IMMUNOGENIC COMPOSITION

TECHNICAL FIELD

The present invention relates to immunogenic compositions comprising a *Clostridium difficile (C. difficile)* polypeptide and an aluminum-free adjuvant. The invention also relates to vaccine compositions and the use of the vaccines and immunogenic compositions of the invention in prophylaxis or therapy or in the manufacture of a medicament.

BACKGROUND

C.difficile is the most important cause of nosocomial intestinal infections and is the major cause of pseudomembranous colitis in humans (Bartlett *et al* Am. J. Clin. Nutr. 11 suppl: 2521-6 (1980)). The overall associated mortality rate for individuals infected with C.difficile was calculated to be 5.99% within 3 months of diagnosis, with higher mortality associated with advanced age, being 13.5% in patients over 80 years (Karas *et al Journal of Infection* 561:1-9 (2010)).The current treatment for *C.difficile* infection is the administration of antibiotics (metronidazole and vancomycin), however there has been evidence of strains which are resistant to these antibiotics (Shah *et al*., Expert Rev. Anti Infect. Ther. 8(5), 555–564 (2010)). Accordingly there is a need for immunogenic compositions capable of inducing antibodies to, and/or a protective immune response to, *C.difficile*.

The enterotoxicity of *C.difficile* is primarily due to the action of two toxins, toxin A ('ToxA') and toxin B ('ToxB'). The C-terminal domains of toxin A and toxin B comprise repeating units, for example the C-terminal domain of toxin A is made up of contiguous repeating units (Dove *et al* Infect. Immun. 58:480-499 (1990)).For this reason the C-terminal domain may be referred to as the 'repeating domain'. These repeat portions can be separated further into short repeats (SRs) and long repeats (LRs) as described in Ho *et al* (PNAS 102:18373-18378 (2005)).

Immunogenic compositions comprising antigens from *C. difficile* have been described. WO96/12802 and WO00/61762 and Lyerly *et al* (Current Microbiology 21:29-32 (1990)) relate to fragments of toxin A, in particular fragments of the C-terminal domain, for inducing a protective immune response in hamsters. WO9920304 relates to a mixture of co-purified toxin A and toxin B inactivated by incubation in formaldehyde. WO00/61762 relates to immunogenic compositions comprising either the full length C-terminal domain or fragments of the C-terminal domain of toxin

A and toxin B of *C.difficile*. Similarly WO2012/163817, WO2012/163811 and WO2012/163810 provide disclosures of possible fragments of toxin A and toxin B as well as fusion proteins thereof.

New compositions or vaccines with improved immunogenicity are needed. As one strategy, adjuvants have been used to try and improve the immune response raised to any given antigen. For example, WO2009035707 describes a composition comprising a toxoid of *C. difficile* toxins A and B and an adjuvant such as aluminum hydroxide compound.

Adjuvants containing combinations of lipopolysaccharide and Quillaja saponins have been disclosed previously, for example in EP0671948. Oil in water emulsions *per se* are well known in the art, and have been suggested to be useful as adjuvant compositions (EP 399843; WO 95/17210).

There is still a need for vaccine and immunogenic compositions that provide a suitable immune response against *C.difficile*.

SUMMARY

In one aspect, the invention relates to an immunogenic composition comprising a polypeptide comprising a *Clostridium difficile (C. difficile)* toxin A fragment and / or a *C. difficile* toxin B fragment and an adjuvant comprising an immunologically active saponin fraction presented in the form of a liposome. The immunologically active saponin fraction may be QS21 and the lipopolysaccharide may be 3D-MPL. Suitably, the immunogenic composition may comprise about 50 µg 3D-MPL, and about 50 µg QS21.

In one aspect, the invention relates to an immunogenic composition comprising a polypeptide comprising a *C. difficile* toxin A fragment and / or a *C. difficile* toxin B fragment and an adjuvant comprising an oil in water emulsion, wherein said oil in water emulsion comprises a metabolisable oil, a tocol and an emulsifier. The metabolisable oil may be present at an amount of about 5.35 mg.The tocol may be present at an amount of about 5.94 mg. Suitably, the emulsifying agent may be present at an amount of about 2.425 mg . Suitably, the metabolisable oil is squalene, the tocol is alpha-tocopherol and the emulsifying agent is polyoxyethylene sorbitan monooleate.

