "comprendono" e "comprende" nella presente si intende da parte degli inventori che sono facoltativamente sostituibili con i termini "costituito da", "sono costituiti da" e "è costituito da", rispettivamente, in ogni caso. Il termine "comprende" significa "include." Pertanto, eccetto se il contesto richiede altrimenti, la parola "comprende," e variazioni quali "comprendono" e "comprendente" sarà compreso che

implicano l'inclusione di un composto oppure una composizione affermati (ad esempio, acido nucleico, polipeptide, antigene) oppure di un passo, oppure di un gruppo di composti oppure di passi, ma non l'esclusione di qualunque altro composto, composizione, passo, oppure gruppi degli stessi. L'abbreviazione, "e.g." è derivata dal Latino *exempli gratia*, e si usa nella presente per indicare un esempio non limitante. Pertanto, l'abbreviazione "e.g." è sinonimo del termine "per esempio".

Forme di realizzazione nella presente relative a "composizioni vaccinali" dell'invenzione sono applicabili anche a forme di realizzazione relative a "composizioni immunogene" dell'invenzione, e vice versa.

Eccetto se altrimenti spiegato, tutti i termini tecnici e scientifici usati nella presente hanno lo stesso significato come comunemente compreso da uno di ordinaria esperienza nella tecnica a cui appartiene la presente descrizione. Definizioni dei comuni termini in biologia molecolare si possono trovare in Benjamin Lewin, *Genes V*, pubblicato da Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew et al. (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); e Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, pubblicato da VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

I termini singolari "un," "uno/a", e "il/lo/la" includono riferimenti plurali eccetto se il contesto indica chiaramente altrimenti. Analogamente, la parola "oppure" si intende che include "e" eccetto se il contesto indica chiaramente altrimenti. Il termine "pluralità" si riferisce a due o più. Si deve inoltre comprendere che tutte le dimensioni di basi o di amminoacidi, e tutti i valori di peso molecolare o massa molecolare, dati per acidi nucleici oppure polipeptidi sono approssimativi, e sono forniti per la descrizione. Inoltre, limitazioni numeriche date rispetto a concentrazioni oppure a livelli di una sostanza, quale un antigene, possono essere approssimative.

Tutti i riferimenti oppure le domande di brevetto citati all'interno di questa descrizione di brevetto sono incorporati per riferimento nella presente nella loro interezza.

Affinché la presente invenzione possa essere meglio compresa, i seguenti esempi sono esposti. Questi esempi sono puramente a fini illustrativi, e non si devono interpretare come limitanti l'ambito dell'invenzione in alcun modo.

## ESEMPI

### **Esempio 1: Progettazione di cinque fusioni ToxA-ToxB da *C.difficile***

Sono state progettate proteine di fusione contenenti frammenti dei domini ripetitivi C-terminali di ToxA e ToxB. Queste fusioni contenevano un frammento del dominio ripetitivo C-terminale di ToxA e un frammento del dominio ripetitivo C-terminale di ToxB, e una giunzione tra l'estremità C-terminale del frammento ToxA e l'estremità N terminale del frammento ToxB. Due strategie sono state concepite, nella prima strategia; la fusione è stata progettata in modo tale che la lunga struttura a solenoide fosse mantenuta a livello della giunzione tra i due frammenti. Nella seconda strategia, i due frammenti delle fusioni sono separati da un linker per permettere loro un corretto ripiegamento indipendente.

La parte C-terminale di ToxA e B è composta da sequenze ripetute: ripetizioni corte (SR) e ripetizioni lunghe (LR) (PNAS 2005 vol 102 : 18373-18378).

La struttura 3D nota parziale per il dominio C-terminale di ToxA (PNAS 2005 Greco et al., vol 102 : 18373-18378 ; Nature Structural & Molecular biology 2006 vol 13(5) : 460-461 ; PDB codifica : 2F6E, 2G7C e 2QJ6).

Gli inventori hanno predetto che vi sono due tipi di interazioni importanti tra residui della parte C-terminale di ToxA e ToxB. La prima interazione avviene tra residui contenuti in una LR e la sua SR precedente ed è importante mantenere la struttura tipo a solenoide. Il secondo tipo di interazione avviene tra residui contenuti in una LR e la SR seguente e questa interazione media la funzione di legame a carboidrati della tossina.

Una nuova ripetizione "strutturale-funzionale" SR-LR-SR è stata definita. La struttura di questa ripetizione è stata mantenuta intatta nelle nostre fusioni progettate.

La Figura 2 rappresenta i domini C-terminali di ToxA e ToxB e la box definita "SR-LR-SR".

Le posizioni delle ripetizioni corte (SR) e lunghe (LR) di ripetizioni di ToxA e ToxB sono presentate nella tabella 1.

Una lista delle box "SR-LR-SR" contenute nel dominio C-terminale di ToxA e ToxB è presentata nella Tabella 2.

Tabella 2

Nome	posizione iniziale	posizione finale
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

Tabella 2

Infine, il numero delle SR tra due LR sarà mantenuto nelle fusioni progettate per mantenere la lunga struttura tipo solenoide.

Prima della progettazione delle giunzioni per le fusioni, sono state definite due ipotesi di lavoro: prima ipotesi, più corte sono le fusioni, più è la probabilità per le fusioni di essere stabilmente sovra-espresse; seconda ipotesi, secondo il concetto di box "SR-LR-SR", la posizione iniziale deve essere scelta allo scopo di garantire un corretto ripiegamento della prima SR di questa box SR-LR-SR precedentemente definita. Pertanto le fusioni iniziano all'inizio della SR che precede la box SR-LR-SR. Usando queste due ipotesi, tre posizioni di partenza sono state analizzate: residuo 2370, 2234 e 2121 di ToxA.

La posizione iniziale 2370 è stata esclusa. La posizione iniziale 2234 è stata anche esclusa poiché uno dei residui implicati nelle interazioni importanti per la stabilità strutturale della proteina non è conservato. Pertanto, è stato deciso che tutta la fusione progettata inizierà in residuo 2121 di ToxA.

Tutte le fusioni termineranno nell'ultimo residuo di ToxB.

Quattro fusioni (F1-4) sono state progettate allo scopo di mantenere l'intera fusione in una lunga struttura tipo solenoide tra i due frammenti di fusione.

Le fusioni 1 (F1) e 2 (F2) sono state progettate usando la stessa ipotesi. Tutte le sequenze di proteine SR di ToxA e ToxB erano state confrontate usando un software di allineamento multiplo (ClustalW - Thompson JD et al. (1994) *Nucleic Acids Res.*, 22, 4673-4680). Le sequenze più simili erano la terza SR VIII di ToxA e la terza SR II di ToxB e la terza SR III di ToxB. Allo scopo di fare una scelta tra queste due SR di ToxB, una modellazione di omologia strutturale (usando l'interfaccia SwissModel - Arnold K et al. (2006) *Bioinformatics*, 22, 195-201) è stata compiuta sulla parte C-terminale di ToxB usando la nota struttura 3D di dominio C-terminale di ToxA parziale (PDB codice : 2QJ6). Usando la terza SR VIII di ToxA, la migliore sovrapposizione strutturale locale (compiuta usando SwissPDBViewer - Guex N et al. (1997), *Electrophoresis* 18, 2714-2723) è stata ottenuta con la terza SR II di ToxB. Pertanto, due giunzioni sono state progettate: la prima è tra la terza SR VI di ToxA e la quarta SR II di ToxB (F1) e la seconda è tra la seconda SR VIII di ToxA e la terza SR II di ToxB (F2). Queste giunzioni sono presentate nella figura 3 e 4 rispettivamente.

Per progettare la fusione 3 (F3), una sovrapposizione strutturale globale è stata compiuta tra sia la struttura nota del dominio C-terminale parziale di ToxA sia la struttura predetta di dominio C-terminale di ToxB (usando i software SwissModel e SwissPDBViewer). La migliore sovrapposizione è stata trovata tra LR VII di ToxA e LR II di ToxB. Pertanto, è stato deciso di fare una giunzione in questa LR simile. La giunzione è stata compiuta innanzitutto in una regione dove la sequenza è conservata tra ToxA e ToxB, dopo di che al fine di mantenere nella parte della fusione di ToxA, i residui in interazione con la SR precedente e infine, al fine di mantenere nella parte di ToxB, i residui in interazione con la SR seguente. Questa giunzione è mostrata nella figura 5.

Per la progettazione di fusione 4 (F4), il dominio C-terminale di ToxB è stato diviso in 4 frammenti e una modellazione di omologia più precisa (SwissModel) è stata compiuta su di loro. La divisione è stata realizzata al fine di mantenere intatte le box "SR-LR-SR" (ogni dominio finisce alla fine della SR che segue una LR). Una sovrapposizione strutturale tra le strutture predette di questi frammenti e la struttura 3D nota di ToxA è stata fatta e la migliore sovrapposizione strutturale è stata ottenuta per la terza SR di ToxB (SR I) e l'ultima SR di ToxA

(terza SR VIII). Pertanto, la giunzione è stata fatta tra la seconda SR VIII di ToxA e la terza SRI di ToxB.

Questa progettazione è presentata nella figura 6.

L'ultima fusione (F5) è stata progettata al fine di permettere un ripiegamento corretto indipendente dei due frammenti della fusione. Il linker è stato aggiunto tra l'ultimo residuo della sequenza proteica di ToxA e l'inizio della quarta SR II di ToxB (sempre prendendo in considerazione l'importanza di una box "SR-LR-SR" intatta). Solamente un residuo (Glicina) esogeno è stato aggiunto come linker e situato tra due Glicine esistenti. Pertanto, il linker può anche essere descritto come composto da 3 Glicine circondate da beta-filamento noto (per ToxA) e predetto (per ToxB). Questa ultima progettazione è mostrata nella figura 7.

### **Esempio 2: Clonaggio espressione e purificazione delle proteine di fusione**

#### **Plasmide di espressione e ceppo ricombinante**

Geni codificanti per le proteine di fusione di domini C-terminali parziali di ToxA e ToxB (SEQ ID NO:3, 4, 5, 6 e 7) e una His tag sono stati clonati nel vettore di espressione pET24b(+) (Novagen) usando i siti di restrizione NdeI/XhoI usando procedure standard. Il costrutto finale è stato generato mediante la trasformazione di ceppo *E. coli* BLR (DE3) con il vettore di espressione ricombinante secondo il metodo standard con cellule trattate con CaCl<sub>2</sub> (Hanahan D. « Plasmid transformation by Simanis. » In Glover, D. M. (Ed), DNA cloning. IRL Press London. (1985): p. 109-135.).

#### **Ceppo ospite:**

BLR(DE3). BLR è un recA derivato di BL21. Ceppi aventi la designazione (DE3) sono lisogeni per un profago  $\lambda$  che contiene una T7 RNA polimerasi inducibile da IPTG. Lisogeni DE3 per  $\lambda$  sono progettati per l'espressione proteica da vettori pET. Questo ceppo è anche deficiente nelle proteasi *Ion* e *ompT*.

Genotipo : ceppo *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>)

#### **Espressione delle proteine ricombinanti:**

Un trasformatante *E.coli* è stato strisciato da agar piastra e usato per inoculare 200 ml di brodo LBT  $\pm$  1% (p/v) di glucosio + kanamicina (50  $\mu$ g/ml) per ottenere una O.D.600nm tra 0.1 -0.2. Le colture sono state incubate per l'intera notte a 37°C, 250 RPM.

Questa coltura per l'intera notte è stata diluita a 1:20 in 500 ml di terreno LBT contenente kanamicina (50 µg/ml) e cresciuta a 37°C ad una velocità di agitazione di 250 rpm finché la O.D.<sub>620</sub> raggiungeva 0.5/0.6.

A O.D.<sub>600nm</sub> intorno a 0.6, la coltura è stata raffreddata prima di indurre l'espressione della proteina ricombinante mediante aggiunta di 1 mM di isopropil β-D-1-tiogalattopiranoside (IPTG; EMD Chemicals Inc., numero catalogo: 5815) e incubata per l'intera notte a 23°C, 250 RPM.

Dopo induzione per l'intera notte (intorno a 16 ore), la O.D.<sub>600nm</sub> è stata valutata dopo induzione e la coltura è stata centrifugata a 14 000 RPM per 15 minuti e i pellet sono stati congelati a -20°C separatamente.

#### Purificazione:

Il pellet batterico è stato risospeso in tampone bicina 20 mM (pH 8.0) contenente 500 mM di NaCl e una miscela di inibitori di proteasi (Complete, Roche). Batteri sono stati lisati usando un sistema a Pressa Francese a 20 000 PSI. Componenti solubili (surnatante) e insolubili (pellet) sono stati separati mediante centrifugazione per esempio a 20 000g per 30 min a 4°C.

La proteina etichettata con 6-His è stata purificata sotto condizioni native su IMAC. I componenti solubili sono stati caricati su una colonna GE (15 ml per esempio) (caricata con Ni) pre-equilibrata con lo stesso tampone usato per la risospensione batterica. Dopo caricamento sulla colonna, la colonna è stata lavata con lo stesso tampone. Eluizione è stata compiuta usando un tampone bicina 20mM (pH 8.0) contenente 500 mM di NaCl e differenti concentrazioni di imidazolo (5-600 mM). Dopo analisi del gel, frazioni più pure sono state selezionate, concentrate e caricate su cromatografia SEC per un ulteriore passo di purificazione.

Frazioni contenenti le proteine di fusione sono state selezionate sulla base della purezza secondo SDS-PAGE e dializzate contro tampone bicina (Bicina 20mM, NaCl 150 mM, con oppure senza EDTA 5mM pH8.0), La concentrazione proteica è stata determinata usando DC Protein Assay di BioRad. Le proteine sono state così unite in pool, filtrate sterilmente su 0.22 µm, conservate a -80°C.

In alternativa, purificazione IMAC è stata preceduta da un passo di purificazione su DEAE usando tampone bicina 2mM (pH 8.0) per caricamento e lavaggio, ed eluizione usando un gradiente con lo stesso tampone ma con aggiunto NaCl 1M.

### **Esempio 3 - Clonaggio espressione e purificazione dei frammenti di Tox A e Tox B da *C.difficile* separati**

#### Plasmide di espressione e ceppo ricombinante.

Geni codificanti per i frammenti di proteina di ToxA e ToxB (SEQ ID NO:8 e SEQ ID NO:9) e una His tag sono stati clonati nel vettore di espressione pET24b(+) (Novagen) usando i siti di restrizione NdeI/XhoI usando procedure standard. Il costrutto finale è stato generato mediante la trasformazione di ceppo *E. coli* BLR (DE3) con il vettore di espressione ricombinante secondo un metodo standard con cellule trattate con CaCl<sub>2</sub> (Hanahan D. « Plasmid transformation by Simanis. » In Glover, D. M. (Ed), DNA cloning. IRL Press London. (1985): p. 109-135.).

#### Ceppo ospite:

BLR(DE3). BLR è un recA derivato di BL21. Ceppi aventi la designazione (DE3) sono lisogeni per un profago  $\lambda$  che contiene una T7 RNA polimerasi inducibile con IPTG. Lisogeni DE3 per  $\lambda$  sono progettati per espressione proteica da vettori pET. Questo ceppo è deficiente anche nelle proteasi *lon* e *ompT*.

Genotipo : ceppo *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 (TetR)

#### Espressione delle proteine ricombinanti:

Un trasformatante *E.coli* è stato strisciato da agar piastra e usato per inoculare 200 ml di brodo LBT  $\pm$  1% (p/v) di glucosio + kanamicina (50  $\mu$ g/ml) per ottenere O.D.<sub>600nm</sub> tra 0.1 -0.2. Colture sono state incubate per l'intera notte a 37°C, 250 RPM.

Questa coltura per l'intera notte è stata diluita a 1:20 in 500 ml di terreno LBT contenente kanamicina (50  $\mu$ g/ml) e cresciuta a 37°C ad una velocità di agitazione di 250 rpm finché l'O.D.620 raggiungeva 0.5/0.6.

Ad una O.D. a 600nm di circa 0.6, la coltura è stata raffreddata prima di indurre l'espressione della proteina ricombinante mediante aggiunta di 1 mM di isopropil  $\beta$ -D-1-tiogalattopiranoside (IPTG; EMD Chemicals Inc., numero catalogo: 5815) e incubata per l'intera notte a 23°C, 250 RPM.

Dopo l'induzione per l'intera notte (intorno a 16 ore), O.D. a 600nm è stata valutata dopo induzione e la coltura è stata centrifugata a 14 000 RPM per 15 minuti e i pellet sono stati congelati a -20°C separatamente.

#### Purificazione:

Il pellet batterico è stato risospeso in tampone bicina 20 mM (pH 8.0) contenente NaCl 500 mM supplementato con una miscela di inibitori di proteasi (Complete senza EDTA, Roche cat 11873580001) e benzonasi. (Roche cat 1.01695.0001). I batteri sono stati lisati usando un sistema a Pressa Francese 2 X 20 000 PSI. Componenti solubili (surnatante) e insolubili (pellet) sono stati separati mediante centrifugazione a 34 000g oppure 48 000g per 25-30 min a 4°C. Il surnatante è stato raccolto e filtrato su filtro da 0.22 µm.

La proteina etichettata con 6-His è stata purificata sotto condizioni native su IMAC. I componenti solubili sono stati caricati su una colonna GE (per esempio 15ml) (caricata con Ni) pre-equilibrata con lo stesso tampone usato per la risospensione batterica. Dopo caricamento, la colonna è stata lavata con lo stesso tampone.

#### Per ToxA:

Eluizione è stata compiuta usando un tampone bicina 20mM (pH 8.0) contenente NaCl 500 mM e differenti concentrazioni di imidazolo (5-100 mM). Dopo analisi del gel, frazioni più pure sono state selezionate, concentrate e caricate su cromatografia SEC (SUPERDEX™ 75) per un ulteriore passo di purificazione nello stesso tampone senza imidazolo.

#### Per ToxB:

Un secondo lavaggio è stato compiuto con tampone bicina 20mM (pH 8.0) contenente NaCl 500 mM e 0.5% di deossicolato oppure lo stesso tampone con 150 mM di NaCl. L'eluizione è stata compiuta usando un tampone bicina 20mM (pH 8.0) contenente NaCl 500 mM e differenti concentrazioni di imidazolo (10-500 mM). Dopo analisi del gel, le frazioni più pure sono state selezionate, supplementate con 5 mM di EDTA e caricate su cromatografia SEC (SUPERDEX™ 200) per ulteriore passo di purificazione nello stesso tampone con EDTA 5 mM.

Frazioni contenenti frammenti di ToxA oppure ToxB sono state selezionate sulla base di purezza secondo SDS-PAGE e dializzate contro tampone bicina (Bicina 20mM, NaCl 150 mM, pH8.0), la concentrazione proteica è stata determinata usando Protein Assay RCDC di BioRad. Le proteine sono state così unite in pool, filtrate sterilmente su 0.22 µm, conservate a -80°C.

#### **Esempio 4 - Valutazione dei pesi molecolari delle cinque fusioni ToxA-ToxB da *C.difficile***

Ultracentrifugazione analitica è usata per determinare l'omogeneità e la distribuzione delle dimensioni in soluzione delle differenti specie all'interno di un campione di proteine misurando la velocità alla quale le molecole si muovono in risposta a una forza centrifuga. Questo si basa sul calcolo dei coefficienti di sedimentazione delle differenti specie che sono ottenute mediante esperimento di velocità di sedimentazione, che dipendono dalla loro forma e massa molecolare.

1. Campioni proteici sono centrifugati in una ultracentrifuga analitica Beckman-Coulter PROTEOMELAB™ XL-1 a 42 000 RPM dopo che il rotore AN-60Ti era stato equilibrato a 15°C.

- a. proteina di fusione F1, 500µg/ml, Bicina 20mM, NaCl 150mM, pH8,0
- b. proteina di fusione F2, 500µg/ml, Bicina 20mM, NaCl 150mM, pH8,0
- c. proteina di fusione F3, 500µg/ml, Bicina 20mM, NaCl 150mM, pH8,0
- d. proteina di fusione F4, 500µg/ml, Bicina 20mM, NaCl 150mM, pH8,0
- e. proteina di fusione F5, 500µg/ml, Bicina 20mM, NaCl 150mM, pH8,0

2. Per la raccolta dei dati, sono state registrate 160 scansioni a 280nm ogni 5 minuti.

3. L'analisi dei dati è stata compiuta usando il programma SEDFIT per la determinazione della distribuzione dei C(S). La determinazione del volume specifico parziale delle proteine è stata compiuta con il software SEDNTERP dalla loro sequenza amminoacidica. SEDNTERP è stato anche usato per determinare la viscosità e la densità del tampone.

4. Il peso molecolare delle differenti specie è stato determinato dal tracciato di distribuzione dei C(S) (concentrazione vs coefficiente di sedimentazione), considerando che esso è una migliore rappresentazione dei dati grezzi rispetto alla distribuzione dei C(M) (concentrazione vs peso molecolare) per caratterizzare la distribuzione delle dimensioni di una miscela.

La Figura 8 descrive la distribuzione delle fusioni ToxA-ToxB come determinata dalla velocità di sedimentazione mediante ultracentrifugazione analitica.

Il peso molecolare delle specie più importanti rilevate dalla distribuzione dei C(S) di tutte le cinque proteine di fusione ToxA-ToxB corrisponde alla loro forma monomerica. I rapporti di attrito best fit determinati per le

cinque fusioni sono tutti tra 2 e 2,2. Questo può indicare che le proteine sono presenti in soluzione come una forma allungata, che sarebbe coerente con la struttura proteica.

**Esempio 5- valutazione di strutture secondarie e terziarie di fusioni ToxA-ToxB di *C.difficile* mediante dicroismo circolare e spettroscopia di fluorescenza**

Il dicroismo circolare si usa per determinare la composizione della struttura secondaria di una proteina misurando la differenza nell'assorbimento di luce polarizzata a sinistra versus luce polarizzata a destra che è dovuta a asimmetria strutturale. La forma e l'entità degli spettri CD nella regione del lontano-UV (190-250nm) sono differenti se una proteina esibisce una struttura a foglietto beta, a alfa-elica oppure a random coil. L'abbondanza relativa di ogni tipo di struttura secondaria in un dato campione di proteine può essere calcolata mediante confronto con spettri di riferimento.

La struttura terziaria di un campione di proteine può essere valutata mediante la valutazione della immobilizzazione degli amminoacidi aromatici. L'osservazione di un segnale CD nella regione del vicino-UV (250-50nm) può essere attribuibile alla polarizzazione di residui fenilalanina, tirosina e triptofano ed è una buona indicazione che la proteina è ripiegata in una struttura ben definita.

Si è usato il seguente protocollo:

1. Spettri Lontano-UV sono misurati usando un cammino ottico di 0,01 cm da 178 a 250nm, con una risoluzione di 1nm e un'ampiezza di banda su uno spettropolarimetro Jasco J-720. La temperatura della cella è mantenuta a 23°C mediante un blocco di cella termostata Peltier RTE-111. Un flusso di azoto di 10L/min è mantenuto durante le misurazioni.
2. Spettri Vicino-UV sono misurati usando un cammino ottico di 0,01cm da 250 a 300nm, con una risoluzione di 1nm e un'ampiezza di banda su uno spettropolarimetro Jasco J-720. La temperatura della cella è mantenuta a 23°C mediante un blocco di cella termostata RTE-111 Peltier. Un flusso di azoto di 6L/min è mantenuto durante le misurazioni.

L'osservazione degli spettri del Lontano-UV (figura 9) per tutte le cinque proteine di fusione ToxA-ToxB suggerisce un debole contenuto di strutture a alfa elica e un alto contenuto di strutture a foglietto beta. Inoltre,

tutte le proteine esibivano un massimo a 230nm, che è inusuale per proteine globulari solubili. Questa particolarità è stata ben caratterizzata nella letteratura ed è associata con un piccolo gruppo di proteine note per la loro assenza di alfa elica e il loro alto contenuto in foglietto beta e in amminoacidi aromatici (Zsila, Analytical Biochemistry, 391( 2009) 154-156). Quelle particolarità sono coerenti con la struttura che ci si aspetta per le proteine di fusione ToxA-ToxB. Sono state confrontate strutture cristalline di 13 proteine esibenti gli spettri CD caratteristici con un segnale positivo a 230nm (Protein Data Bank). Il contenuto di struttura secondaria medio di quelle proteine è 42% di foglietto beta  $\pm 9\%$  e 7% di alfa elica  $\pm 6\%$ . Questo indica fortemente che la firma spettrale delle proteine di fusione ToxA-ToxB è diagnostica di una proteina contenente alto contenuto di foglietto beta e basso di alfa elica.

L'osservazione della forma degli spettri vicino-UV (figura 10) per tutte le cinque proteine di fusione indica che almeno alcuni degli amminoacidi aromatici sono immobilizzati, il che è una forte indicazione di una struttura terziaria compatta e specifica. Inoltre, il trattamento della proteina con una concentrazione di urea denaturante causava la scomparsa del segnale vicino-UV, che è una indicazione aggiuntiva che questi spettri caratteristici erano dovuti a ripiegamento della proteina.

#### **Esempio 6 - Immunizzazione di topi con frammenti di Tox A oppure Tox B e fusioni ToxA-ToxB**

Topi Balb/C sono stati immunizzati con i costrutti descritti negli esempi 2 e 3.

##### Immunizzazione di topi

Gruppi di 15 topi Balb/c femmine sono stati immunizzati IM nei giorni 0, 14 e 28 con 3 $\mu$ g o 10  $\mu$ g dei frammenti separati di toxA e toxB (si veda esempio 2) come pure con proteine di fusione ToxA-ToxB (si veda esempio 3) adiuvantate con AS03B. Un gruppo di controllo di 10 topi è stato vaccinato con solo AS03B.

Titoli ELISA anti-ToxA e anti-ToxB sono stati determinati in singoli sieri raccolti nel giorno 42 (post III).

Titoli di inibizione di emoagglutinazione sono stati determinati in sieri Post III uniti in pool.

##### Risposta ELISA anti-ToxA e anti-ToxB: Protocollo

Campioni dei frammenti di toxA oppure toxB sono stati adesi a 1  $\mu$ g/ml in soluzione salina fosfato tamponata (PBS) su piastre di microtitolazione ad alto legame (Nunc MAXISORP™), per l'intera notte a 4°C. Le piastre

sono state saturate con PBS-BSA 1% per 30 min a TA con agitazione. Gli anti-sieri di topo sono pre-diluiti 1/500 in PBS-BSA 0.2%-TWEEN™ 0.05%. e poi, ulteriori diluizioni su base due sono state fatte nelle micropiastre e incubate a TA per 30 min con agitazione. Dopo lavaggio, anticorpo murino legato è stato rilevato usando Anti-Topo IgG (H+L) di Capra affiniPure coniugato a perossidasi da Jackson ImmunoLaboratories Inc. (rif: 115-035-003) diluito 1:5000 in PBS-BSA 0.2%-tween 0.05%. Gli anticorpi di rilevazione sono stati incubati per 30 min. a temperatura ambiente (TA) con agitazione. Il colore è stato sviluppato usando 4 mg di O-fenilenediammina (OPD) + 5 µl di H<sub>2</sub>O<sub>2</sub> per 10 ml a pH 4.5 di tampone citrato 0.1M per 15 minuti al buio a temperatura ambiente. La reazione è stata fermata con 50 µl di HCl, e la densità ottica (OD) è stata letta a 490 nm rispetto a 620 nm.

Il livello di anticorpi anti-ToxA oppure anti-ToxB presenti nei sieri è stato espresso in titoli del punto medio. Una GMT è stata calcolata per i 15 campioni in ogni gruppo di trattamento (10 per il gruppo di controllo).

#### Saggio di inibizione di emoagglutinazione: Protocollo

Diluizioni su base due seriali di antisieri uniti in pool di topo (25µl) sono state compiute in soluzione salina fosfato tamponata (PBS) in micropiastre a 96 pozzetti a fondo a U.

25 µl di Tossina A nativa (0,2 µg/pozzetto) sono stati poi aggiunti e le piastre sono state incubate a temperatura ambiente per 30 minuti.

Dopo incubazione, 50 µl di eritrociti di coniglio purificati diluiti a 2% sono stati aggiunti a ogni pozzetto. Le piastre sono state incubate a 37°C per 2 ore.

Le piastre sono state analizzate visivamente, con l'emoagglutinazione che si presenta come globuli rossi diffusi nel pozzetto e l'inibizione di emoagglutinazione osservata come un punto rosso depositato nel pozzetto.

I titoli di inibizione sono stati definiti come il reciproco della più alta diluizione del siero inibente l'emoagglutinazione.

#### Saggio di citotossicità

Cellule di fibroblasti IMR90 sono state coltivate a 37°C con 5% di CO<sub>2</sub>, in EMEM + 10% di siero bovino fetale + 1% di glutammina + 1% di antibiotici (penicillina-streptomicina-amfotericina) e sono state seminate in piastre

per coltura di tessuto a 96 pozzetti ad una densità di  $5 \cdot 10^4$  cellule/pozzetto.

Dopo 24h, il terreno delle cellule è stato rimosso dai pozzetti.

Diluizioni su base due seriali di antisieri uniti in pool di topi (50 $\mu$ l) sono state compiute in terreno per cellule.

50  $\mu$ l di Tossina B nativa (0.5ng/ml) sono poi aggiunti e le piastre incubate a 37°C con 5% di CO<sub>2</sub> per 24 ore.

Le cellule sono state osservate dopo 24 ore, e la percentuale di cellule arrotondate è stata determinata.

I titoli di inibizione sono stati definiti come il reciproco della diluizione più alta del siero inibente il 50% di arrotondamento delle cellule.

#### Risultati:

Risultati Elisa, usando anticorpi ToxA sono descritti nella figura 11. Anticorpi anti-Tox A sono stati indotti dopo immunizzazione con il solo ToxA ma anche con ognuna delle 5 fusioni.

Le proprietà funzionali di questi anticorpi sono state testate nel saggio di emoagglutinazione. Questo saggio è adatto soltanto alla valutazione di ToxA poiché nessuna emoagglutinazione è osservata con ToxB.

Titoli di inibizione di emoagglutinazione sono descritti nella figura 12. L'inibizione di emoagglutinazione è stata osservata con i sieri anti-frammento di ToxA oppure i sieri diretti contro ognuna delle fusioni ToxA-ToxB.

Un ELISA utilizzando anticorpi anti ToxB è stato anche compiuto; i risultati di questo sono illustrati nella Figura 13. Anticorpi anti-Tox B sono stati indotti dopo immunizzazione con il solo frammento ToxB ma anche con le fusioni F2, F3 e F4.

Titoli di inibizione di citotossicità sono descritti nella figura 14. I titoli di inibizione ottenuti usando sieri da topi immunizzati con il frammento di ToxB oppure le fusioni ToxA-ToxB sono stati maggiori di quelli ottenuti usando i sieri di controllo.

#### **Esempio 7 Progettazione, clonaggio, espressione e purificazione di 4 ulteriori proteine di fusione**

Quattro ulteriori proteine di fusione sono state progettate usando i principi di progettazione descritti nell'esempio 1, queste sono state chiamate F54 Gly (SEQ ID NO:21), F54 New (SEQ ID NO:23), F5 ToxB (SEQ ID NO:25) e F52 New (SEQ ID NO:27).

Queste proteine di fusione sono state espresse secondo la procedura descritta in esempio 2.

**Esempio 8 - valutazione di peso molecolare delle fusioni ToxA-ToxB da *C.difficile* descritte in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, e SEQ ID NO:27**

Il peso molecolare delle fusioni descritte in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, e SEQ ID NO:27 è stato determinato come descritto nell'esempio 4.

La Figura 15 descrive la distribuzione di queste quattro ulteriori proteine di fusione come determinata da velocità di sedimentazione mediante ultracentrifugazione analitica.

Il peso molecolare delle principali specie determinato dalla distribuzione dei C(S) di tutte le quattro fusioni di proteine descritte in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, e SEQ ID NO:27 corrisponde alla loro forma monomerică e tutte le proteine esibiscono proprietà di sedimentazione simili alle fusioni da F1 a F5.

**Esempio 9 - valutazione di strutture secondarie e terziarie di fusioni ToxA-ToxB da *C.difficile* descritte in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, e SEQ ID NO:27**

Le strutture secondarie e terziarie delle fusioni descritte in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, e SEQ ID NO:27 sono state valutate secondo il metodo descritto nell'esempio 5. Il CD nel lontano UV per queste proteine di fusione si può trovare nella figura 16, e gli spettri vicino UV per queste fusioni si possono trovare nella figura 17.

L'analisi degli spettri CD nel vicino e nel lontano UV delle proteine descritte in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, e SEQ ID NO:27 mostra che tutte e quattro hanno la stessa struttura ad alto contenuto di foglietto beta rispetto alle fusioni da F1 a F5. Inoltre, l'osservazione degli spettri nel vicino UV non mostra alcuna differenza significativa nella posizione degli amminoacidi aromatici nella struttura terziaria rispetto alle fusioni da F1 a F5.

**Esempio 10 - immunizzazione di topi con fusioni ToxA-ToxB**

Topi Balb/c sono stati immunizzati con i quattro costrutti di proteine di fusione F54 Gly (SEQ ID NO:21), F54 New (SEQ ID NO:23), F5 ToxB (SEQ ID NO:25) e F52 New (SEQ ID NO:27) come descritti nell'esempio 6.

Un ELISA è stato effettuato usando la risposta ELISA anti-ToxA e anti-ToxB: il protocollo è descritto nell'esempio 6 eccetto che qui i campioni dei frammenti di toxA oppure toxB sono stati adesi a 2µg/ml in

soluzione salina fosfato tamponata su piastre di microtitolazione ad alto legame. Un saggio di inibizione di emoagglutinazione è stato compiuto come descritto nell'esempio 6. Un saggio di citotossicità di toxB è stato compiuto come descritto nell'esempio 6. Un ulteriore saggio di citotossicità di toxA è stato compiuto come descritto sotto.

#### Saggio di citotossicità di ToxA

Cellule HT29 sono state coltivate a 37°C con 5% di CO<sub>2</sub> in DMEM +10% di siero bovino fetale +1% di glutammina +1% di antibiotici (penicillina-streptomina-amfotericina) e sono state seminate in piastre per coltura di tessuto a 96 pozzetti ad una densità di 5.10<sup>4</sup> cellule/pozzetto.

Dopo 24h, il terreno per cellule è stato rimosso dai pozzetti.

Diluizioni su base due seriali di antisieri uniti in pool di topi (50µl) sono state compiute in terreno per cellule.

50µl di Tossina B nativa (0.15ng/ml) sono poi aggiunti e le piastre incubate a 37°C con 5% di CO<sub>2</sub> per 48 ore.

Le cellule sono state osservate dopo 48 ore e la proporzione di cellule arrotondate è stata determinata.

I risultati dell'Elisa anti-toxA, dell'Elisa anti-toxB, dei saggi di inibizione dell'emoagglutinazione e di citotossicità sono descritti nelle Figure 18, 19, 20, 21 e 22 rispettivamente.

### RIVENDICAZIONI

1. Polipeptide comprendente un primo frammento e un secondo frammento, in cui

- (i) il primo frammento è un frammento di dominio ripetitivo di tossina A;
- (ii) il secondo frammento è un frammento di dominio ripetitivo di tossina B;
- (iii) il primo frammento ha una prima estremità prossimale;
- (iv) il secondo frammento ha una seconda estremità prossimale; e

in cui il primo frammento e il secondo frammento sono separati da meno di oppure esattamente 5 amminoacidi nella struttura primaria, in cui il polipeptide suscita anticorpi che neutralizzano sia tossina A sia tossina B, in cui la prima estremità prossimale è all'interno di una ripetizione corta, la seconda estremità prossimale è all'interno di una ripetizione corta e la prima estremità prossimale e la seconda estremità prossimale non interrompono porzioni ripetizione corta-ripetizione lunga-ripetizione corta.

2. Polipeptide secondo la rivendicazione 1 in cui il polipeptide suscita una risposta immunitaria protettiva in un ospite mammifero contro ceppi di *C.difficile*.

3. Polipeptide secondo una qualunque delle rivendicazioni da 1 a 2 in cui il primo frammento e/o il secondo frammento comprendono meno di 25%, 20%, 18% o 15% di struttura a alfa elica.

4. Polipeptide secondo qualunque precedente rivendicazione in cui il primo frammento e/o il secondo frammento comprendono più di 25%, 30%, 35%, 38% o 40% di struttura a foglietto beta.

5. Polipeptide secondo qualunque precedente rivendicazione in cui la prima estremità prossimale non è all'interno di amminoacidi 1878-1940, 2012-2074, 2146-2208, 2258-2322, 2394-2456, 2507-2569 oppure 2598-2660 di tossina A.

6. Polipeptide secondo qualunque precedente rivendicazione in cui la seconda estremità prossimale non è all'interno di amminoacidi 1881-1942, 2012-2074, 2144-2208 oppure 2278-2339 di tossina B.

7. Polipeptide secondo qualunque precedente rivendicazione in cui il primo frammento comprende almeno 100, 250, 400 oppure 450 amminoacidi.

8. Polipeptide secondo qualunque precedente rivendicazione in cui il secondo frammento comprende almeno

100, 200, 300 oppure 400 amminoacidi.

9. Polipeptide secondo qualunque precedente rivendicazione in cui la prima estremità prossimale è all'interno della porzione di ripetizioni VIII (amminoacidi 2645-2710) di tossina A.

10. Polipeptide secondo una qualunque delle rivendicazioni da 1 a 9 in cui la prima estremità prossimale è all'interno di amminoacidi 2700-2710 oppure 2680-2690 di tossina A.

11. Polipeptide secondo qualunque precedente rivendicazione in cui la seconda estremità prossimale è all'interno di porzione di ripetizioni I (amminoacidi 1834-1926) di tossina B.

12. Polipeptide secondo qualunque precedente rivendicazione in cui la seconda estremità prossimale è all'interno di amminoacidi 1860-1878 oppure 1854-1876 di tossina B.

13. Polipeptide secondo una qualunque delle rivendicazioni da 1 a 10 in cui la seconda estremità prossimale è all'interno di porzione di ripetizioni II (amminoacidi 1927-2057) di tossina B.

14. Polipeptide secondo una qualunque delle rivendicazioni da 1 a 10 oppure 13 in cui la seconda estremità prossimale è all'interno di amminoacidi 1960-1970, 1988-1998 oppure 1867-1877 di tossina B.

15. Polipeptide secondo una qualunque delle rivendicazioni 9, 10, 11 oppure 12 in cui la prima estremità prossimale è all'interno di porzione di ripetizioni VIII (amminoacidi 2645-2710) di tossina A e in cui la seconda estremità prossimale è all'interno di porzione di ripetizioni I (amminoacidi 1834-1926) di tossina B.

16. Polipeptide secondo una qualunque delle rivendicazioni 9, 10, 13 oppure 14 in cui la prima estremità prossimale è all'interno di porzione di ripetizioni VIII (amminoacidi 2645-2710) di tossina A e in cui la seconda estremità prossimale è all'interno di porzione di ripetizioni II (amminoacidi 1927-2057) di tossina B.

17. Polipeptide secondo una qualunque delle rivendicazioni 9, 10, 13 oppure 14 in cui la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2687-2710) e la seconda estremità prossimale è all'interno di ripetizione corta 4 di porzione di ripetizioni II di tossina B (amminoacidi 1988-2007).

18. Polipeptide secondo una qualunque delle rivendicazioni 9, 10, 13 oppure 14 in cui la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VIII (amminoacidi 2665-2686) di tossina

A e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni II di tossina B (amminoacidi 1968-1987).

19. Polipeptide secondo una qualunque delle rivendicazioni 9, 10, 11 oppure 12 in cui la prima estremità prossimale è all'interno di ripetizione corta 2 di porzione di ripetizioni VIII di tossina A (amminoacidi 2665-2686) e la seconda estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni I di tossina B (1877-1896).

20. Polipeptide secondo una qualunque delle rivendicazioni da 1 a 8, 9, 10, 11, 12 e 15 in cui la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2645-2686) e la seconda estremità prossimale è all'interno di ripetizione corta 1 di porzione di ripetizioni I di tossina B (amminoacidi 1834-1854).

21. Polipeptide secondo una qualunque delle rivendicazioni 9, 10, 13 oppure 14 in cui la prima estremità prossimale è all'interno di ripetizione corta 3 di porzione di ripetizioni VIII di tossina A (amminoacidi 2687-2710) e la seconda estremità prossimale è all'interno di ripetizione corta 4 di porzione di ripetizioni II di tossina B (amminoacidi 1988-2007).

22. Polipeptide secondo la rivendicazione 16 in cui la prima estremità prossimale è all'interno di amminoacidi 2700-2710 di tossina A e la seconda estremità prossimale è all'interno di amminoacidi 1960-1970 di tossina B.

23. Polipeptide secondo la rivendicazione 16 in cui la prima estremità prossimale è all'interno di amminoacidi 2680-2690 di tossina A e la seconda estremità prossimale è all'interno di amminoacidi 1960-1970 di tossina B.

24. Polipeptide secondo la rivendicazione 16 in cui la prima estremità prossimale è all'interno di amminoacidi 2700-2710 di tossina A e la seconda estremità prossimale è all'interno di amminoacidi 1988-1998 di tossina B.

25. Polipeptide secondo la rivendicazione 16 in cui la prima estremità prossimale è all'interno di amminoacidi 2680-2690 e la seconda estremità prossimale è all'interno di amminoacidi 1860-1878.

26. Polipeptide secondo qualunque precedente rivendicazione in cui il polipeptide è parte di una proteina di fusione più grande.

27. Polipeptide secondo qualunque precedente rivendicazione in cui il polipeptide comprende un frammento

immunogeno di SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7 SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25 oppure SEQ ID NO:27.

28. Polipeptide secondo qualunque precedente rivendicazione in cui il polipeptide comprende un frammento immunogeno di almeno 500, 600, 700, 750, 800, 850 oppure 900 amminoacidi di SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7 SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25 oppure SEQ ID NO:27.

29. Polipeptide secondo qualunque precedente rivendicazione in cui il polipeptide comprende più di 450, 475, 500, 525, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825 oppure 850 amminoacidi da tossina A.

30. Polipeptide secondo qualunque precedente rivendicazione in cui il polipeptide comprende meno di 850, 825, 800, 775, 750, 725, 700, 675, 650, 625, oppure 600 amminoacidi da tossina A.

31. Polipeptide secondo qualunque precedente rivendicazione in cui il polipeptide comprende più di 350, 375, 400, 425, 450, 475, 500 oppure 525 amminoacidi da tossina B.

32. Polipeptide secondo qualunque precedente rivendicazione in cui il polipeptide comprende meno di 525, 500, 475, oppure 450 amminoacidi da tossina B.

33. Polinucleotide codificante per il polipeptide secondo una qualunque delle rivendicazioni da 1 a 32.

34. Vettore comprendente il polinucleotide secondo la rivendicazione 33 legato a un promotore inducibile.

35. Composizione immunogena comprendente il polipeptide secondo una qualunque delle rivendicazioni da 1 a 32 e un eccipiente farmaceuticamente accettabile.