Suitably, the immunogenic composition according to any aspect of the invention comprises a *C. difficile* toxin A repeating domain fragment and a *C. difficile* toxin B repeating domain fragment and elicits antibodies that neutralize toxin A and toxin B. In a further embodiment of the invention, the immunogenic composition comprises a *C.difficile* toxin A fragment and a *C. difficile* toxin B

fragment wherein one of the fragments is a N-terminal fragment and the other is a repeating domain fragment.

Suitably, the immunogenic composition according to any aspect of the invention comprises a polypeptide variant of 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 or SEQ ID NO:27.

Suitably, the immunogenic composition according to any aspect of the invention comprises a polypeptide comprising 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,SEQ ID NO:34 or SEQ ID NO:35 or a variant or fragment of any of these sequences.

The immunogenic compositions according to the invention may comprise additional antigens such as antigens derived from a bacterium selected from the group consisting of *Streptococcus pneumoniae (S. pneumoniae)*, *Haemophilus influenzae (H.influenzae)*, *Neisseria meningitidis (N.meningitidis)*, *Escherichia coli (E.coli)*, *Moraxella catarrhalis (M.cattarhalis)*, *Clostridium tetani (C. tetani), Corynebacterium diptherieriae (C. diptherieriae), Bordetella pertussis (B. pertussis), Staphylococcus epidermidis* (*S.epidermidis)*, enterococci, and *Staphylococcus aureus (S.aureus)*.

In one aspect, the invention relates to a vaccine comprising the immunogenic composition according to the invention and a pharmaceutically acceptable excipient.

In one aspect, the invention relates to a use of the immunogenic composition or the vaccine according to the invention in the treatment or prevention of *C.difficile* disease.

In one aspect, the invention relates to the immunogenic composition or the vaccine according to the invention for use in the treatment or prevention of *C.difficile* disease.

In one aspect, the invention relates to use of the immunogenic composition or the vaccine of according to the invention in the preparation of a medicament for the prevention or treatment of *C.difficile* disease.

In one aspect, the invention relates to a method of preventing or treating *C.difficile* disease comprising administering the immunogenic composition or the vaccine according to the invention to a subject.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 – Graphs showing anti-ToxA and anti-ToxB ELISA results for hamsters immunised with F2 formulated with Adjuvant A; Adjuvant A without QS21 and Adjuvant A without MPL.

FIG. 2 - Graphs showing anti-ToxA and anti-ToxB ELISA results for hamsters immunised with non-adjuvanted F2 or F2 formulated with Adjuvant A, Adjuvant B and Alum.

FIG. 3 – Graphs showing anti-ToxA and anti-ToxB ELISA results for Post I sera of male and female hamsters immunised with a dose-range of F2 fusion proteins formulated in Adjuvant B, AS04D, Adjuvant A or Adjuvant C.

FIG. 4 – Graphs showing anti-ToxA and anti-ToxB ELISA results for Post II sera of male and female hamsters immunised with a dose-range of F2 fusion proteins formulated in Adjuvant B, AS04D, Adjuvant A or Adjuvant C.

FIG. 5 – Graphs showing anti-ToxA and anti-ToxB ELISA results for Post III sera of male and female hamsters immunised with a dose-range of F2 fusion proteins formulated in Adjuvant B, AS04D, Adjuvant A or Adjuvant C.

FIG. 6 – Graphs showing anti-ToxA and anti-ToxB ELISA results for Post III day 87 sera of male and female hamsters immunised with a dose-range of F2 fusion proteins formulated in Adjuvant B, AS04D, Adjuvant A or Adjuvant C.

FIG. 7 – Graphs showing anti-ToxA and anti-ToxB ELISA results for male hamsters immunised with a dose-range of F2 fusion proteins formulated in Adjuvant B, AS04D, Adjuvant A or Adjuvant C on Post I, II and III sera.

FIG. 8 – Graphs showing anti-ToxA and anti-ToxB ELISA results for male and female hamsters immunised with a Mix (ToxA+ToxB) or with the F2 fusion protein formulated in Alum or Adjuvant A.