36. Composizione immunogena secondo la rivendicazione 35 comprendente inoltre un adiuvante.

37. Composizione immunogena secondo una qualunque delle rivendicazioni da 35 a 36 comprendente inoltre antigeni aggiuntivi.

38. Composizione immunogena secondo la rivendicazione 37 in cui gli antigeni aggiuntivi sono antigeni derivati da un batterio selezionato dal gruppo composto da *S.pneumoniae*, *H.influenzae*, *N.meningitidis*, *E.coli*, *M.cattarhalis*, tetanus, diphtheria, pertussis, *S.epidermidis*, enterococchi, e *S.aureus*.

39. Vaccino comprendente la composizione immunogena secondo una qualunque delle rivendicazioni da 35 a 38.

40. Composizione immunogena secondo una qualunque delle rivendicazioni 35-38 oppure il vaccino secondo la rivendicazione 39 per uso nel trattamento oppure nella prevenzione di malattia da *C.difficile*.

**FIGURA 1**

**SEQ ID NO:1 – sequenza di tossina A**

MSLISKEELIKLAYSIRPRENEYKTI LTNLDEYNKLTNNNNENKYQLKKNL NESIDVFMN  
KYKTSSRRNALS NLKDI LKEVILIKNSNTSPVEKNLHFVWIGGEVSDIALEYIKQWADI  
NAEYNIKLWYDSEAFVNTLTKKAIVESSTTEALQLLEEEIQNPQFDNMKFYKRRMEFIYD  
RQKRFINIYKYSQINKPTVPTIDDIKSHLVSEYNRDET VLESYRTNSLRKINSNHGIDIR  
ANSLFTEQELLNISYQELLNRGNLAAASDIVRLLALKNFVGGVYLDVDMPLGIHSDLFKTI  
SRPSSIGLDRWEMIKLEAIMKYKYYINNYTSENFDKLDQQLKDNFKLIESKSEKSEIFS  
KLENLNVSDLEIKIAFALGSVINQALISKQGSYLTNLVIEQVKNRYQFLNQHLNPAIESD  
NNFTDTTKIFHDSL FNSATAENSMFLTKIAPYLQVGFMPPEARSTISLSGPGAYASAYYDF  
INLQENTIEKTLKASDLIEFKFPENNLSQLTEQEI NSLWSFDQASAKYQFEKYVRDYTG  
SLSEDNQVDFNKN TALDKNYLLNKKI PSNNVEEAGSKNYVHYIQQLQDDISYEATCNLF  
SKNPKNSII IQRNMNESAKSYFLSDDGESILELNKYRI PERLKNKEKVKVTFIGHGKDEF  
NTSEFARLSVDSLSNEISSFLDTIKLDISPKNVEVNLGCMNMFSDYDFNVEETYPGKLLLS  
IMDKITSTLPDVNKSITIGANQYEVRI NSEGRKELLAHSGKWINKEEAIMSDLSSKEYI  
FFDSIDNKLKAKSKNIPGLASISEDIKTL LLDASVSPDTKFI LNNLKLNI ESSIGDYIYY  
EKLEPVKNI IHNSIDDLIDEFNLL ENVSDELYELKKNLNLDEKYLISFEDI SKNNSTYSV  
RFINKSNGESVYVETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNT  
LNAAFFIQSLIDYSSNKDVLNLDLSTSVKQQLYAQLFSTGLNTIYDSIQLVNLISNAVNDT  
INVLPTITEGPIVSTILDGINLGAAIKELLD EHDPLKKELEAKVGVLA INMSLSIAAT  
VASIVGIGAEVTIFLLPIAGISAGI PSLVNNELI LHDKATSVVNYFNHLSSEKKYGPLKT  
EDDKILVPIDDLVI SEIDFNNSIKLGT CNILAMEGGSGHTVTGNI DHFFSSPSISSHIP  
SLSIYSAIGIETENLDFSKKIMMLPNAPSRVFWWETGAVPGLRSL ENDGTRLLDSIRDLY  
PGKFYWRFYAFFDYAITTLKPVYEDTNIKIKLDKDRNFIMPTITTNEIRNKLSYSFDGA  
GGTYSLLLSSYPIS TNINLSKDDLWIFNIDNEVREISIENGTIKKGLIKDVL SKIDINK  
NKLIIGNQTI DFGSDIDNKDRYIFLTCELD DKSILIEINLVAKSYSLLSGDKNYLISN  
LSNTIEKINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSI IHYKDKSKNILEFYNDST  
LEFNKSKDFAEDINVMKDDINTITGKYVDNNTDKSIDFSISLVSKNQVKVNGLYL NES  
VYSSYLDVFKNSDGHHTSNFMNLF LDNISFWKLF GFENIN FVIDKYFTLVGKTNLGYVE  
FICDNNKNI DIYFGEWKTS SSKSTIFSGNGRNVVVEPIYNPDTGEDI STSLDFSYEPLYG  
IDRYINKVLIAPDLYTSLININTNYYSNEYYPEI IVLNPNTFHKKVINL DSSSFEYKWS  
TEGSDFILVRYLEESNKILQKIRIKGILSNTQS FNKMSIDFKDIK KLSLGYIMSNFKSF  
NSENELDRDHLGFKI IDNKTYYYDEDSKLVKGLININNSLFYFDPIEFNLVTGWQTINGK  
KYYFDINTGAALTSYKINGKH FYFNNDGVMQLGVFKGPDGF EYFAPANTQNNNIEGQAI  
VYQSKFLT LNKYYFDNNSKAVTGWRI INNEKYFNPNNIAA AVGLQVIDNNKYFNP  
TAIISKGWQTVNGSRYYFDTDTAIAFN GYKTIDGKH FYFSDCVVKIGVFST SNGFEYFA  
PANTYNNNIEGQAI VYQSKFLT LNKYYFDNNSKAVTGLQTI DSKYYFNTNTAE AATG  
WQTI DGKKYYFNTNTAE AATGWQTI DGKKYYFNTNTAIASTGYTI INGHFYFNTD GIMQ  
IGVFKGPNGF EYFAPANTDANNIEGQAILYQNEFLT LNKYYFGSDSKAVTGWRI INNK  
KYYFNPNNIAA IHLCTINNDKYYSYDGI LQNGYITIERNNFYFDANNESKMVTGVFKG  
PNGFEYFAPANTHNNNIEGQAI VYQNKFLT LNKYYFDNNSKAVTGWQTI DGKKYYFN  
NTAE AATGWQTI DGKKYYFN LNTAE AATGWQTI DGKKYYFNTNTFI ASTGYTS INGHFY  
FNTD GIMQIGVFKGPNGF EYFAPANTDANNIEGQAILYQNKFLT LNKYYFGSDSKAVT  
GLRTIDGKKYYFNTNTAVAVTGWQTINGK KYFNTNTSIASTGYTI ISGKH FYFNTD GIM  
QIGVFKGPDGF EYFAPANTDANNIEGQAIRYQNRFLYLHDNI YFGNNSKAATGWVTIDG  
NRYFEPNTAMGANGYKTI DNKNFYFRNGLPQIGVFKG SNGFEYFAPANTDANNIEGQAI  
RYQNRFLHLLGKI YFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGV  
DGVKAPGIY G

**SEQ ID NO:2 – Sequenza di tossina B**

MSLVNRKQLEKMANVRFRTQDEDEYVAI LDAL EYHNMSENTVVEKYLKLDINS LTDIYI  
DTYKKSGRNKALKKFKEYLVTEVLELKNNNLTPVEKNLHFVWIGGQINDTA INYINQWKD  
VNSDYNVNVFYDSNAFLINTLKKTVVESAINDTLESFRENLDNPRFDY NKFFRKRMEI IY  
DKQKNFINIYKAQREENPELI IDDIVKTYLSNEYSKEIDELNTYIEESLNKI TQNSGNDV  
RNFEFKNGESFNLYEQELVERWNLAAASDILRISALKEIGGMYLDVDMPLGIQPDLFES

IEKPSVTVDFWEMTKLEAIMKYKEYIPEYTSEHFDMLDEEVQSSFESVLASKSDKSEIF  
SSLGDMEASPLEVKIAFNKGIINQGLISVKDSYCSNLIIVKQIENRYKILNNSLNPAISE  
DNDFNNTTNTFIDSIMAEANADNRFMMELGKYLRVGFPPDVKTTINLSGPEAYAAAYQD  
LLMFKEGSMNIHLIEADLRNFEISKTNISQSTEQEMASLWSFDDARAKAQFEYKRNIFE  
GSLGEDDNLDFSQNIIVVDKEYLLEKISSLARSSERYIHYIVQLQGDKISYEACNLFAK  
TPYDSVLFQKNIEDSEIAYYYNPGDGEIQEIDKYKIPSIISDRPKIKLTFIGHGKDEFNT  
DIFAGFDVDSLSTEIEAAIDLAKEDISPKSIEINLLGCNMFSYSINVEETYPGKLLKVK  
DKISELMPSISQDSIIVSANQYEVIRINSEGRRELLDHSGEWINKEESI IKDISSKEYISF  
NPKENKITVKSLELSTLLQEI RNNNSNSDIELEEKVMLTECEINVISNIDTQIVEER  
IEEAKNLTSDSINYIKDEFKLIESI DALCDLKQQNELED SHFISFEDISETDEGFSIRF  
INKETGESIFVETEKTFSEYANHITTEEISKIKGTIFDVTNGKLVKKNLDTTHEVNTLN  
AAFFIQSLIEYNSKESLNSLVAMKVQVYAQLFSTGLNTITDAKVVELVSTALDETID  
LLPTLSEGLPIIATI IDGVSLGAAIKELSETSDPLLQEI EAKIGIMAVNLTTATTAIIT  
SSLGIASGFSILLVPLAGISAGIPSLVNNELVLRDKATKVVDYFKHVSIVETEGVFTLLD  
DKIMPPQDDLVI SEIDFNNSIVLGKCEIWRMEGGSGHTVTDIDHFFSAPSITYREPHL  
SIYDVLEVQKEELDLSKDLMLPNAPNRVFAWETGWTPLRSLDNDGTKLLDRIRDNYEG  
EFYWRVFAFIADALITTLKPRYEDTNRINLDSNTRSFIVPIITTEYIREKLSYSFYGSG  
GTIALSLSQYNGINIELSESDVWIIDVDNVVRDVTIESDKIKKGDLEIGILSTLSIEEN  
KIILNSHEINFSGEVNNGSNGFVSLTFSILEGINAIEVDLLSKSYKLLISGELKILMLNS  
NHIQQKIDYIGFNSELQKNIPYSFVDSEKENGFI NGSTKEGLFVSELDPVVLISKVYMD  
DSKPSFGYYSNNLKDVKVITKDNVNI LTGYLKD DIKISLSLTLQDEKTIKLN SVHLD  
GVAEILKFMNRKGNNTSDSLMSFLESMNIKSIFVNFLQSNIKFILDANFIISGTTSIGQ  
FEFICDENDNIQPYFIKFNLTNYTLYVGNRQNMIVEPNYDLDSDGDISSTVINFSQKY  
LYGIDSCVNKVVISPNITYTDEINITPVYETNNTYPEVIVLDANYINEKINVNINDLSIRY  
VWSNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSKQDVPVSEIILSFT  
PSYEDGLIGYDLGLVSLYNEKFYINNFMMVSGLIYINDSLYFKPPVNNLITGFVTVG  
DDKYYFNPINGGAASIGETIIDKNYYFNQSGVLQTVGFSTEDGFKYFAPANTLDENLEG  
EAI DFTGKLIIDENIYFFDDNYRGAVEWKELDGEMHYFS PETGKAFKGLNQIGDYKYYFN  
SDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFA  
HHNEDLGNEEGEEISYSGILNFNKKIYFDDSFTAVVGWKDLEDGSKYFDEDTAEAYIG  
LSLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGIIESGVQNI DDNYFYIDDNGIVQIG  
VFDTS DGKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESD  
KYYFNPETKKACKGINLIDDIKYYFDEKIMRTGLISFENNNYFNENGENMQFGYINIED  
KMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYI  
AATGSVIDGEEYFFDPDTAQLVISE

**SEQ ID NO:3 – sequenza di fusione 1**

MGWQTI DGKKYYFNTNTAIASTGYTIINGKHFFYFNTDGMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKYYFSGSDSKAVTGWRI INNKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAI VYQNKFLTLNGKYYFDNDSKAVTGWQTI DGKKYYFNLTAE  
AATGWQTI DGKKYYFNLTAEAAATGWQTI DGKKYYFNTNTFI ASTGYTISNGKHFFYFNTDGMQIGVFKGPNGFE  
YFAPANTDANNIEGQAILYQNKFLTLNGKYYFSGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKYYFN  
TNTSIASTGYTIISGKHFFYFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAI RYQNRFLYLHDNIYFFGNNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTI DNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAI RYQN  
RFLHLLGKIYFFGNNSKAVTGWQTINGKVYFMPDTAMAAAGLFEIDGVIYFFGVGVKAPGFVSINDNKHYFD  
DSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNKKIYFDDSFTA  
VVGWKDLEDGSKYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGIIESGVQNI DDNYFY  
IDDNGIVQIGVFDTS DGKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESDKYYFN  
PETKKACKGINLIDDIKYYFDEKIMRTGLISFENNNYFNENGENMQFGYINIEDKMFYFGEDGVMQIGVFNTPD  
GFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIDGEEYFFDPDTAQLVISE

**SEQ ID NO:4 – sequenza di fusione 2**

MGWQTI DGKKYYFNTNTAIASTGYTIINGKHFFYFNTDGMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKYYFSGSDSKAVTGWRI INNKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAI VYQNKFLTLNGKYYFDNDSKAVTGWQTI DGKKYYFNLTAE

AATGWQTI DGKKYFNLNTAE AATGWQTI DGKKYFNTNTFI ASTGYTS INGKHFFYFNTDGIMQIGVFKGPNGFE  
YFAPANTDANNIEGQAILYQNKFLT LNGKKYFSGSDSKAVTGLRTI DGKKYFNTNTAVAVTGWQTINGKKYFNFN  
TNTSI ASTGYTI ISGKHFFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYFNGNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTI DNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQ  
RFLHLLGKIYFNGNSKAVTGWQTINGKVYFMPDTAMAAAGGLNQIGDYKYFNSDGVMMQKGFVS INDNKHYFD  
DSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHNNEDLGNEEGEEISYSGILNFNNKIYFDDSFTA  
VVGWKDLEDGSKYFDEDTAEAYIGLSLINDGQYFNDGIMQVGFVTINDKVYFSDSGIIESGVQNI DDNYFY  
IDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESDKYFNF  
PETKKACKGINLIDDIKYFDEKIMRTGLISFENNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPD  
GFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDTAQLVISE

**SEQ ID NO:5 – sequenza di fusione 3**

MGWQTI DGKKYFNTNTAI ASTGYTI INGKHFFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKKYFSGSDSKAVTGWRI INNKKYFNPNAIAA IHLCTINNDKYFSDGILQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAI VYQNKFLT LNGKKYFDNDSKAVTGWQTI DGKKYFNLNTAE  
AATGWQTI DGKKYFNLNTAE AATGWQTI DGKKYFNTNTFI ASTGYTS INGKHFFYFNTDGIMQIGVFKGPNGFE  
YFAPANTDANNIEGQAILYQNKFLT LNGKKYFSGSDSKAVTGLRTI DGKKYFNTNTAVAVTGWQTINGKKYFNFN  
TNTSI ASTGYTI ISGKHFFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYFNGNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTI DNKNFYFRNGLPQIGVFKGSNGFEYFAHNNEDLGNEEGEEISYSG  
ILNFNNKIYFDDSFTAVVGWKDLEDGSKYFDEDTAEAYIGLSLINDGQYFNDGIMQVGFVTINDKVYFSD  
SGIIESGVQNI DDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIET  
TGWIYDMENESDKYFNF PETKKACKGINLIDDIKYFDEKIMRTGLISFENNYYFNENGEMQFGYINIEDKMF  
YFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPD  
TAQLVISE

**SEQ ID NO:6 – sequenza di fusione 4**

MGWQTI DGKKYFNTNTAI ASTGYTI INGKHFFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKKYFSGSDSKAVTGWRI INNKKYFNPNAIAA IHLCTINNDKYFSDGILQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAI VYQNKFLT LNGKKYFDNDSKAVTGWQTI DGKKYFNLNTAE  
AATGWQTI DGKKYFNLNTAE AATGWQTI DGKKYFNTNTFI ASTGYTS INGKHFFYFNTDGIMQIGVFKGPNGFE  
YFAPANTDANNIEGQAILYQNKFLT LNGKKYFSGSDSKAVTGLRTI DGKKYFNTNTAVAVTGWQTINGKKYFNFN  
TNTSI ASTGYTI ISGKHFFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYFNGNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTI DNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQ  
RFLHLLGKIYFNGNSKAVTGWQTINGKVYFMPDTAMAAAGGETI IDDKNYFNQSGVLQTVGFSTEDGFKYFA  
PANTLDENLEGEAIDFTGKLI IDENIYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYFNSDGV  
MQKGFVS INDNKHYFDDSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHNNEDLGNEEGEEISYSGI  
LNFNNKIYFDDSFTAVVGWKDLEDGSKYFDEDTAEAYIGLSLINDGQYFNDGIMQVGFVTINDKVYFSDS  
GIIESGVQNI DDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIET  
GWIYDMENESDKYFNF PETKKACKGINLIDDIKYFDEKIMRTGLISFENNYYFNENGEMQFGYINIEDKMFY  
FGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDT  
AQLVISE

**SEQ ID NO:7 – sequenza di fusione 5**

MGWQTI DGKKYFNTNTAI ASTGYTI INGKHFFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKKYFSGSDSKAVTGWRI INNKKYFNPNAIAA IHLCTINNDKYFSDGILQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAI VYQNKFLT LNGKKYFDNDSKAVTGWQTI DGKKYFNLNTAE  
AATGWQTI DGKKYFNLNTAE AATGWQTI DGKKYFNTNTFI ASTGYTS INGKHFFYFNTDGIMQIGVFKGPNGFE  
YFAPANTDANNIEGQAILYQNKFLT LNGKKYFSGSDSKAVTGLRTI DGKKYFNTNTAVAVTGWQTINGKKYFNFN  
TNTSI ASTGYTI ISGKHFFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYFNGNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTI DNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQ  
RFLHLLGKIYFNGNSKAVTGWQTINGKVYFMPDTAMAAAGGLFEIDGVIYFFGVDGKAPGIYGGGFVS INDN

KHYFDSDGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNANKIYYFD  
DSFTAVVGWGDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNI  
DNYFYIDDNGIVQIGVFDTSBGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESD  
KYYFNPETKKACKGINLIDDIKYYFDEKGMRTGLISFENNNYFNENGEMQFGYINIEDKMFYFGEDGVMQIGV  
FNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDTAQLVISE

**SEQ ID NO:8 Sequenza di singolo frammento di Tossina A**

MASTGYTSINGKHFYFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNKFLTLNGKYYFSGSDSKAVT  
GLRTIDGKYYFNTNTAVAVTGWQTINGKYYFNTNTSIASTGYTIISGKHFYFNTDGMQIGVFKGPDGFYFA  
PANTDANNIEGQAIRYQNRFLYLHDNIYFNGNSKAATGWVTIDGNRYFEPNTAMGANGYKTDNKNFYFRNGL  
PQIGVFKSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYFNGNSKAVTGWQTINGKVYFMPDTAMAAA  
GGLFEIDGVIYFFGVDGPKAP

**SEQ ID NO:9 – Sequenza di singolo frammento di Tossina B**

MILMSTSEENKVSQVKIRFVNFKDKTLANKLSFNFSKQDVPVSEIILSFTPSYEDGLIGYDLGLVSLYNEKF  
YINNFMMVSGLIYINDSLYFKPPVNNLITGFVTVGDDKYYFNPINGGAASIGETIIDKNYFNQSGVLQTV  
FSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIG  
DYKYYFNSDGMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGN  
EEGEEISYSGILNFNANKIYYFDSDFTAVVGWGDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVT  
INDKVFYFSDSGIIESGVQNIIDNYFYIDDNGIVQIGVFDTSBGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDV  
YFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGMRTGLISFENNNYFNENGEMQF  
GYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVI  
IDGEEYFDPDTA

**SEQ ID NO:10 – sequenza di frammento di tossina A da fusione 1**

MGWQTIIDGKYYFNTNTAIASTGYTIINGKHFYFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNEF  
LTLNGKYYFSGSDSKAVTGWRIINNKKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNGFYFAPANTHNNNIEGQAILYQNKFLTLNGKYYFDNDSKAVTGWQTIIDGKYYFNLNTAE  
AATGWQTIIDGKYYFNLNTAEAAATGWQTIIDGKYYFNTNTFIASSTGYTSINGKHFYFNTDGMQIGVFKGPNGF  
YFAPANTDANNIEGQAILYQNKFLTLNGKYYFSGSDSKAVTGLRTIDGKYYFNTNTAVAVTGWQTINGKYYFN  
TNTSIASTGYTIISGKHFYFNTDGMQIGVFKGPDGFYFAPANTDANNIEGQAIRYQNRFLYLHDNIYFNGNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTDNKNFYFRNGLPQIGVFKSNGFEYFAPANTDANNIEGQAIRYQ  
RFLHLLGKIYFNGNSKAVTGWQTINGKVYFMPDTAMAAAAGGLFEIDGVIYFFGVDGPKAP

**SEQ ID NO:11 – sequenza di frammento di tossina A da fusione 2**

MGWQTIIDGKYYFNTNTAIASTGYTIINGKHFYFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNEF  
LTLNGKYYFSGSDSKAVTGWRIINNKKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNGFYFAPANTHNNNIEGQAILYQNKFLTLNGKYYFDNDSKAVTGWQTIIDGKYYFNLNTAE  
AATGWQTIIDGKYYFNLNTAEAAATGWQTIIDGKYYFNTNTFIASSTGYTSINGKHFYFNTDGMQIGVFKGPNGF  
YFAPANTDANNIEGQAILYQNKFLTLNGKYYFSGSDSKAVTGLRTIDGKYYFNTNTAVAVTGWQTINGKYYFN  
TNTSIASTGYTIISGKHFYFNTDGMQIGVFKGPDGFYFAPANTDANNIEGQAIRYQNRFLYLHDNIYFNGNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTDNKNFYFRNGLPQIGVFKSNGFEYFAPANTDANNIEGQAIRYQ  
RFLHLLGKIYFNGNSKAVTGWQTINGKVYFMPDTAMAAAAG

**SEQ ID NO:12 – sequenza di frammento di tossina A da fusione 3**

MGWQTIIDGKYYFNTNTAIASTGYTIINGKHFYFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNEF  
LTLNGKYYFSGSDSKAVTGWRIINNKKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNGFYFAPANTHNNNIEGQAILYQNKFLTLNGKYYFDNDSKAVTGWQTIIDGKYYFNLNTAE  
AATGWQTIIDGKYYFNLNTAEAAATGWQTIIDGKYYFNTNTFIASSTGYTSINGKHFYFNTDGMQIGVFKGPNGF  
YFAPANTDANNIEGQAILYQNKFLTLNGKYYFSGSDSKAVTGLRTIDGKYYFNTNTAVAVTGWQTINGKYYFN

TNTSIASTGYTII SGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAI RYQNRFLYLHDNIYYFGNNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGNSNGFEY

**SEQ ID NO:13 – sequenza di frammento di tossina A da fusione 4**

MGWQTI DGKKYFNTNTAIASTGYTI INGKHFYFNTDGIMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKKYYFGSDSKAVTGWRI INNKKYFNPNNIAAAIHLCTINNDKYYSYDGI LQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNNGFEYFAPANTHNNNIEGQAI VYQNKFLTLNGKKYYFDNDSKAVTGWQTI DGKKYFNLNTAE  
AATGWQTI DGKKYFNLNTAEAAATGWQTI DGKKYFNTNTFIASSTGYTSINGKHFYFNTDGIMQIGVFKGPNNGFE  
YFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYFNTNTAVAVTGWQTINGKKYFNT  
TNTSIASTGYTII SGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAI RYQNRFLYLHDNIYYFGNNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGNSNGFEYFAPANTDANNIEGQAI RYQNR  
RFLHLLGKIYYFGNNSKAVTGWQTINGKVYFMPDTAMAAAG

**SEQ ID NO:14 – sequenza di frammento di tossina A da fusione 5**

MGWQTI DGKKYFNTNTAIASTGYTI INGKHFYFNTDGIMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQNEF  
LTLNGKKYYFGSDSKAVTGWRI INNKKYFNPNNIAAAIHLCTINNDKYYSYDGI LQNGYITIERNNFYFDANN  
ESKMVTGVFKGPNNGFEYFAPANTHNNNIEGQAI VYQNKFLTLNGKKYYFDNDSKAVTGWQTI DGKKYFNLNTAE  
AATGWQTI DGKKYFNLNTAEAAATGWQTI DGKKYFNTNTFIASSTGYTSINGKHFYFNTDGIMQIGVFKGPNNGFE  
YFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYFNTNTAVAVTGWQTINGKKYFNT  
TNTSIASTGYTII SGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAI RYQNRFLYLHDNIYYFGNNS  
KAATGWVTIDGNRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGNSNGFEYFAPANTDANNIEGQAI RYQNR  
RFLHLLGKIYYFGNNSKAVTGWQTINGKVYFMPDTAMAAAGGLFEIDGVIYFFGVGDGKAPGIYG

**SEQ ID NO:15 – Sequenza di frammento di tossina B da fusione 1**

GFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHNNEDLGNEEGEEISYSGILNF  
NNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGI I  
ESGVQNI DDNYFYIDDNGIVQIGVFDTS DGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYFGETYTIETGWI  
YDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYFNENGEMQFGYINIEDKMFYFGE  
DGMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDTAQL  
VISE

**SEQ ID NO:16 – Sequenza di frammento di tossina B da fusione 2**

GLNQIGDYKYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAH  
NEDLGNEEGEEISYSGILNFNNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIM  
QVGFVTINDKVFYFSDSGI IESGVQNI DDNYFYIDDNGIVQIGVFDTS DGYKYFAPANTVNDNIYGQAVEYSGLV  
RVGEDVYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYFNE  
NGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAA  
TGSVIIDGEEYFDPDTAQLVISE

**SEQ ID NO:17 – Sequenza di frammento di tossina B da fusione 3**

FAHNNEDLGNEEGEEISYSGILNFNNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFN  
DGIMQVGFVTINDKVFYFSDSGI IESGVQNI DDNYFYIDDNGIVQIGVFDTS DGYKYFAPANTVNDNIYGQAVEY  
SGLVRVGEDVYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNY  
YFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDE  
YIAATGSVIIDGEEYFDPDTAQLVISE

**SEQ ID NO:18 – Sequenza di frammento di tossina B da fusione 4**

GETIIDDKNYFNQSGVLQTVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKE  
L DGMHYFSPETGKAFKGLNQIGDYKYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQ

IGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTAEAYIGL  
SLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGIIESGVQNI DDNYFYIDDNGIVQIGVFDTS DGYKYFAPANT  
VNDNIYGQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESDKYFNPETKKACKGINLIDDIKYFDEKGM  
RTGLISFENNNYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWL  
DLDEKRYF TDEYIAATGSVIDGEEYFDPDTAQLVISE

**SEQ ID NO:19 – Sequenza di frammento di tossina B da fusione5**

GFVSINDNKHYFDDSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNF  
NNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTAEAYIGL SLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGI I  
ESGVQNI DDNYFYIDDNGIVQIGVFDTS DGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYFGETYTIETGWI  
YDMENESDKYFNPETKKACKGINLIDDIKYFDEKGMRTGLISFENNNYFNENGEMQFGYINIEDKMFYFGE  
DGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYF TDEYIAATGSVIDGEEYFDPDTAQL  
VISE

**SEQ ID NO :20 – Sequenza nucleotidica di F54 Gly**

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATATTTTTAATACCAACCCGCAATTGCAAGCACCGGCTATACCATTATCAACGGCA  
AACACTTTTATTTAACACCGACGGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACGGCTTGAATACTTTGCACCGGCAATACCGGATGC  
CAATAATATGAAGGCCAGGCCATTCTGTATCAGAAATGAAATTTCTGACCCCTGAACGGCAAAAAATACTACTTTGGCAGCGATAGCAAAGCAGTT  
ACCGGTTGGCGCATCATCAAAATAAGAAATATTACTTCAACCCGAAATGCAATTCAGCAATTCATCTGTGCACCATTAACAACGACAAAT  
ATTATTTACGATATGACGGTATTTCTGCAGAAATGGCTACATACCATCGAACCGCAACAACCTTTTATTTTCGATGCCAACACGAAAGCAAATGGT  
GACCGGTGTTTTCAAAGGCCCTAATGGTTTTGAGTATTTCTGCTCCGGCAACACCCATAATAACAACATGAAGGTCAGGCGATCGTTTTATCAG  
AACAAATTCCTGACGCTGAATGGTAAGAAATACTATTTGATAATGACAGCAAAGCCGTGACCGGCTGGCAGACAATTGACGGGAAGAAATAT  
ACTTTAATCTGAATACCGCAGAACGACCAACCGGTTGGCAACGATCGACGGTAAAGGTACTACTTCAACCTGAACACAGCCGAAGCAGCCAC  
AGGATGGCAGACTATGATGGAAAAAATACTATTTCAACACCAACACCTTTATTCATCTACCGGTTATACCGCATTAACGGTAAACATTTCT  
TACTTCAACACCGATGGTATCATGCAGATCGCGCTTTCAAAGTCCAAATGGTTTCAATACCTTTGCCCCGCCAATACAGATGCAAAATAACA  
TCGAGGGTCAGGCAATCTGTACCAAAACAAATTTCTGACCCGTAATGGGAAAAAATACTACTTTGGTAGCGATTCTAAAGCCGTTACCGGCT  
GCGTACCATTGATGGTAAAAATACTACTTTAATACGAATACAGCCGTGCGGTTACAGGCTGGCAGACCATTAACGGGAAAAAATACTATTTT  
AACACAAATACCGCATTTGCCCAACGGGTTATACCATTATTTGCGGTTAAACACTTCTACTTTAATACCGATGGTATTAACAATCGGAGTCT  
TTAAAGGACCTGATGGGTTCAAAATATTTGCGCCGCGAACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAAATCGCTTCT  
GTATCTGCAGACAACATCTATTTTGGCAACAATTCAAAAGCAGCCACCGGCTGGGTTACAATTGATGGCAACCGCTACTATTTCAACCG  
AATACCGCAATGGGTGCAAAATGGCTACAAAACCATCGATAATAAAAAATTTCTATTTTCGCAACGGTCTGCCGAGATCGGGGATTTAAAGGTA  
GCAACGGCTTCGAATACCTGCTCCAGCGAATACGGACGCGAACAATATTGAGGGTCAAGCGATTCTGTTATCAAAACCGTTTTCTGCATCTGCT  
GGGCAAAATCTACTACTTTGGCAATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGGATACCGCC  
ATGGCAGCACCGGTTGGCTGTTTGAATGATGGCGTATCTATTTTTGGTGTGGATGGTGTAAAGCACCGGGAATATACGGTGTACCG  
GCTTTGTGACCGTGGGTGATGATAAATACTATTTCAATCCGATTAACGGTGGTGCAGCGAGCATTGGCGAAACCATCATCGATGCAAAAACTA  
TTATTTCAACCCAGAGCGGTGTGCTGCAGACCGGTGTGTTTAGCACCGAAGATGGCTTAAATATTTTGGCCAGCAACCCCTGGATGAAAC  
CTGGAAGGCGAAGCGATTGATTTTACCGGCAAACTGATCATCGATGAAACATCTATTACTTTCGATGATAACTATCTGTTGCGGTGGTAAAGGTA  
AAGAACTGGATGGCGAAATGCATTTTCTCCGAAACCGGTTAAAGCGTTTAAAGGCTGAACAGATCGCGGATTACAAATACTACTTCAA  
CAGCGATGGCGTGTGCAAGAGGCTTTGTGAGCATCAACGATAACAACACTATTTTCGATGATAGCGGTGTGATGAAAGTGGCTATACCGAA  
ATTGATGGCAAACTTTCTACTTCCGGGAAACCGGCAAAATGCAGATTGGCGTGTTCATACCGAAGATGGTTTCAAATACTTCCGCGACCATA  
ACGAAGATCTGGGTAACGAAGAAGCGAAGAAATFAGCTATAGCGGCATCTGAACTTCAACAACAAATCTACTACTTTGATGATAGCTTTAC  
CGCGTGGTGGGTAAGGCTTGGAAAGATGGCAGCAAAATATTTGATGAAAGATACCGCGGAAGCGTATATTGGCCTGAGCCTGAACTTAAC  
GATGGCCAGTACTATTTAACGATGATGGCATTATGCAGGTGGGTTTCTGACCATTAAATGATAAAGTGTCTATTTAGCGGATAGCGGCATTA  
TTGAAAGCGCGCTGCAGAACATGATGATAACTACTTCTACATCGATGATAACGGCATTGTGCAGATCGCGCTTTTGTATACCGGATGGCTA  
CAAATATTTGCGACCGGCAATACCGTGAACGATAAATTTATGGCCAGCGGTTGAAATATAGCGGTCTGGTCCGTGGGCGAAGATGTGTAT  
TATTTCCGGGAAACCTATACCATCGAAACCGGCTGGATTTATGATATGGAAACGAAAGCGATAAATATTACTTTAATCCGGAAACGAAAAAG  
CGTGCAAAAGGCATTAACCTGATCGATGATACAAATACFATTTTGTGATGAAAGGCAATTTATGCGTACCCTGATAGCTTCCGAAACCAACA  
CTATTACTTCAACGAAACCGTGAATGCAGTTCGGCTACATCAACATCGAAGATAAATGTTCTACTTCCGGCAAGATGGTGTATGCAGATT  
GGTGTTTTTAACACCCCGGATGGCTTCAAATACTTTGCCCATCAGAATACCTGGATGAAATTTCAAGGTGAAAGCATTAACATACCGGCT  
GGCTGGATCTGGATGAAACGCTACTACTTCACCGATGAATACATTGCGGCGACCGGCGAGCGTATTATGATGGCGAAGAACTACTTCCGA  
TCCGGATACCGCGCAGCTGGTATTAGCGAACATCATCATCACCAT

**SEQ ID NO :21 – Sequenza amminoacidica di F54Gly**

MATGWQIDGKYYFNTNTAIASSTGYTIINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQNEFLTLN  
GKYYFGSDSKAVTGWRIINNKYYFNPNNIAAIIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDANNESKVMVTGVF  
KGNPFYFAPANTHNNNIEGQAIIVYQNKFLTLNGKYYFDNDSKAVTGWQIDGKYYFNLNTAEAAATGWQIDGKYY  
FNLNTAEAAATGWQIDGKYYFNTNTFIASSTGYTSINGKHFFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILY

QNKFLTLNGKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTWQTINGKYYFNTNTSIASTGYTIIISGKHFFYFNTDG  
IMQIGVFKGPDGFYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTDGNRYFEPNTAMGANGYKT  
IDNKNFYFRNGLPQIGVFKGSDGFYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTWQTINGKYYFMP  
DTAMAAAGGLFEIDGVIYFFGVDGVKAPGIYGGTGFVTVGDDKYYFNPINGGAASIGETIIDDKNYFNFQSGVLQTVFVS  
TEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQGIDYKYFFN  
SDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFFYFAENGEMQIGVFNTEGDKYFAHHNEDLGNEEGEIEISYSGIL  
NFNKKIYFDDSFYVVGWKLDEGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGIIESG  
VQNIDDNYFYIDDNGIVQIGVFDTSQYKYFAPANTVNDNIYQAVEYSGLVVRVGEDVYFGETYTIETGWIYDMENESD  
KYYFNPETKKACKGINLIDDIKYYFDEKIMRTGLISFENNNYFNENGENMQFGYINIEDKMFYFGEDGVMQIGVFNTPD  
GFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTEDEYIAATGSVIDGEEYFDPDTAQLVISEHHHHHH

**SEQ ID NO :22 – Sequenza nucleotidica di F54 New**

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATATTTTAAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTTATCAACGGCA  
AACACTTTTATTTTAAACACCGACGGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACGGCTTTGAATACCTTGCACCGGCAAAATACCGATGC  
CAATAATATTGAAGCCAGGCCATTCTGTATCAGAATGAATTTCTGACCCCTGAACGGCAAAAAATACTACTTTGGCAGCGATAGCAAAGCAGTT  
ACCGGTTGGCCATCATCAACAATAAGAAAATATTACTTCAACCCGAAATAATGCAATGCGAGCAATTCATCTGTGACCAATTAACAACGACAAAAT  
ATTATTTACAGTATGACGGTATTTCTGCAGAAATGGCTACATTACCATCGAACGCAACAACCTTTTATTTTCGATGCCAACACGAAAGCAAAATGGT  
GACCGGTGTTTCAAAGGCCCTAATGGTTTTGAGTATTTCCGCTCCGGCAACACCCATAATAACAACATGAAAGGTGACGGCATCGTTTATCAG  
AACAAATTCCTGACGCTGAATGGTAAGAAAATACTATTTTCGATAATGACAGCAAGCCGTGACCGGCTGGCAGACAATGACGGGAAGAAAATAT  
ACTTTAATCTGAATACCGCAGAAGCAGCAACCGGTTGGCAACGATCGACGGTAAAAAGTACTACTTCAACCTGAACACAGCCGAAGCAGCCAC  
AGGATGGCAGACTATGATGGAIAAAAAATACTATTTCAACACCAACACCTTTATTGCATCTACCGGTTATACCAGCATTAACGGTAAACATTTTC  
TACTTCAACACCGATGGTATCATGCAGATCGCGGTTTTCAAAGGTCCAAATGGTTTTGCAATACCTTTCGCCCTGCCAATACAGATGCAAAATAACA  
TCGAGGGTCAGGCAATCCTGTACCAAAAACAAATTTCTGACCCCTGAATGGGAAAAAATATTAATTTGGTAGCGATTCAAAGCCGTTACCGGCTC  
GCGTACCATTGATGGTAAAAAATACTACTTTAATACGAAATACCGGTTACAGGCTGGCAGACCAATTAACGGGAAAAAATACTATTTT  
AACACAATAACAGCATTCGCTCAACGGGTTATACCATTATTTCCGGTAAACACTTCTACTTTAATACCGATGGTATTAAGCAATCGGAGTCT  
TTAAAGGACCTGATGGGTTCGAATATTTTCGCCCTGCCAACCTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAATCGCTTTCT  
GATCTGCACGACAACATCTATATTTTGGCAACAAATCAAAGCAGCCACCGGCTGGGTACAAATGATGGCAACCGCTACTATTTTCGAACCG  
AATACCGCAATGGGTGCAAAATGGCTACAAAACCATCGATAATAAAAAATTTCTATTTTCGCAACGGCTTCGCCGAGATCGGGGTATTTAAAGGTA  
GCAACCGGCTCGAATACTTCGCCAGCGAATACGGACGGCAACAATATTGAGGGTCAAGCGATTCTGTTATCAAACCGTTTTCTGCATCTGCT  
GGCAAAATCTACTTTGGCAATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGGATACCGCC  
ATGGCAGCAGCCGGTGTCTGTTGAAATGATGGCGTATCTATTTTTTGGTGTGGATGGTGTAAAGCAGTTACCGGCTTTGTGACCGTGG  
GTGATGATAAATACTATTTCAATCCGATTAACGGTGGTGCAGCGAGCATTGGCGAAACCATCATCGATGACAAAACTATATTTCAACCGAG  
CGGTGTCTGCAGACCGGTTGTGTTAGCACCAGATGGCTTTAAATATTTTCGCCAGCGAACAACCTTGGATGAAACCTGGAAGGCGAAGCG  
ATTGATTTTACCGGCAACTGATCATCGATGAAAACATCTATTTACTTCGATGATACTATCTGTTGGTGGGTTGAATGGAAGAAGTGGATGGCG  
AAATGCATTTATTTCTCCGGAAACCGGTAAGCGTTTAAAGCCCTGAACCGATCGGGGATTACAAATACTACTTCAACAGCGATGGCGTGAT  
GCAGAAAGGCTTTGTGAGCATCAACGATAACAACACTATTTTCGATGATAGCGGTGTGATGAAAGTGGGCTATACCGAAATGATGGCAAAACAT  
TTCTACTTCGCCGAAAACGGCGAAATGCAGATTGGCGTGTCAATACCGAAGATGGTTTTCAAATACTTCGCCGACCATAACGAAGATCTGGGTA  
ACGAAGAAGGCGAAGAAATAGCTATAGCGGCATCCTGAACTTCAACAACAAAATCTACTACTTTGATGATAGCTTTACCGCGGTGGTGGGCTG  
GAAAGATCTGGAAGATGGCAGCAAAATATATTTTCGATGAAGATACCGCGAAGCGTATATTTGGCCTGAGCCTGATTAACGATGGCCAGTACTAT  
TTTAAACGATGATGGCATTATGCAGGTGGTTTCGTGACCATTAATGATAAAGTGTCTATTTTCAGCGATAGCGGCATTATTTGAAAGCGGCTGC  
AGAACATTTGATGATAACTACTTCTACATCGATGATAACCGCATTTGTCAGATCGCGGTTTTTGTATACCAGCGATGGCTACAAATATTTTCGCACC  
GGCCAAATACCGTGAACGATAACATTTATGGCCAGGCGGTGAAATATAGCGGCTGTTGGTGGTGGGCGAAGATGTGATTTATTTTCGCCGAAACC  
TATACCATCGAAACCGGCTGGATTTATGATATGGAACGAAAGCGATAAATATTACTTTAATCCGAAACGAAAAAGCGTGCAAAGGCATTA  
ACCTGATCGATGATATCAATACTATTTTGTATGAAAAAGGCATTATGCGTACCGGCTGATTTAGCTTCGAAAACAACAACATTTACTTCAACGA  
AAACGGTGAATGCACTTCGGCTACATCAACATCGAAGATAAATGTCTACTTCGGCGAAGATGGTGTATGCAGATTGGTGTTTTTAACACC  
CCGATGGCTTCAAATACTTTGCCATCAGAATACCCCTGGATGAAAAATTTCCGAAGGTGAAAGCATTAACATACCCGCTGGCTGGATCTGGATG  
AAAAACGCTACTACTTCCCGATGAATACATTTGCCGACCGCGCAGCGTATTTATGATGGCAAGAAATACTACTTCGATCCGGATACCGCGCA  
GCTGGTATTAGCGAACATCATCATCACCAT

**SEQ ID NO :23 Sequenza amminoacidica di F54 New**

MATGWQITIDGKKYYFNTNTAIASTGYTIIINGKHFFYFNTDGMQIGVFKGPDGFYFAPANTDANNIEGQAILYQN  
EFLTLNGKYYFGSDSKAVTWGRIINNKYYFNPNNIAIAIHLCTINNDKYYFSYDGLQNGYITIERNNFYFDANNESK  
MVTGVFKGPDGFYFAPANTHNNNIEGQAIYVQNKFLTLNGKYYFNDNSKAVTWQITIDGKKYYFNLNTAEAAATGWQITI