FIG. 9 – Graph showing anti-ToxA immunogenicity in mice immunised with a fragment of the C-terminus of toxin A (aa 2387-2706), a fragment of the C-terminus of toxin B (aa 1750-2360), or fusions 1, 2, 3, 4 or 5 formulated in Adjuvant B.

FIG. 10 - Graph showing hemagglutination inhibition in mice immunised with a fragment of the C-terminus of toxin A (aa 2387-2706), a fragment of the C-terminus of toxin B (aa 1750-2360), or fusions 1, 2, 3, 4 or 5 formulated in Adjuvant B.

FIG. 11 – Graph showing anti-ToxB immunogenicity in mice immunised with a fragment of the C-terminus of toxin A (aa 2387-2706), a fragment of the C-terminus of toxin B (aa 1750-2360), or fusions 1, 2, 3, 4 or 5 formulated in Adjuvant B.

FIG. 12 – Cyotoxicity inhibition titres from mice immunised with a fragment of the C-terminus of toxin A (aa 2387-2706), a fragment of the C-terminus of toxin B (aa 1750-2360), or fusions 1, 2, 3, 4 or 5 formulated in Adjuvant B.

FIG. 13 – Graph showing anti-ToxA ELISA results for mice immunised with the F2, F52New, F54Gly, G54New or F5 ToxB fusions formulated in Adjuvant B.

FIG. 14 – Graph showing anti-ToxB ELISA results for mice immunised with the F2, F52New, F54Gly, F54New or F5ToxB fusions formulated in Adjuvant B.

FIG. 15 – Graph showing hemagglutination inhibition in mice immunised with the F2, F52New, F54Gly, F54New or F5ToxB fusions formulated in Adjuvant B.

FIG. 16 – Graph showing cytotoxicity titres in HT29 cells from mice immunised with the F2, F52New, F54Gly, F54New or F5ToxB fusions formulated in Adjuvant B.

FIG.17 – Graph showing cytotoxicity titres in IMR90 cells from mice immunised with the F2, F52New, F54Gly, F54New or F5ToxB fusions formulated in Adjuvant B.

FIG. 18 – Graph showing anti-ToxA ELISA results for mice immunised with *C. difficile* ToxA-Cter (aa 2387-2706), ToxB-Cter (aa 1750-2360) and fusion proteins (non-adjuvanted).

FIG. 19 – Graph showing anti-ToxB ELISA results for mice immunised with *C. difficile* ToxA-Cter (aa 2387-2706),, ToxB-Cter (aa 1750-2360) and fusion proteins (non-adjuvanted)

FIG. 20 – Graph showing anti-ToxA ELISA results for mice immunised with *C. difficile* ToxA-Cter (aa 2387-2706), ToxB-Cter (aa 1750-2360) and fusion proteins formulated in Adjuvant B: Post II, III, Pre-Boost, Post IV sera.

FIG. 21 – Graph showing anti-ToxB ELISA results for mice immunised with *C. difficile* ToxA-Cter (aa 2387-2706), ToxB-Cter (aa 1750-2360)and fusion proteins formulated in Adjuvant B: Post II, III, Pre-Boost, Post IV sera.

FIG. 22 – Graph describing the Far-UV spectrum of Fusions, 2, 3, 4, and 5 measured using circular dichroism. The spectrum for fusion 2 is represented by a line with the points depicted as small squares, the spectrum for fusion 3 is represented by a line with the points depicted as small diamond shapes, fusion 4 is represented by a line with the points depicted as circles, and fusion 5 is represented by a line with the points depicted as cross shapes.

FIG. 23 - Graph describing the near-UV spectrum of Fusions 2, 3, 4, and 5 measured using circular dichroism. The spectrum for fusion 2 is represented by a line with the points depicted as cross shapes, the spectrum for fusion 3 is represented by a line with the points depicted as circles, the spectrum for fusion 4 is represented by a line with the points depicted as triangles, and the spectrum for fusion 5 is represented by a line with the points depicted as small diamond shapes.