DGKKYYFNLNTAEAAATGWQIDGKKYYFNTNTFIASSTGYTSINGKHFYFNTDGMQIGVFKGPNGFYFAPANTDANNIE  
GQAILYQNKFLTLNGKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTWQTINGKYYFNTNTSIASSTGYTIIISGKHF  
YFNTDGMQIGVFKGPDGFYFAPANTDANNIEGQAIYRQNRFLYLDHNIYFGNNSKAATGWVTIDGNRYFEPNTAMG  
ANGYKTIDNKNFYFRNGLPQIGVFKGSDNGFYEYFAPANTDANNIEGQAIYRQNRFLHLGKIYFNGNSKAVTWQTINGK  
VYFMPDTAMAAAGGLFEIDGVIYFFGVGDKAVTGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTVF  
STEDGFYKIFAPANTLDENLEGEAIDFTGKLI IDENIYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYF  
NSDGVMMQKGFVSINDNKHVFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDEGFKYFAHNNEDLGNEEGEEISYSGI  
LNFNNKIYFDDSFYAVGVKDLLEDGSKYYFDEDTAEAYIGLSLINDGQYFNDGIMQVGFVTINDKVFYFSDSGIIES  
GVQNIIDDNYFIDNNGIVQIGVFDTSYGYKYFAPANTVNDNIYQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENES  
DKYYFNPETKACKGINLIDDIKYYFDEKGMRTGLISFENNYYFNENGENMQFGYINIEDKMFYFGEDGVMQIGVFNTP  
DGFYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPDTAQLVISEHHHHH

**SEQ ID NO :24 Sequenza nucleotidica di F5 ToxB**

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATATTTTAAATACCAACACCGCAATTGCAAGCACCGGCTATACCATATCAACGGCA  
AACACTTTTATTTAACACCGACCGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACCGGCTTGAATACTTTGCACCGGCAAAATACCGATGC  
CAATAATATGAAGGCCAGGCCATTCGTATCAGAAATGAATTTCTGACCCGTAACCGCAAAAAATACACTTTGGCAGCGATAGCAAAGCAGTT  
ACCGGTTGGCGCATCATCAACAATAAGAAAATATTACTTCAACCCGAATAATGCAATTCAGCAATTCATCTGTGCACCAATAACAACGACAAA  
ATTTATTTTCAGCTATGACGGTATTTCTGCAGAAATGGCTACATTTACCATCGAACGCAACAACCTTTTATTTTCGATGCCAACACGAAAGCAAAATGGT  
GACCGGTGTTTCAAAAGCCCTAATGGTTTTGTAGTATTTCCGTCGGCAACACCCATAATAACAACATGAAAGGTGACGGCATCGTTTATCAG  
AACAAATTCCTGACGCTGAATGTAAGAAAATACTATTTTCGATAATGACAGCAAAAGCCGTGACCGGCTGCCAGACAATGACGGGAAGAAAATAT  
ACTTTAATCTGAATACCGCAGAAAGCAGCAACCGGTTGGCAAAACGATCGACGGTAAAAAGTACTACTTCAACCTGAACACAGCCGAAGCAGCCAC  
AGGATGGCAGACTATGATGAAAAAATACTATTTCAACACCAACACCTTTATTTGCATCTACCGGTTATACCGCATTAACGGTAAACATTTTC  
TACTTCAACACCGCATGGTATCATGACGATCGCGGTTTTCAAAAGGTCCAAAATGGTTTCGAATACTTTGCCCTGCAATACAGATGCAAAATAACA  
TCGAGGGTCAGGCAATCCTGTACCAAAAACAAATTTCTGACCCGTAATGGGAAAAAATAATTAATTTGGTAGCGATTCAAAGCCGTACCGGCT  
GGGTACCATTTGATGTAATAAATACTACTTTAATACGAAATACAGCCGTTGGGGTTACAGGCTGGCAGACCAATAACGGGAAAAAATACTATTTT  
AACCAAAATACCGCATTTGCCTCACCGGGTTATACCATATTTTCGGGTAACACATTTCTACTTTAATACCGATGGTATTAAGCAATCGGATCT  
TTAAAGGACCTGATGGGTTTCGAATATTTTGGCCCTGGCAACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAAATCGCTTTCT  
GTATCTGCAGCACAACAATCTATTTATTTGGCAACAATTCAAAAGCAGCCACCGGCTGGGTTACAATTTGATGGCAACCGCTACTATTTTCGAACCG  
AATACCGCAATGGGTGCAATGGCTACAAAACCATCGATAAATAAATAATTTCTATTTTCGCAACGGCTGCGCGAGATCGGGGTATTTAAAGGTA  
GCAACCGCTTCGAATACCTCCAGCGAATACGGACCGCAACAATATGAGGGTCAAGCGATTCTGTTATCAAACCGTTTTCTGCATCTGCT  
GGGCAAAATCTACTACTTTGGCAATAACAGTAAAGCAGTTACTGGATGGCAGACAAATCAATGGTAAAGTGTACTATTTTATGCGGGATACCGCC  
ATGGCAGCAGCCGCTGCTCTGTTGAAATTTGATGGCGTGATCTATTTTTTTGGTGTGGATGGTGTAAAGCAGTACCGGCTGATTTATATTA  
ACGATAGCTGTATTTAAACACCGGTTGAATAACCTGATTTACCGGCTTTGTGACCGTGGGTGATGATAAATACTATTTCAATCCGATTTAA  
CGGTGGTGCAGCGACATTTGGCGAAACCATCATCGATGACAAAACATATTTTCAACACAGAGCGGTGTGCTGCAGACCGGTTGTTTAGCACC  
GAAGATGGCTTTAAATATTTTGGCCAGCGAACACCCCTGGATGAAACCTGGAAGGCGAAGCGATTGATTTTACCGGCAAACTGATCATCGATG  
AAAACATCTATTTACTTCGATGATAACTATCTGTTGGTGGTGGAAAGAACTGGATGGCGAAATGCATTTATTTCTCCGGAACCGGTTAA  
AGCGTTTAAAGGCTGAACAGATCGCGCATTTACAATACTACTTTCAACAGCGATGECGTGATGCGAAAAGGCTTTGTGAGCACTCAACGATAAC  
AAACACTATTTTCGATGATAGCGGTGTGATGAAAGTGGGCTATACCGAAATTTGATGGCAAACTTTCTACTTCCGCGAAAACCGCGAAATGCAGA  
TTGGCGGTGTTCAAATACCGAAGATGGTTTCAAATACTTCGCGCAACATAACGAAGATCTGGGTAACGAAGAAGGCGAAGAAATTAGCTATAGCGG  
CATCTGAACTTCAACAACAAAATCTACTACTTTGATGATAGCTTTACCGCGGTGGTGGGCTGGAAAGATCTGGAAGATGGCAGCAAAATATTTAT  
TTGATGAAGATACCGCGAAGCGTATATTTGGCCGAGCCTGATTAACGATGGCCAGTACTATTTTAAACGATGATGGCATATGCAGGTGGGTT  
TCGTGACCATTAATGATAAAGTGTCTATTTTCAGCGATAGCGCATTTTGAAGCGCGGTGCAGAACATTGATGATAACTACTTCTACATCGA  
TGATAACCGCATTTGTCAGATCGCGTTTTTGGATACCGCATGGCTACAAAATATTTTCGACCGGCCAAATACCGTGAACGATAACATTTATGGC  
CAGGCGGTGGAATATAGCGGTCTGGTGCCTGTGGGCGAAGATGTGTATTTTCGGCGAAACCTATACCATGAAACCGGCTGGATTTATGATA  
TGAAAACGAAAGCGATAAATAATTTACTTTAATCCGAAACGAAAAAGCGTCAAAAGGCAATTAACCTGATGATGATAAATACTATTTTGA  
TGAAAAGGCAATATGCGTACCGGCTGATAGCTTCGAAAACAACAATACTACTTCAACGAAACCGTGAATGCAGTTCGGCTACATCAAC  
ATCGAAGATAAATAAGTTCTACTTCCGCGAAGATGGTGTATGAGATGGTGTTTTAAACACCCCGGATGGCTTCAAATACTTTGCCATCAGA  
ATACCTGGATGAAATTTGAAAGGTGAAAGCATTAACATATACCGGCTGGCTGGATCTGGATGAAAAACCGTACTACTTACCGATGAATACAT  
TGGCGGACCGGACCGGTGATTTGATGGCGAAGAAATACTACTTTCGATCCGATACCGCGCAGCTGGTGTATAGCGAACATCATCATCAC  
CAT

**SEQ ID NO :25 Sequenza amminoacidica di F5 ToxB**

MATGQWTIDGKKYYFNTNTAIASTGYFI INKHFYFNTDGMQIGVFKGPNGFYFAPANTDANNIEGQAILYQN

EFLTLNGKKYYFGSDSKAVTCWRI INNKYYFNPNNALAAIHLCTINNDKYYFSYDGLONGYITIERNNFYFDANNESK  
MVTGVFKGPNPGEYFAPANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDNDSKAVTGWQTI DGKKYYFNLNTAEAAATGWQTI  
DGKKYYFNLNTAEAAATGWQTI DGKKYYFNTNTFIASGTGYTSINGKHFFYFNTDGMQIGVFKGPNPGEYFAPANTDANNIE  
GQAILYQNKFLTLNGKKYYFGSDSKAVTGLRFTI DGKKYYFNTNTFAVAVTGWQTINGKKYYFNTNTS IASGTGYT IESKHFF  
YFNTDGMQIGVFKGPNPGEYFAPANTDANNIEGQAIIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYFEPNTAMG  
ANGYKTI DNKNFYFRNGLPQIGVFKGPNPGEYFAPANTDANNIEGQAIIRYQNRFLYHLGKIYYFGNNSKAVTGWQTINGK  
VYFMPD TAMAAAGLFEIDGVIYFFGVGDKAVSGLIYINDSLYFKPPVNNLITGFVTVGDDKYYFNPINGGAASIGE  
TIIIDKNYYFNQSGVLQTVGFSTEDGFKYFAPANTLDENLEGEAIDFTGKLI IDENIYYFDDNYRGAVEWKELDGEMHYF  
SPETGKAFKGLNQIGDYKYYFNSDGMVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFFYFAENGEMQIGVFNTEGFKY  
FAHNEDLGNREGEESI SYSGILNFNK IYFDDSEF TAVVGWKDLEDGSKYYFDEDTAEAY IGLSLINDGQYYFNDDGIMQ  
VGFVTINDKVFYFSDSGIIESGVQNI DDNYFYIDDNGIVQIGVFDTS DGYKYFAPANTVNDNIYQAVEYSGLVRVGEDV  
YFGETYTIETGWLYDMENESDKYYFNPETKACKGINLIDDIKYYFDEKGMRTGLISFENNYYFNENEMQFGYINI  
EDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYGWLDLDEKRYFTDEYIAATGSVIIDGEEYFDPD  
TAQLVI SEHHHHHH

**SEQ ID NO :26 - Sequenza nucleotidica di F52 new**

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTATCAACGGCA  
AACACTTTTATTTTAAACACCGACGGCATTATGCAGATTGGTGTGTTTTAAAGGTCGCAACCGGCTTGAATACCTTGCACCGGCAAAATACCGATGC  
CAATAATATTGAAGCCAGGCCATTCTGTATCAGAATGAATTTCTGACCCCTGAACGGCAAAAAATACCTACTTTGGCAGCGATAGCAAAGCAGTT  
ACCGGTTGGCCGATCATCAACAATAAGAAAATATTACTTCAACCCGAAATAATGCAATFGCAGCAATTCATCTGTGCACCAATAACACGCAAAAT  
ATTATTTAGCAGTATGACGGTATTCTGCAGAAATGGCTACATTACCATCGAACGCAACAACCTTTTATTTCCGATGCCAACACGAAAGCAAAATGGT  
GACCGGTGTTTTCAAAGCCCTAATGGTTTTGAGTATTTCCGCTCCGGCAACACCCATAATAACAACATGAAGGTCAGGCGATCGTTTTATCAG  
AACAAATTCCTGACCTGAATGTAAGAAAATACTATTTCGATAATGACAGCAAGCCGCTGACCGGCTGGCAGACAATTGACGGGAAGAAAATATT  
ACTTTAATCTGAATACCGCAGAAGCAGCAACCGGTTGGCAACAGATCGACGGTAAAAAGTACTACTTCAACCTGAACACAGCCGAAGCAGCCAC  
AGGATGGCAGACTATGATGGAAAAAATACTATTTCACACCAACACCTTTATTGCATCTACCGGTTATACCAGCATTAACGGTAAACATTTCT  
TACTTCAACACCGATGGTATCATGCAGATCGCGGTTTTCAAAGTCCAAATGGTTTTCGAATACCTTTGCCCTGCCAATACAGATGCAAAATAACA  
TCGAGGTCAGGCAATCCTGTACCAAAACAAATTTCTGACCCCTGAATGGGAAAAAATATTACTTTGGTAGCGATTCAAAGCCGTTACCGGCTCT  
GCGTACCATTGATGTAATAAATACTACTTTAATAGCAATACAGCCGTTGCGGTTACAGGCTGCAGACCAATAACGGGAAAAAATACTATTTT  
AACACAAATACCGCATTCCTCAACGGGTTATACCATTATTTCCGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAAATCGGAGTCT  
TTAAAGGACCTGATGGTTTCGAATATTTGCGCCTGCGAACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAATCGCTTTCT  
GTATCTGCACGCAACATCTATTATTTGGCAACAATTCAAAAGCAGCCACCGGCTGGGTTACAATGATGGCAACCGCTACTATTTCGAACCG  
AATACCGCAATGGGTGCAAAATGGCTACAAAACCATCGATAATAAATAATTTCTATTTCGCAACCGGCTCTGCCGAGATCGGGGATTTAAAGGTA  
GCAACCGCTTCGAATACTTCGCTCCAGCGAATACGGACGGCAACAATATTGAGGGTCAAGCGATTCTGTTATCAAACCGTTTTCTGCATCTGCT  
GGGCAAAATCTACTACTTTGGCAATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGGATACCGCC  
ATGGCAGCAGCCGGTGGTCTGTTGAAATGATGGCGTATCTATTTTTTTGGTGTGGATGGTGTAAAGCAGTGAAGGCGCTGAACAGATCG  
GCGATTACAAATACTACTTCAACAGCGATGGCGTGATGCGAGAAAGGCTTTGTGAGCATCAACGATAACAACAACACTATTTTCGATGATAGCGGTGT  
GATGAAAGTGGGCTATACCGAAATGATGGCAACATTTCTACTTCCGGGAAAAACGGCAAAATGCAGATTGGCGTGTTCRAATACCGAAGATGGT  
TTCAAATACTTCGCGCACCATAACGAAGATCTGGGTAACGAAGAGGCGAAGAAATAGCTATAGCGGCATCCTGAACTTCAACAACAAAATCT  
ACTACTTTGATGATAGCTTTACCGCGTGGTGGGCTGAAAGATCTGGAAGATGGCAGCAAAATATTATTTCGATGAAGATACCGCGAAGCGTFA  
TATTGGCCTGAGCCTGATTAACGATGGCCAGTACTATTTAACGATGATGGCATTATGCAGGTGGGTTTCGTTACCATTAATGATAAAGTGTTC  
TATTTACCGGATAGCGGCTATTGAAAGCGGCTGCAGAACATGATGATAACTACTTCTACATCGATGATAACCGCATTTGTCGATACCGGCT  
TTTTTGATAACCGGATGGCTACAAATATTTCCGACCGGCCAATACCCTGAACGATAACATTTATGGCCAGGCGGTGGAATATAGCGGCTGGT  
CGGTGTGGCGAAGATGTGTATTTTCGGCGAAACCTATACCATCGAAACCGGCTGGATTATGATATGGAAAACGAAAGCGATAAATATTAC  
TTAATCCGGAACGAAAGCGTGCAAAGCATTAACTGATCGATGATATCAAATACTATTTTGATGAAAAGGCAATATGCGTACCGGCT  
TGATTAGCTTCGAAAACAACAATACTACTTCAACGAAACCGTGAATGCAGTTCCGGCTACATCAACATCGAAGATAAATGTTCTACTTCGG  
CGAAGATGGTGTATGCAGATTGGTGTTTTTAACACCCCGGATGGCTCAAATACCTTTGCCATCAGAAATACCGTGGATGAAAATTCGAAGT  
GAAAGCATTAACTATACCGGCTGGATGAAAACCGCTACTACTTACCCGATGAATACATFGCCGCGACCGGCAAGCGTGAATATT  
ATGGCGAAGAACTACTTTCGATCCGGATACCGCGCAGCTGGTATTAGCGAACATCATCATCACCAT

**SEQ ID NO :27 – Sequenza amminoacidica di F52 New**

MATGWQTI DGKKYYFNTNTAIASGTGYTI INKHFYFNTDGMQIGVFKGPNPGEYFAPANTDANNIEGQAILYQNEFLTLN

GKKYYFGSDSKAVTGWRIINNKKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDANNESKMTVGVF  
KGPNGFEYFAPANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDNDSKAVTGWQTDGKKYYFNLNTAEAAATGWQTDGKKYY  
FNLNTAEAAATGWQTDGKKYYFNTNTFFIASTGYTSINGKHFYFNTDGMQIGVFKGPNNGFEYFAPANTDANNIEGQAILY  
QNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTTAVAVTGWQTINGKYYFNTNTSIASTGYTIIISGKHFYFNTDG  
IMQIGVFKGPDGFEYFAPANTDANNIEGQAIIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYFEPNTAMGANGYKT  
IDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYFMP  
DTAMAAAGGLFEIDGVIYFFGVDGKAVKGLNQIGDYKYYFNSDGMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFY  
FAENGEMQIGVFNTEGDKFYFAHNEDLGNEEGEEISYSGILNFNNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTAEAY  
IGLSLINDGQYYFNDDGIMQVGFVTINDKVYFSDSGIIESGVQNI DDNYFYIDDNGIVQIGVFDTS DQYKYFAPANTVN  
DNIYGQAVEYSGLVRVGEDVYVYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGMRTGLISF  
ENNNYFNFENGEQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDE  
YIAATGSVIDGEEYFDPDTAQLVISEHHHHH

**SEQ ID NO:28 – Sequenza di frammento di tossina A di F54 Gly**

MATGWQTDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQN  
EFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDA  
NNEKMTVGVFKGPNNGFEYFAPANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDNDSKAVTGWQTDGKKYYFNLNT  
AEAAATGWQTDGKKYYFNLNTAEAAATGWQTDGKKYYFNTNTFFIASTGYTSINGKHFYFNTDGMQIGVFKGPNNG  
FEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTTAVAVTGWQTINGKYY  
FNTNTSIASTGYTIIISGKHFYFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAIIRYQNRFLYLHDNIYYFGN  
NSKAATGWVTIDGNRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIIRY  
QNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYFMPDTAMAAAGGLFEIDGVIYFFGVDGKAPGIYG

**SEQ ID NO:29 – Sequenza di frammento di tossina A di F54 New**

MATGWQTDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQN  
EFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDA  
NNEKMTVGVFKGPNNGFEYFAPANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDNDSKAVTGWQTDGKKYYFNLNT  
AEAAATGWQTDGKKYYFNLNTAEAAATGWQTDGKKYYFNTNTFFIASTGYTSINGKHFYFNTDGMQIGVFKGPNNG  
FEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTTAVAVTGWQTINGKYY  
FNTNTSIASTGYTIIISGKHFYFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAIIRYQNRFLYLHDNIYYFGN  
NSKAATGWVTIDGNRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIIRY  
QNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYFMPDTAMAAAGGLFEIDGVIYFFGVDGKAV

**SEQ ID NO:30 – Sequenza di frammento di tossina A di F5 ToxB**

MATGWQTDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGMQIGVFKGPNNGFEYFAPANTDANNIEGQAILYQN  
EFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAAIAIHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDA  
NNEKMTVGVFKGPNNGFEYFAPANTHNNNIEGQAIIVYQNKFLTLNGKKYYFDNDSKAVTGWQTDGKKYYFNLNT  
AEAAATGWQTDGKKYYFNLNTAEAAATGWQTDGKKYYFNTNTFFIASTGYTSINGKHFYFNTDGMQIGVFKGPNNG  
FEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTTAVAVTGWQTINGKYY  
FNTNTSIASTGYTIIISGKHFYFNTDGMQIGVFKGPDGFEYFAPANTDANNIEGQAIIRYQNRFLYLHDNIYYFGN  
NSKAATGWVTIDGNRYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIIRY  
QNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYFMPDTAMAAAGGLFEIDGVIYFFGVDGKAV

**SEQ ID NO:31 – Sequenza di frammento di tossina A di F52 New**

MATGWQTDGKYYFNTNTAIASTGYTIINGKHFFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQN  
EFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNAIAAHLCTINNDKYYFSYDGI LQNGYITIERNNFYFDA  
NNEskMVTGVFKGPNGFEYFAPANTHNNNIEGQAI VYQNKFLTLNGKKYYFDNDSKAVTGWQTDGKYYFNLNT  
AEAATGWQTDGKYYFNLNTAEAATGWQTDGKYYFNTNTFIASSTGYTSINGKHFFYFNTDGIMQIGVFKGPNG  
FEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKYYFNTNTAVAVTGWQTINGKKYY  
FNTNTSIASTGYTIISGKHFFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAI RYQNRFLYLHDNIYYFGN  
NSKAATGWVTIDGNRYFEPNTAMGANGYKTDNKNFYFRNGLPQIGVFKGSDGFEYFAPANTDANNIEGQAI RY  
QNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGKAV

**SEQ ID NO:32 – Sequenza di frammento di tossina B di F54Gly**

TGFVTVGDDKYYFNPIINGGAASIGETIIDDKNYFNFQSGVLQTVGFSTEDGFKYFAPANTLDENLEGEAIDFTGK  
LIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGMQKGFVSINDNKHYFDDSG  
VMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHNNEDLGNEEGEEISYSGILNFNNKIYYFDDSF TAVVG  
WKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNI DDNYFYIDD  
NGIVQIGVFDTS DGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESDKYYFNPET  
KKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFK  
YFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIDGEEYFDPDTAQLVISE

**SEQ ID NO:33 – Sequenza di frammento di tossina B di F54 New**

TGFVTVGDDKYYFNPIINGGAASIGETIIDDKNYFNFQSGVLQTVGFSTEDGFKYFAPANTLDENLEGEAIDFTGK  
LIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGMQKGFVSINDNKHYFDDSG  
VMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHNNEDLGNEEGEEISYSGILNFNNKIYYFDDSF TAVVG  
WKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNI DDNYFYIDD  
NGIVQIGVFDTS DGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYFGETYTIETGWIYDMENESDKYYFNPET  
KKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFK  
YFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIDGEEYFDPDTAQLVISE

**SEQ ID NO:34 – Sequenza di frammento di tossina B di F5 ToxB**

SGLIYINDSLYFKPPVNNLITGFVTVGDDKYYFNPIINGGAASIGETIIDDKNYFNFQSGVLQTVGFSTEDGFKY  
FAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSD  
GVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAHNNEDLGNEEGEEISYS  
GILNFNNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYF  
DSGIIESGVQNI DDNYFYIDDNGIVQIGVFDTS DGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYFGETYTI  
ETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYFNENGEMQFGYINIEDK  
MYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIAATGSVIDGEEYFDP  
DTAQLVISE

**SEQ ID NO:35 – Sequenza di frammento di tossina B di F52 New**

KGLNQIGDYKYYFNSDGMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFFAENGEMQIGVFNTEDGFKYFAH  
NNEDLGNEEGEEISYSGILNFNNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGI  
MQVGFVTINDKVFYFSDSGIIESGVQNI DDNYFYIDDNGIVQIGVFDTS DGYKYFAPANTVNDNIYGQAVEYSGL  
VRVGEDVYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGI MRTGLISFENNNYFN  
ENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYFTDEYIA  
ATGSVIDGEEYFDPDTAQLVISE

FIGURA 2

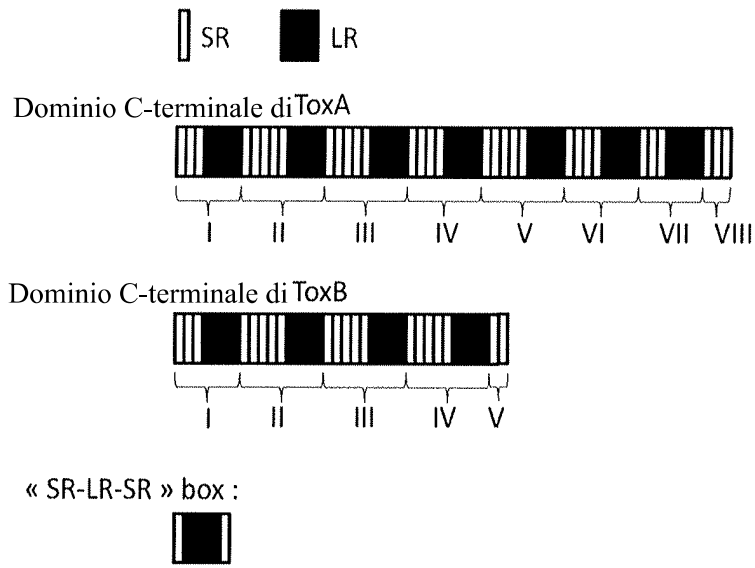


FIGURA 3

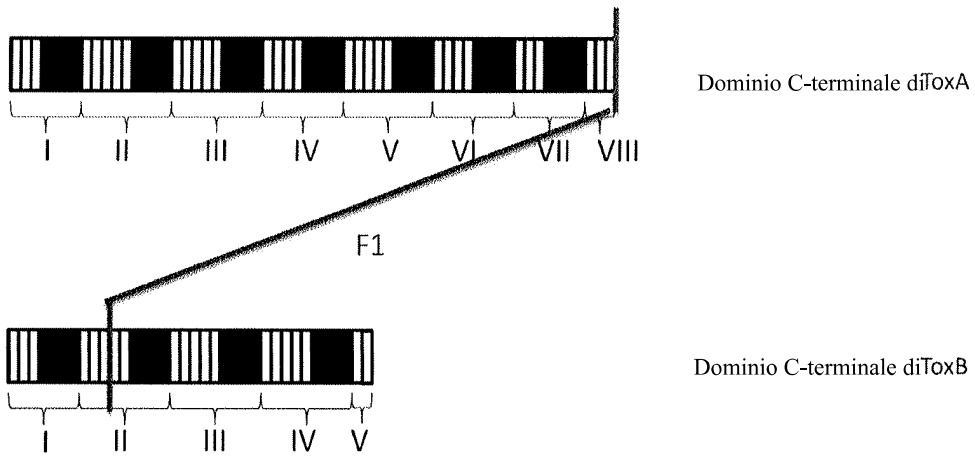


FIGURA 4

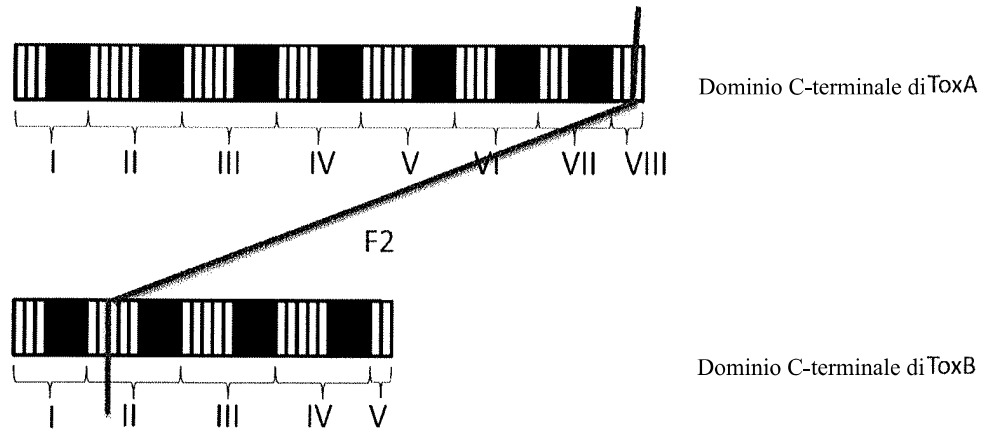


FIGURA 5

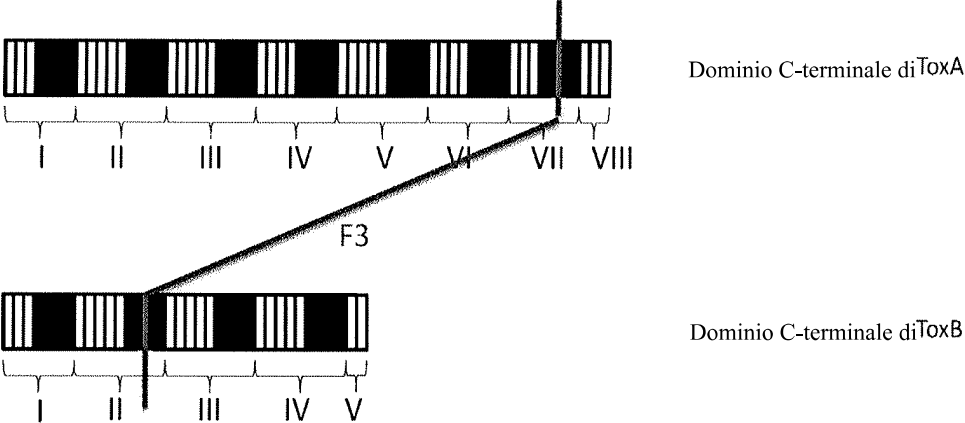


FIGURA 6

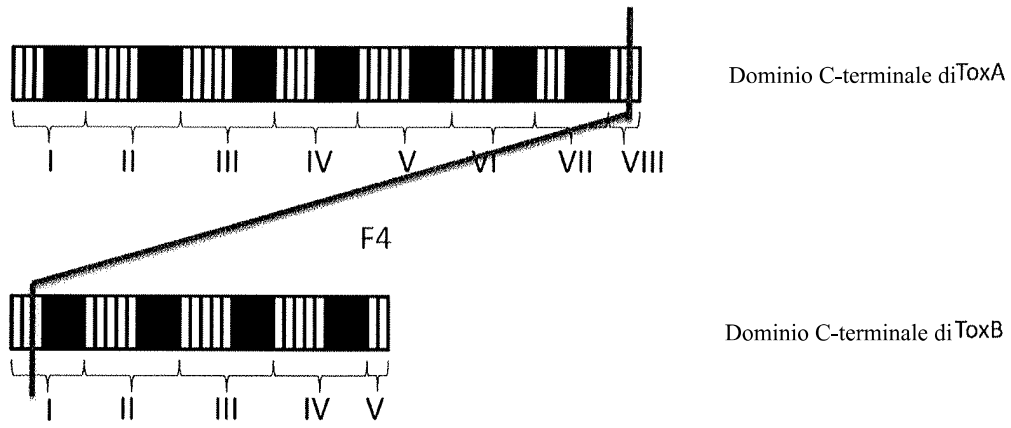


FIGURA 7

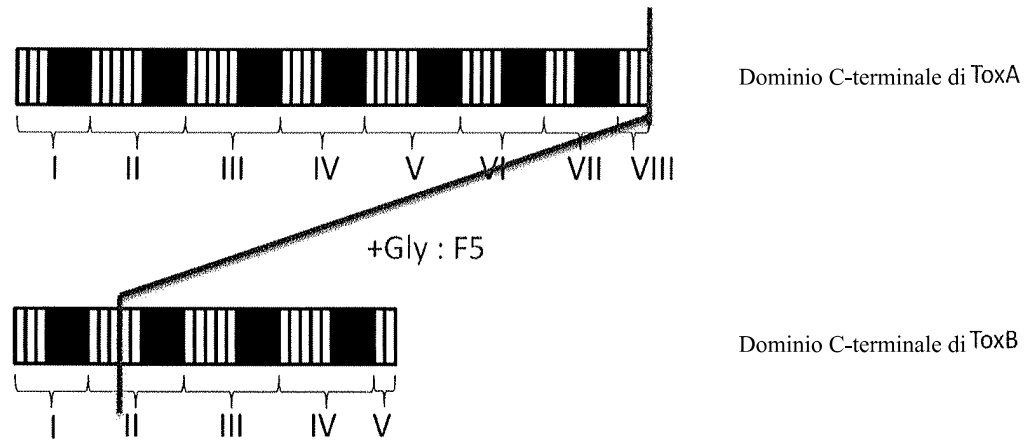
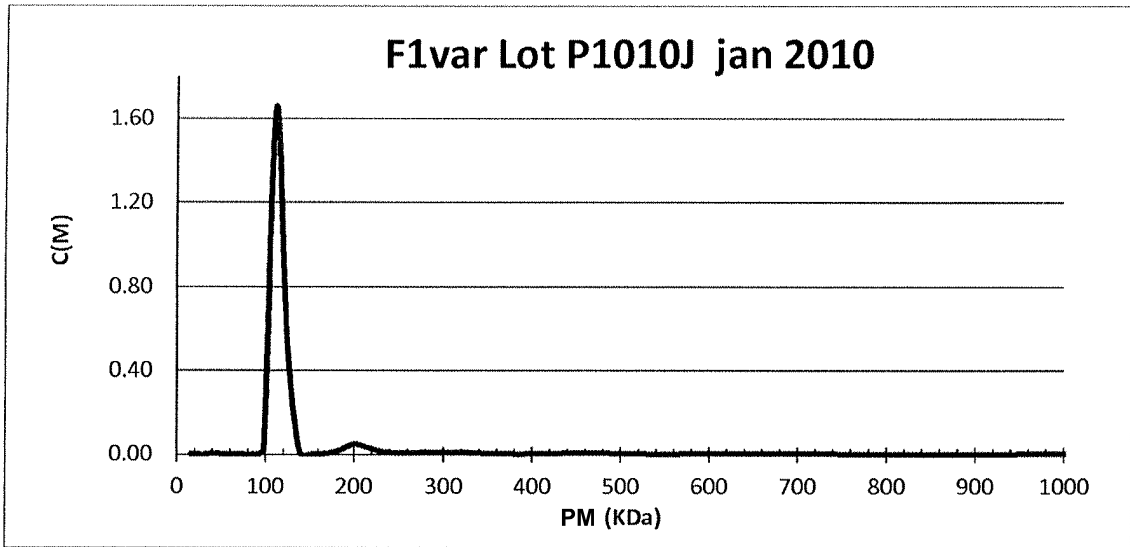
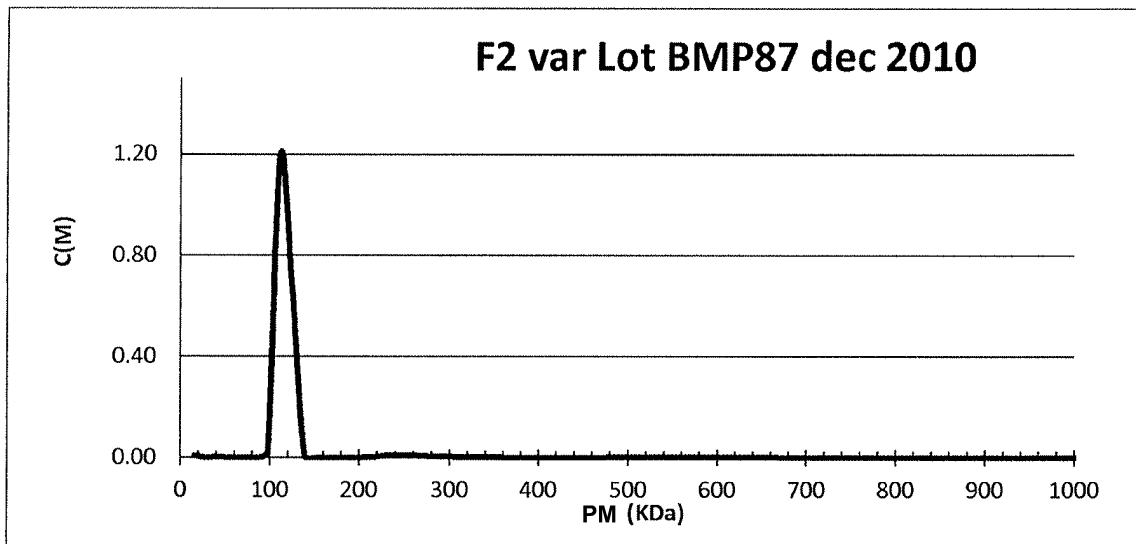


FIGURA 8

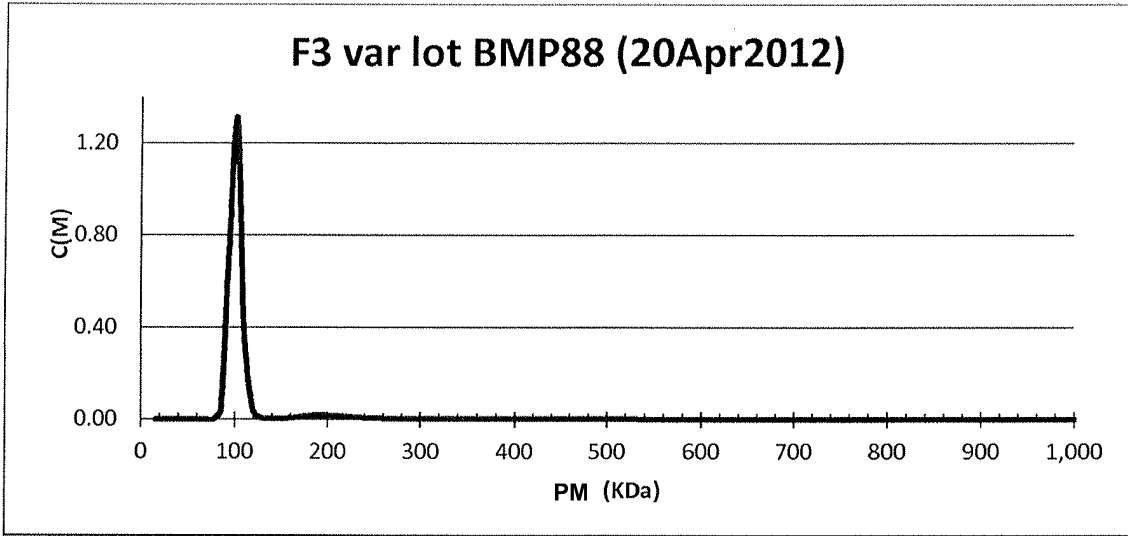
a)



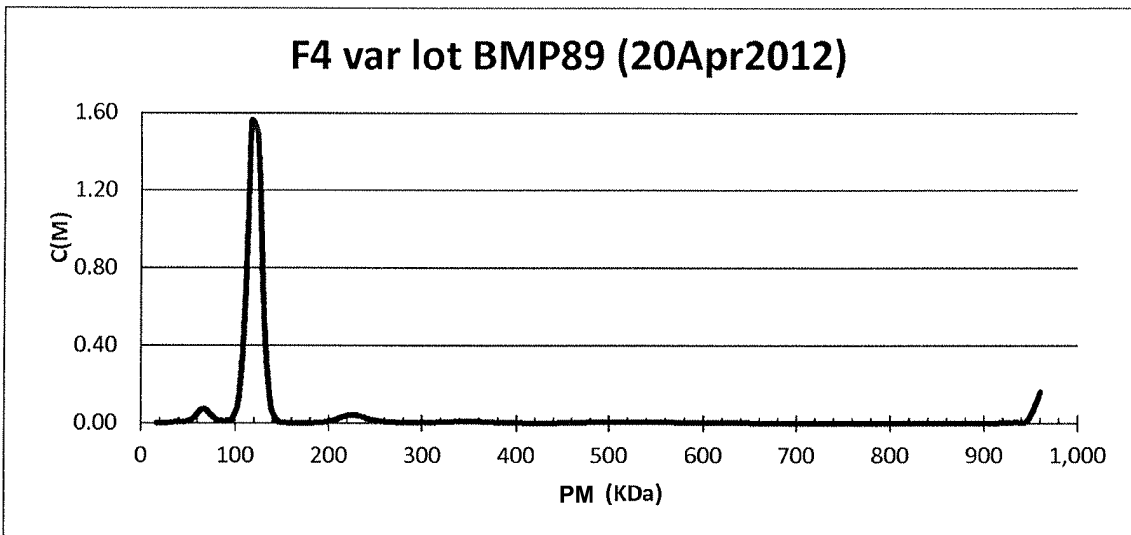
b)



c)



d)



e)