FIG. 24 - Graph describing the Far-UV spectrum of fusions F52New, F54Gly, F54New and F5ToxB measured using circular dichroism. The spectrum for F52New is represented by a line

with the points depicted as double crosses, the spectrum for F54Gly is represented by a line with the points depicted as triangles, F54New is represented by a line with the points depicted as squares, and F5ToxB is represented by a line with the points depicted as cross shapes.

FIG. 25 - Graph describing the Near-UV spectrum of fusions F52New, F54Gly, F54New and F5ToxB measured using circular dichroism. The spectrum for F52New is represented by a line with the points depicted as double crosses, the spectrum for F54Gly is represented by a line with the points depicted as triangles, F54New is represented by a line with the points depicted as squares, and F5ToxB is represented by a line with the points depicted as cross shapes.

DETAILED DESCRIPTION

The inventors have shown immunogenic compositions comprising a polypeptide comprising a *C. difficile* toxin A fragment and/or a *C. difficile* toxin B fragment formulated with alum free adjuvants, particularly adjuvants such as Adjuvant A and Adjuvant B described herein, provide an improved immune response against *C.difficile* as compared to non-adjuvanted *C. difficile* polypeptide containing compositions. Immunogenic compositions according to the invention may also provide an improved immune response against *C.difficile* as compared to *C. difficile* polypeptides formulated with alum.

POLYPEPTIDES

The polypeptide comprising a *C. difficile* toxin A fragment and / or a *C. difficile* toxin B fragment may comprise purified toxin A or toxin B which have been inactivated, suitably by incubation in formaldehyde, such as those described in WO9920304.

The polypeptide may comprise a *C. difficile* toxin A N-terminal fragment and a *C. difficile* toxin B repeating domain fragment. The polypeptide may contain a *C. difficile* toxin A repeating domain fragment and a *C. difficile* toxin B N-terminal fragment.The polypeptide may comprise a *C. difficile* toxin A repeating domain fragment and a *C. difficile* toxin B repeating domain fragment.

The polypeptide may comprise either the full length N – terminal and/ or C-terminal domain of toxin A and toxin B of *C.difficile* or fragments thereof.

In one embodiment, the polypeptide in the immunogenic compositions according to the invention does not contain the full length N-terminal domain of toxin A and / or toxin B of *C.difficile*. In one embodiment, the polypeptide in the immunogenic compositions according to the invention does not contain an N-terminal domain fragment of toxin A and / or toxin B of *C.difficile*.

The polypeptide may be a fusion protein. For example, the fusion protein may comprise a fusion sequence such as Seq ID NOS: 36 and 37 or other sequences described in WO2012/028741.

The polypeptide may comprise a first fragment and a second fragment, wherein

(i) the first fragment is a toxin A repeating domain fragment;

(ii) the second fragment is a toxin B repeating domain fragment;

(iii) the first fragment comprises a first proximal end within a first repeat portion;

(iv) the second fragment comprises a second proximal end within a second repeat portion; and

wherein the first fragment and the second fragment are adjacent to one another and wherein the first repeat portion and the second repeat portion have structural similarity.

The term polypeptide refers to a contiguous sequence of amino acids.

The term 'toxin A repeating domain' refers to the C-terminal domain of the toxin A protein from *C.difficile* comprising repeated sequences. For example, the C-terminal domain of the toxin A protein may be amino acids 1832-2710 of toxin A from strain VPI10463 (ATCC43255) and/or their equivalent in a different strain. Amino acids 1832-2710 of toxin A from strain VPI10463 (ATCC43255) corresponds to amino acids 1832-2710 of SEQ ID NO:1.

The term 'toxin B repeating domain' refers to the C-terminal domain of the toxin B protein from *C.difficile*. For example, the C-terminal domain of the toxin B protein may be amino acids 1834-2366 of toxin B from strain VPI10463 (ATCC43255) and/or their equivalent in a different strain. Amino acids 1834-2366 of toxin B from strain VPI10463 (ATCC43255) corresponds to amino acids 1834-2366 of SEQ ID NO:2.

The full length N-terminal domain of toxin A may be amino acids 1-542 of toxin A from strain VPI10463 (ATCC43255) and/or their equivalent in a different strain. Thus the term "a *C. difficile* toxin A N-terminal fragment" or "an N-terminal domain fragment of toxin A" refers to a fragment of amino acids 1-542 of toxin A from strain VPI10463 (ATCC43255) and/or their equivalent in a different strain. Amino acids 1-542 of toxin A from strain VPI10463 (ATCC43255) corresponds to amino acids 1-542 of SEQ ID NO: 1.

The full length N-terminal domain of toxin B may be amino acids 1-543 of toxin B from strain VPI10463 (ATCC43255) and/or their equivalent in a different strain. Thus the term "a *C. difficile* toxin B N-terminal fragment" or "an N-terminal domain fragment of toxin B" refers to a fragment of amino acids 1-543 of toxin B from strain VPI10463 (ATCC43255) and/or their equivalent in a different strain. Amino acids 1-543 of toxin B from strain VPI10463 (ATCC43255) corresponds to amino acids 1-543 of SEQ ID NO: 2.

C.difficile toxins A and B are conserved proteins. However the sequence differs a small amount between strains. Moreover the amino acid sequence for toxins A and B in different strains may differ in number of amino acids.

In one embodiment, the toxin A repeating domain and/or toxin B repeating domain may be a sequence which is a variant with at least 90%, 95%, 98%, 99% or 100% sequence identity to amino acids 1832-2710 of SEQ ID NO:1 or a variant with at least 90%, 95%, 98%, 99% or 100% sequence identity to amino acids 1834-2366 of SEQ ID NO:2.

In one embodiment, the toxin A N-terminal domain may be a sequence which is a variant with at least 90%, 95%, 98%, 99% or 100% sequence identity to amino acids 1-542 of SEQ ID NO:1. In one embodiment, the toxin B N-terminal domain may be a sequence which is a variant with at least 90%, 95%, 98%, 99% or 100% sequence identity to amino acids 1-543 of SEQ ID NO: 2.

A 'variant' is a polypeptide that varies from the referent polypeptides by conservative amino acid substitutions, whereby a residue is substituted by another with the same physico-chemical properties. Typically such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln, and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr.

Furthermore the amino acid numbering may differ between the C-terminal domains or N-terminal domains of toxin A (or toxin B) from one strain and toxin A (or toxin B) from another strain. For this reason the term 'equivalent in a different strain' refers to amino acids which correspond those of a reference strain (e.g., *C. difficile* VPI10463), but which are found in a toxin from a different strain and which may thus be numbered differently. A region of 'equivalent' amino acids may be determined by aligning the sequences of the toxins from the different strains. The amino acids numbers provided throughout refer to those of strain VPI10463.

The term 'fragment' of a polypeptide or protein refers to a contiguous portion, such as at least 100, 150, 180, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 or 600 amino acids, from that polypeptide or protein.

"N-terminal domain fragment of toxin A" of *C.difficile* refers to a contiguous portion, such as at least 100, 150, 180, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500 or 530 amino acids, from the full length N-terminal domain of toxin A. "N-terminal domain fragment of toxin B" of *C.difficile* refers to a contiguous portion, such as at least 100, 150, 180, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500 or 530 amino acids, from the full length N-terminal domain of toxin B.

Toxin A repeating domain fragment may be a contiguous portion of at least 100, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 or 600 amino acids from the toxin A repeating domain. Toxin A repeating domain fragment may be a contiguous portion of less than 750, less than 700, less than 650, less than 600 or less than 580 amino acids. Toxin B repeating domain fragment may be a contiguous portion of at least 100, 200, 230, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 or 600 amino acids from the toxin B repeating domain. Toxin B repeating domain fragment may be a contiguous portion of less than 530, less than 500, less than 480, or less than 450 amino acids.

The term 'first fragment' refers to a contiguous portion of at least 100, such as 150, 180, 250, 300, 350, 380, 400, 450, 480, 500, 530, 550, 580 or 600, amino acids of the toxin A repeating domain. The term 'second fragment' refers to a contiguous portion of at least 100, such as 200, 230, 250, 280, 300, 350, 400, 450 or 500, amino acids of the toxin B repeating domain.

The term 'first proximal end' refers to the end of the first fragment (Tox A fragment) which is covalently linked to the second fragment (Tox B fragment) or covalently linked to a linker sequence between the first and second fragment. The term 'second proximal end' refers to the end of the second fragment which is closest to the first fragment in primary structure (amino acid sequence).