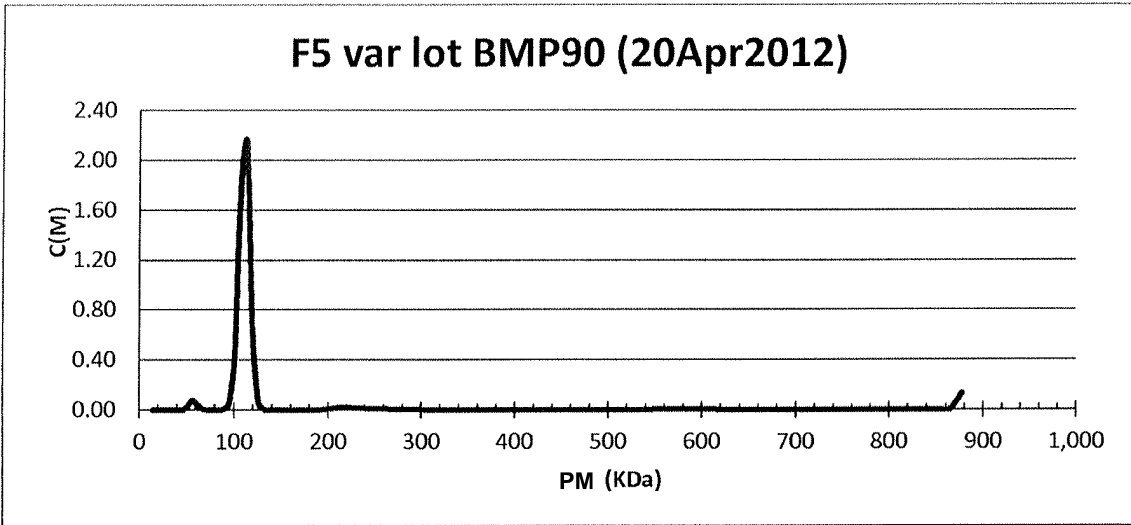


FIGURA 9

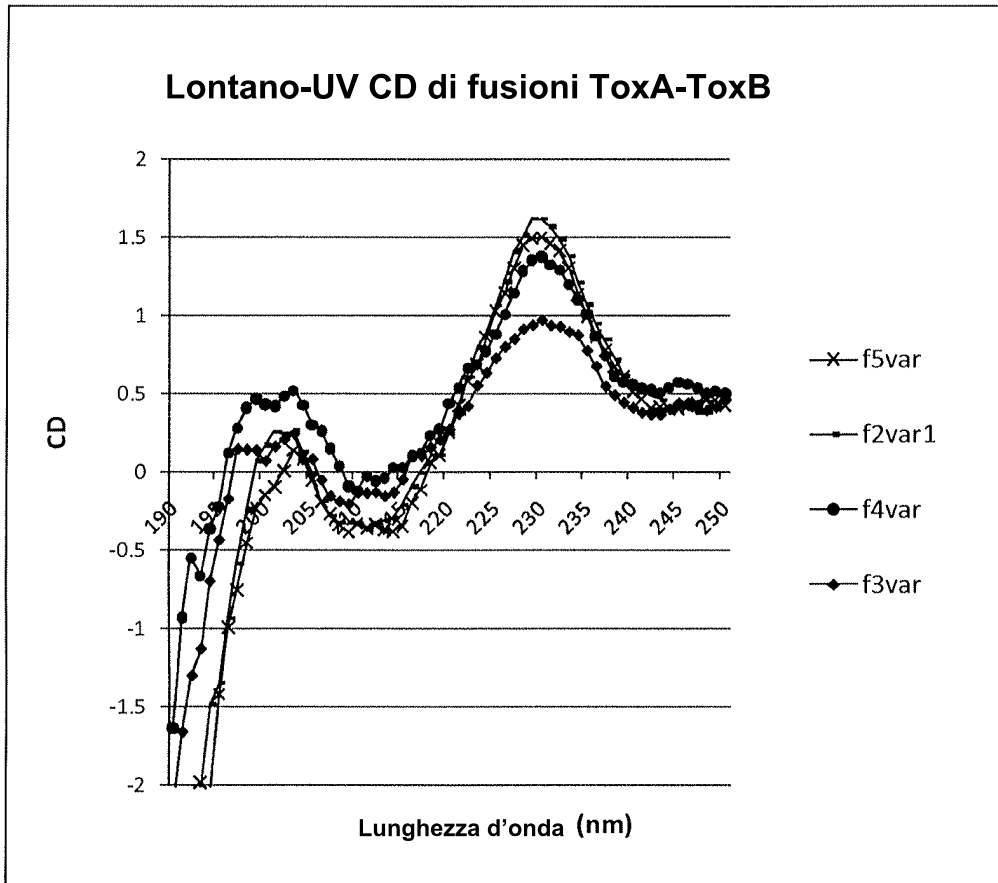


FIGURA 10

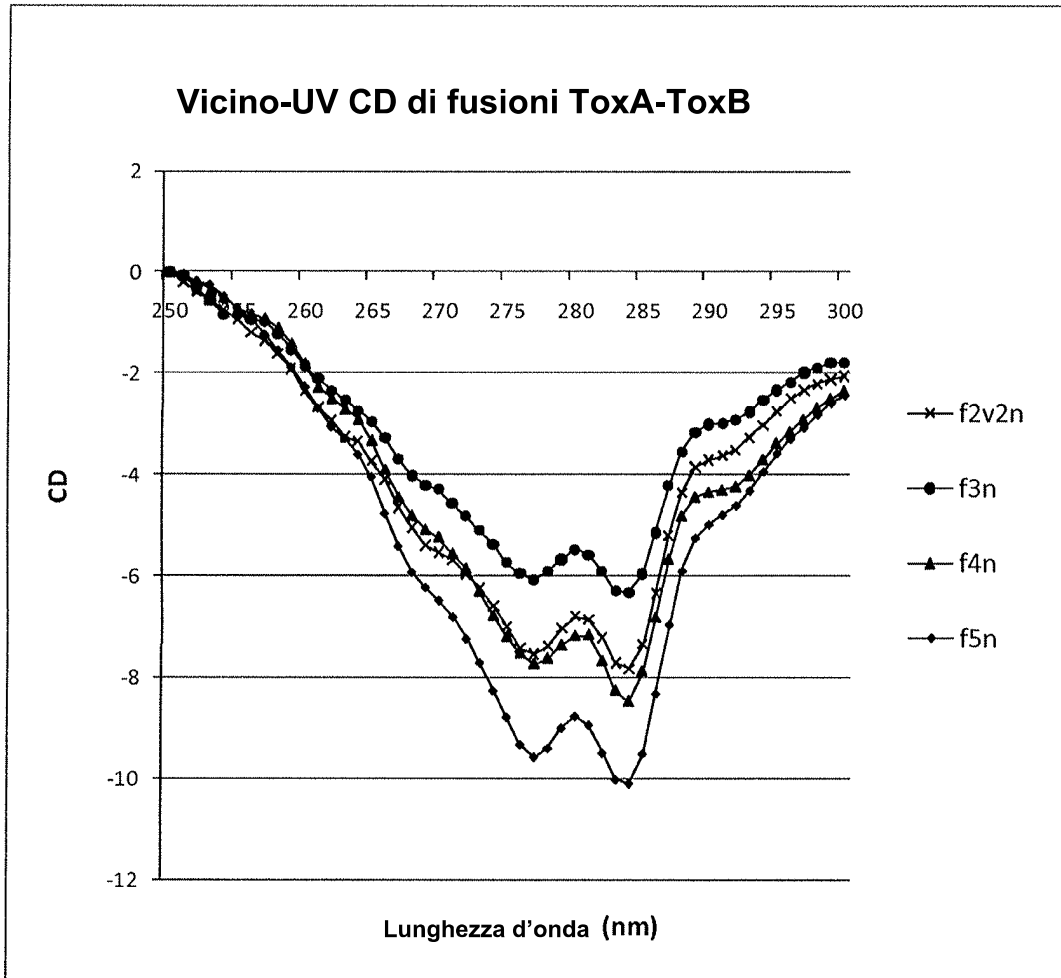


FIGURA 11

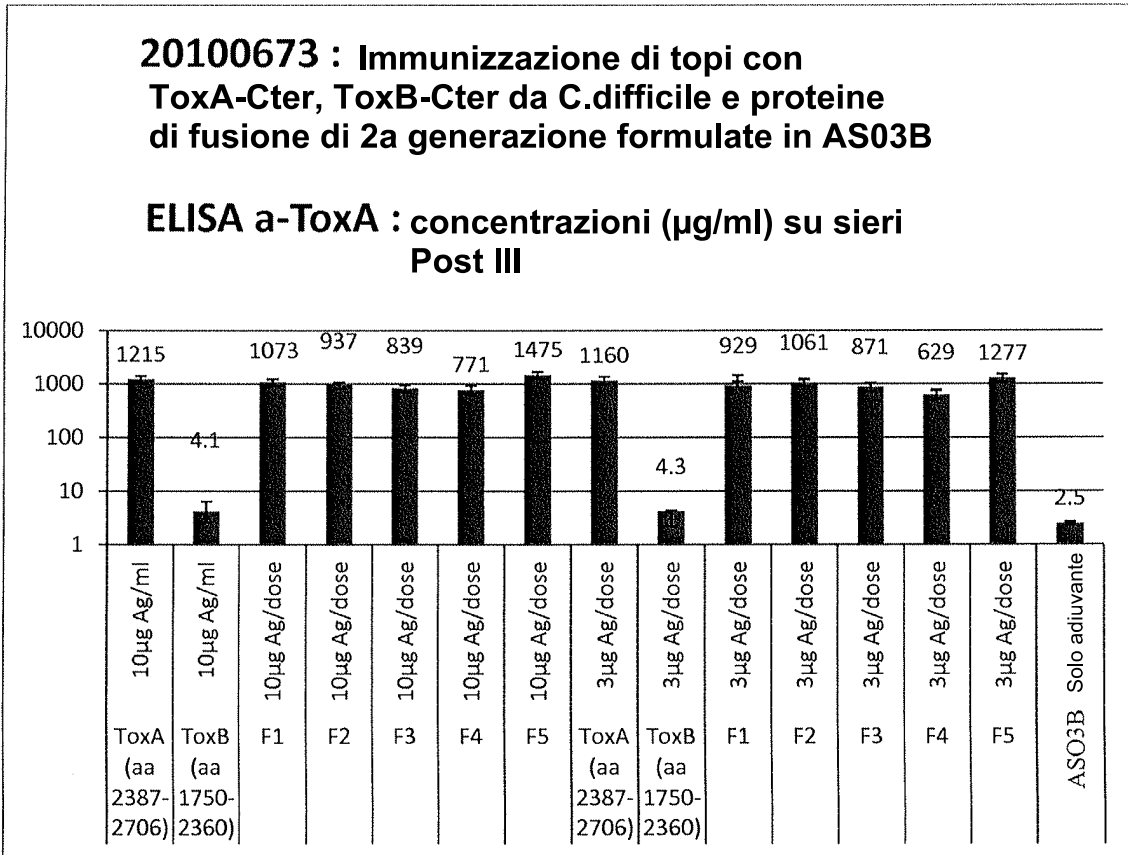


FIGURA 12

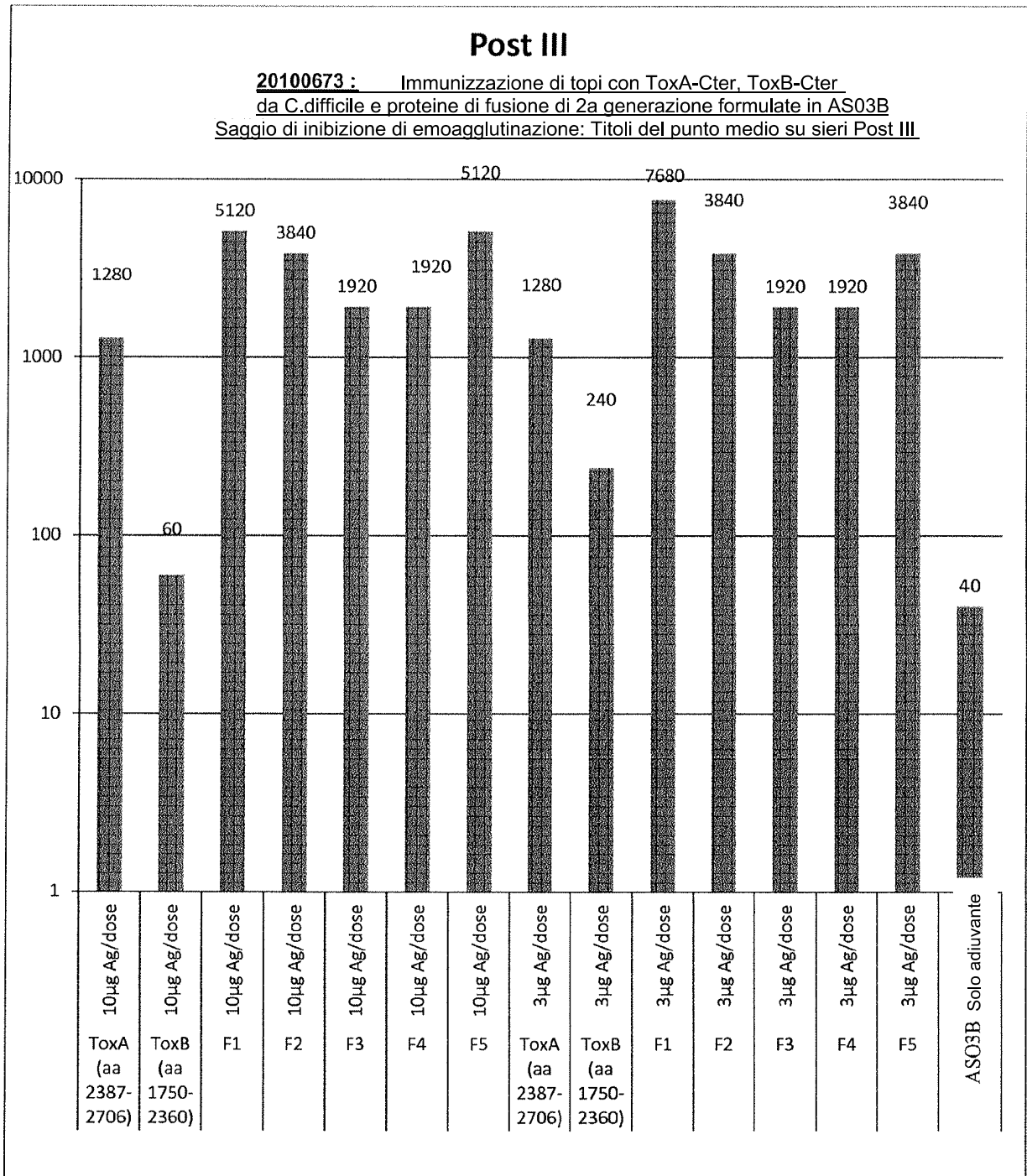


FIGURA 13

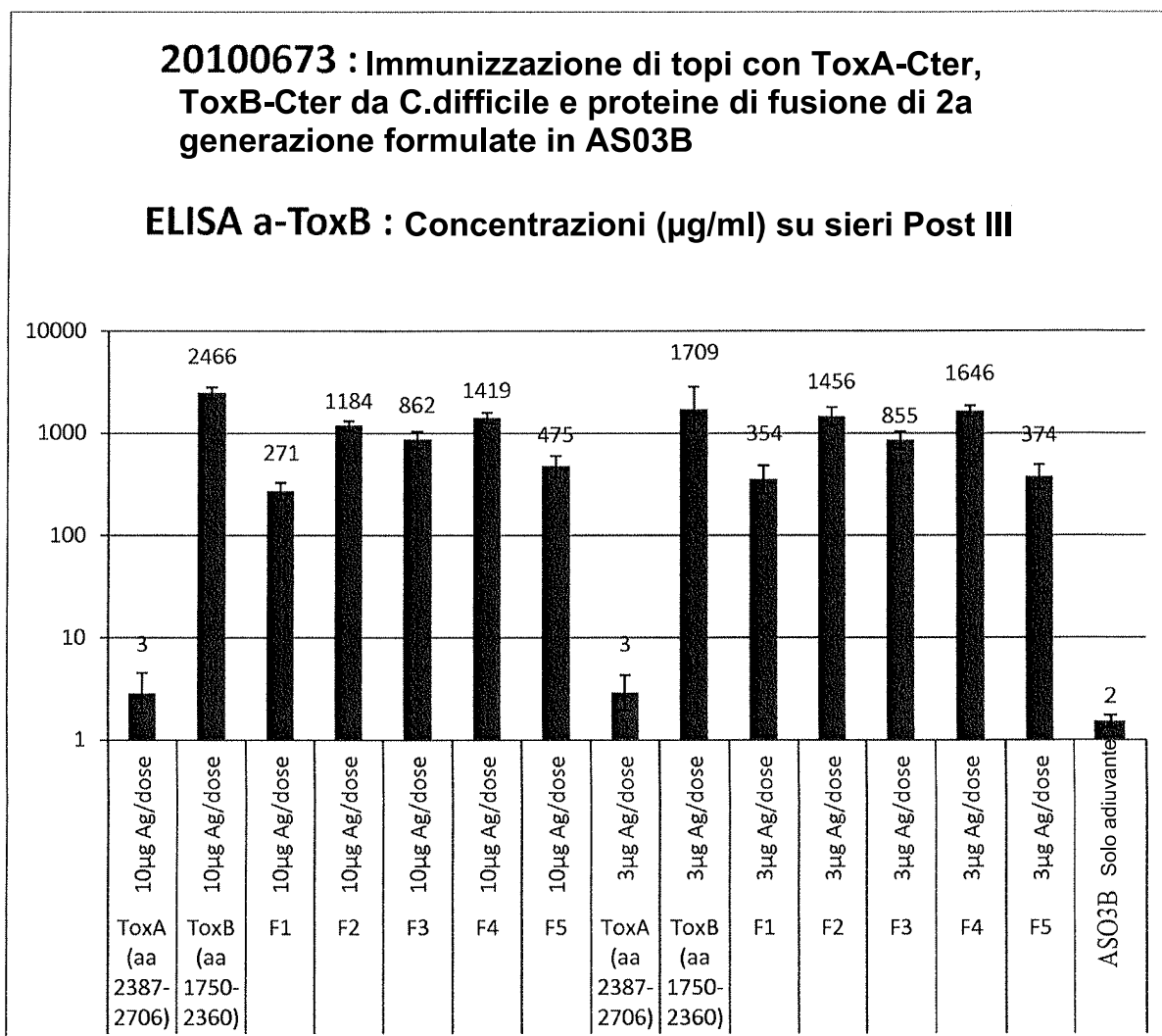


FIGURA 14

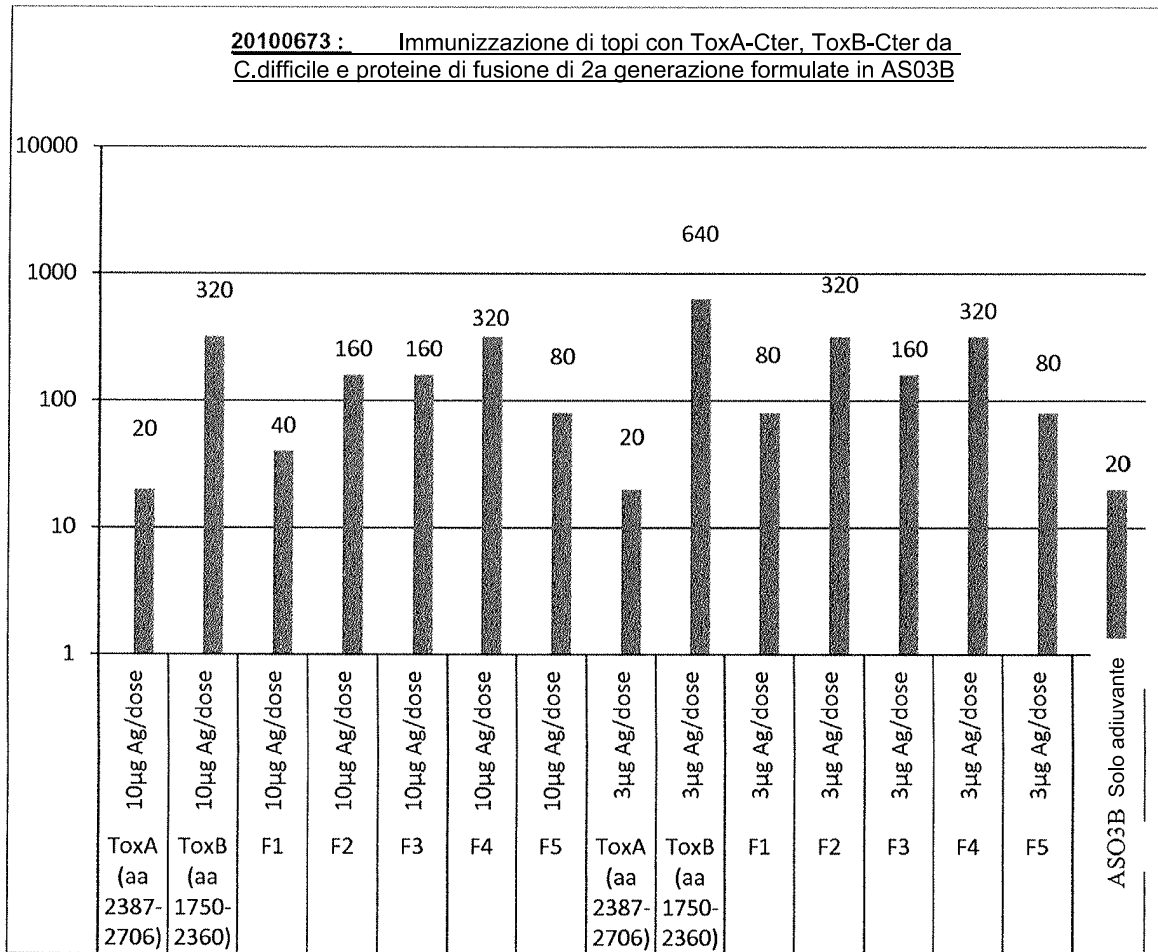
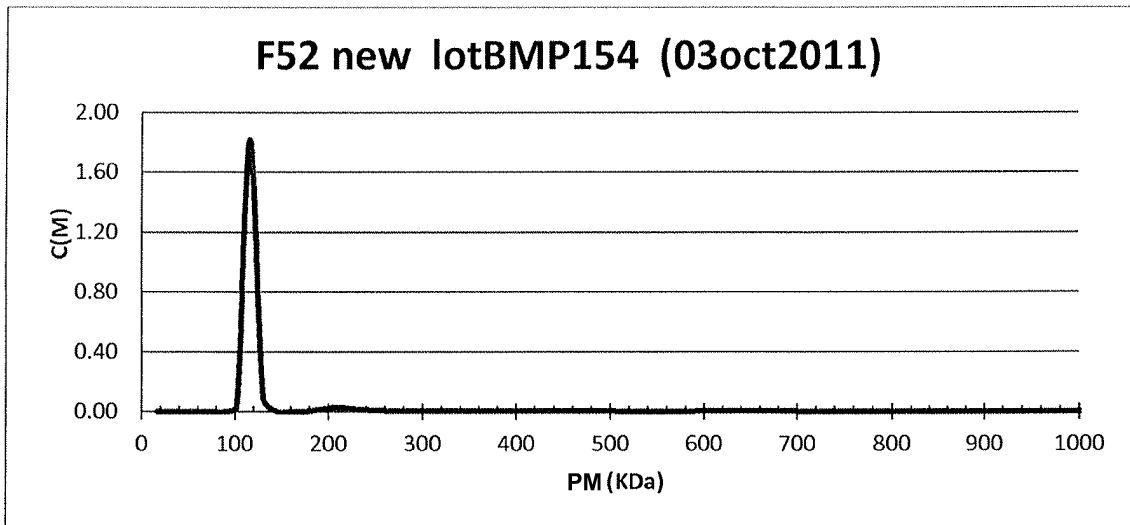
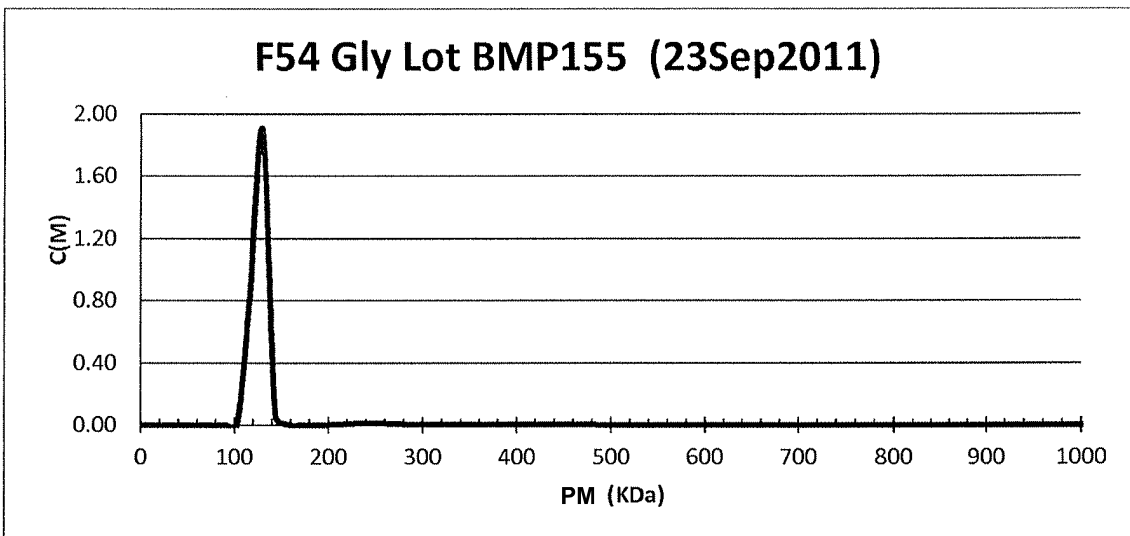


FIGURA 15

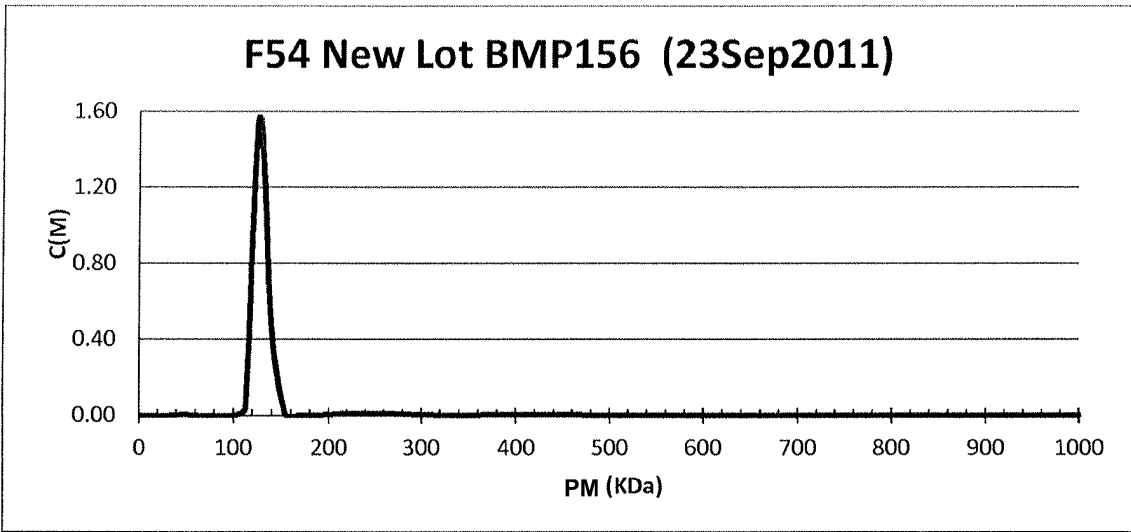
a)



b)



c)



d)

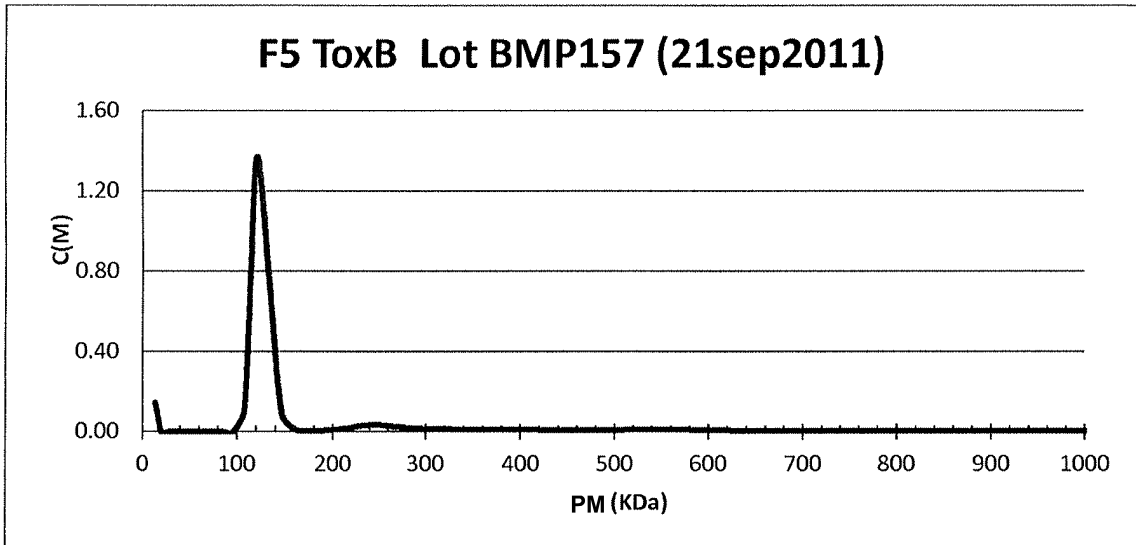


FIGURA 16

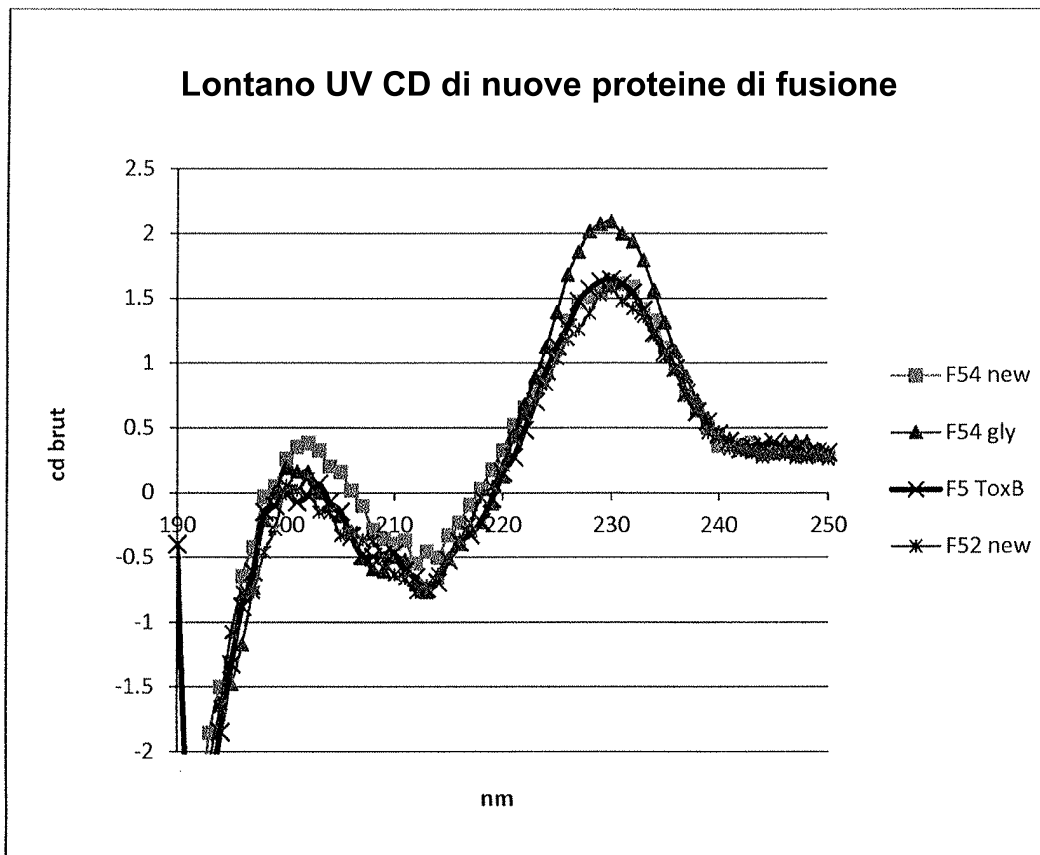


FIGURA 17

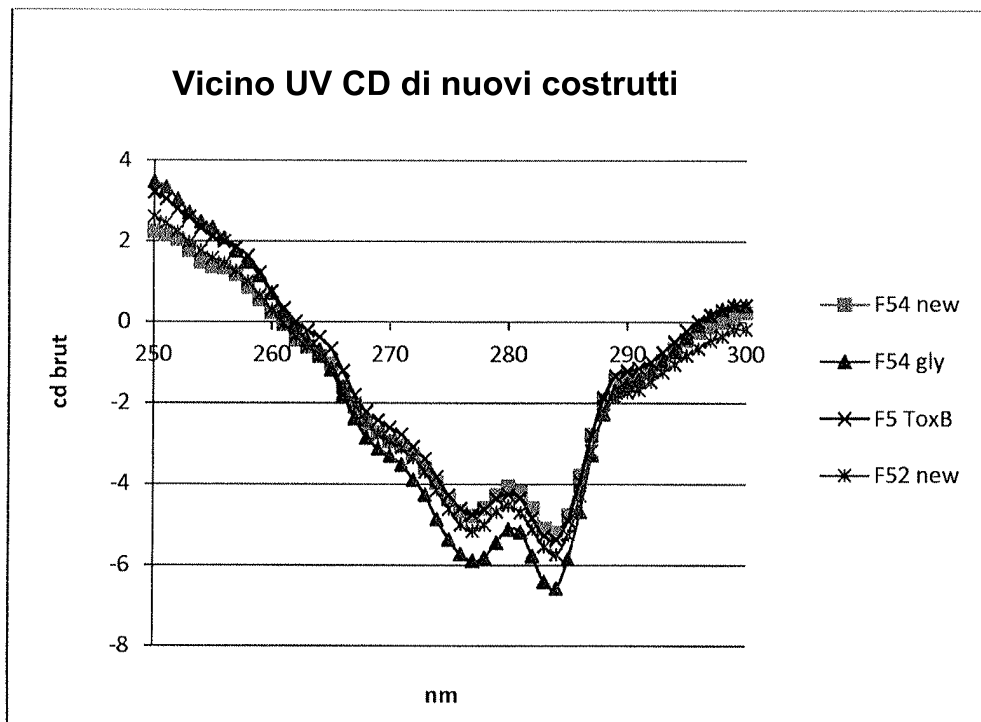


FIGURA 18

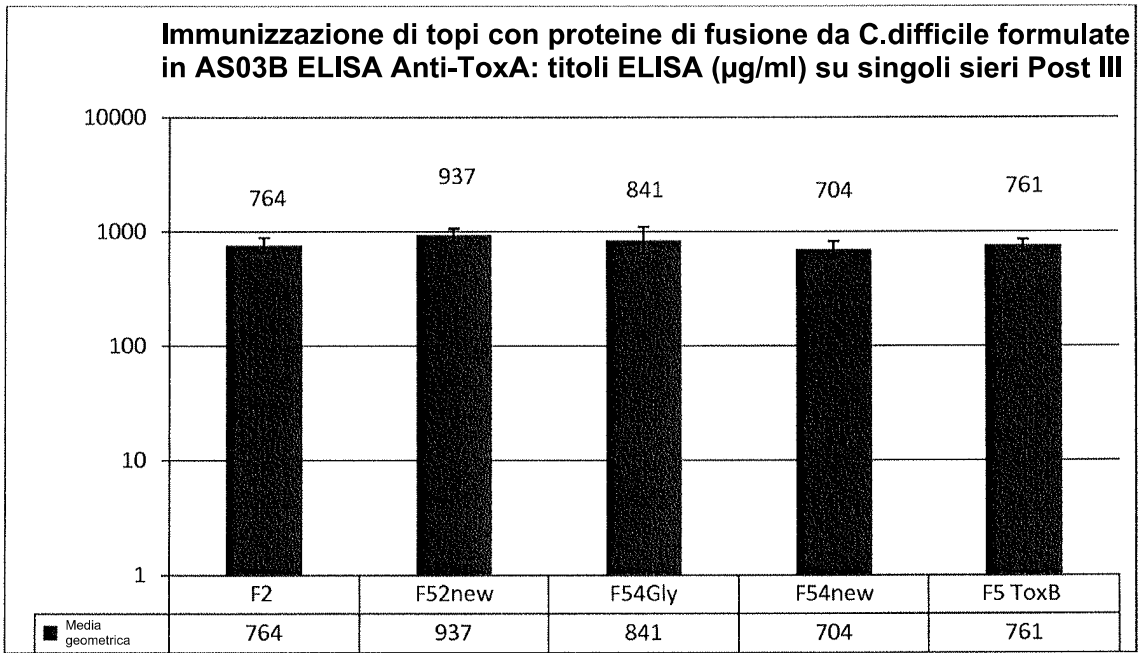


FIGURA 19

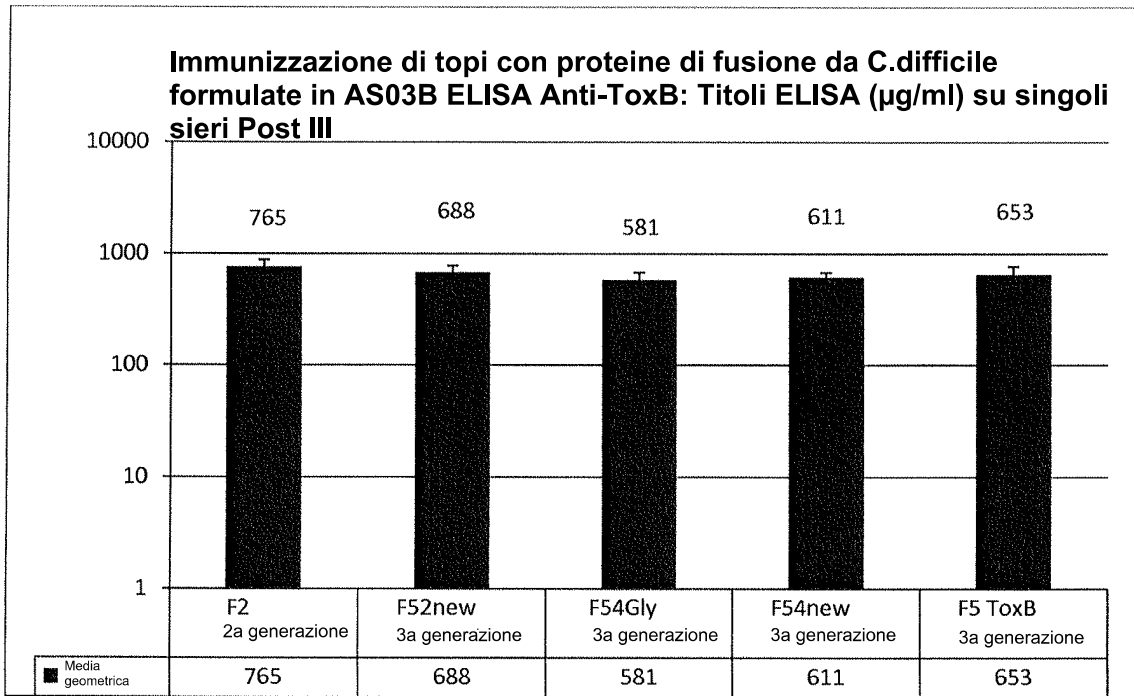


FIGURA 20

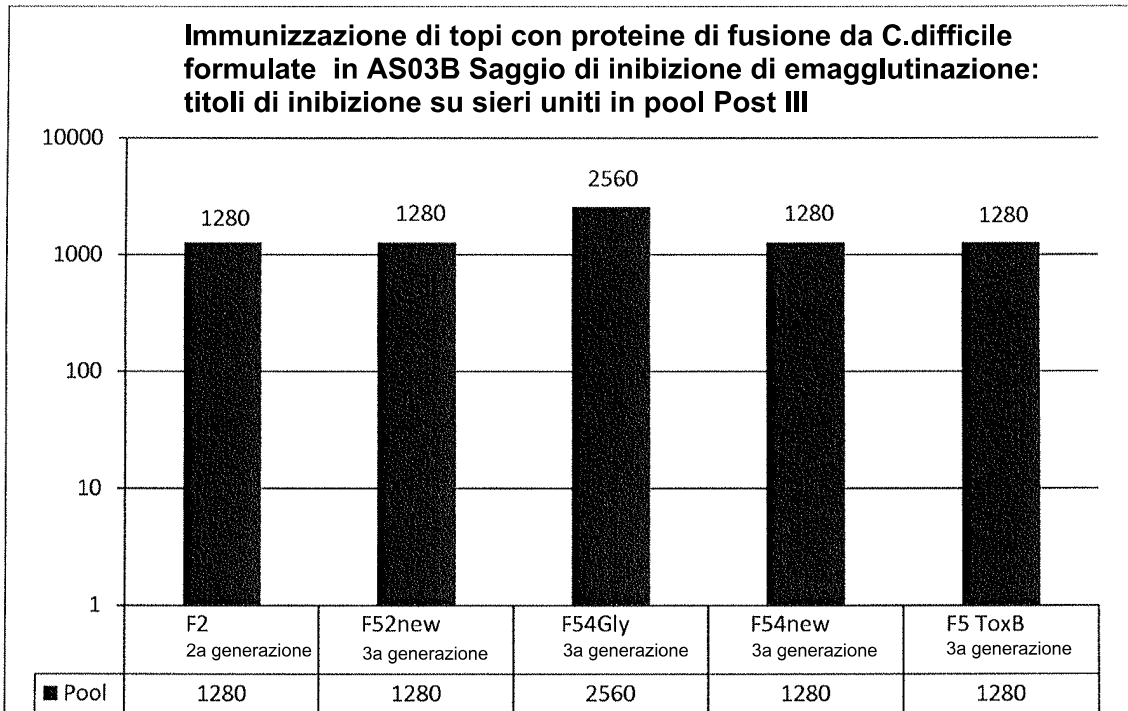


FIGURA 21

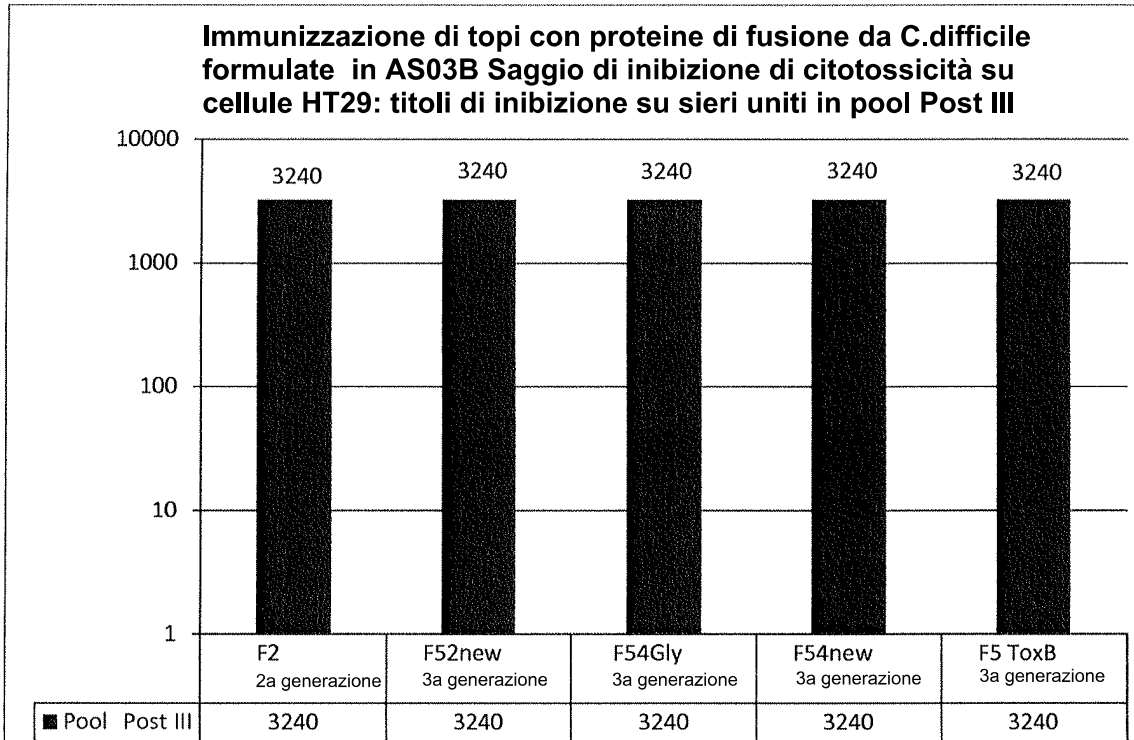


FIGURA 22