The C-terminal domain of toxin A is made up of 8 repeat portions (designated repeat portion I, repeat portion II, repeat portion III, repeat portion IV, repeat portion V, repeat portion VI, repeat portion VII and repeat portion VIII). Each of these repeat portions can be further divided into short repeats (SRs) and long repeats (LRs) (except for Tox A repeat portion VIII which does not have a long repeat). Each of the long repeats has some structural and sequence similarity to the other long repeats. Similarly the short repeats have some sequence and structural similarity to one another. Similarly the C-terminal domain of toxin B is made up of 5 repeat portions subdivided into SRs and LRs. Each repeat portion contains one LR and between 2 and 5 SRs (except for Tox B repeat portion V which does not have a long repeat). For the purposes of the disclosure 'a'repeat portion' refers to one of the eight repeat portions of ToxA (designated I, II, III, IV, V, VI, VII or VIII) or one of the five repeat portions of ToxB (designated I, II, III, IV or V) or a partial repeat portion from the toxin A or toxin B repeating domain. As used herein the term 'first repeat portion' refers to a repeat portion (or partial repeat portion) from the toxin A repeating domain. The term 'second repeat portion' refers to a repeat portion (or partial repeat portion) from the toxin B repeating domain.

For example, repeat portion I of ToxA contains three SRs and one LR, which can be referred to as the first SRI of ToxA, the second SRI of ToxA, the third SRI of ToxA and the LRI of ToxA, respectively.

The first proximal end is considered to be within a 'repeat portion' if the first fragment ends in an amino acid that is within that repeat portion (i.e., the first proximal end contains only part of the repeat portion sequence). Similarly the second proximal end is considered to be within a 'repeat portion' if the second fragment ends in an amino acid that is within that repeat portion. For example the first proximal end is within 'repeat portion I of ToxA if the first fragment ends with any one of amino acids 1833-1923 (inclusive) of VPI10463 or their equivalent in another strain. The first proximal end is within a 'long repeat' or a 'short repeat' if the first fragment ends in an amino acid that is within a 'long repeat' or a 'short repeat', similarly the second proximal end is within a

'long repeat' or a 'short repeat' if the second fragment ends in an amino acid that is within a 'long repeat' or a 'short repeat'.

The amino acid positions of each domain has been defined for toxin A and toxin B from strain VPI10463 (ATCC43255). These are as follows (Table 1)

Name		Start position	End position
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

In one embodiment, the repeat portion of toxin A refers to amino acids 1832-1924, 1925-2058, 2059-2192, 2193-2306, 2307-2440, 2441-2553, 2554-2644 or 2645-2710 of toxin A (SEQ ID NO:1) or an equivalent in a different strain of *C. difficile*. In another embodiment, the repeat portion of toxin B refers to amino acids 1834-1926, 1927-2057, 2058-2189, 2190-2323 or 2324-2366 of toxin B (SEQ ID NO:2) or an equivalent in a different strain of *C.difficile*.

The term 'short repeat' may refer to amino acids 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 or 2687-2710 of toxin A (SEQ ID NO:1) or amino acids 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 or 2344-2366 of toxin B (SEQ ID NO:2) or their equivalents in a different strain of *C.difficile*.

Similarly the term 'long repeat' may refer to amino acids 1894-1924, 2028-2058, 2162-2192, 2276-2306, 2410-2440, 2523-2553 or 2614-2644 of toxin A (SEQ ID NO:1) or amino acids 1897-1926, 2028-2057, 2160-2189 or 2294-2323 of toxin B (SEQ ID NO:2) or their equivalents in a different strain of *C.difficile*.

The polypeptides of the invention may be part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.

The word 'adjacent' means separated by less than or exactly 20, 15, 10, 8, 5, 2, 1 or 0 amino acids in the primary structure.

The fragments may be positioned such that the N-terminus of the first fragment is adjacent to the C-terminus of the second fragment, alternatively the C-terminus of the first fragment may be adjacent to the N-terminus of the second fragment, or the C-terminus of the first fragment may be adjacent to the C-terminus of the second fragment, or the N-terminus of the first fragment may be adjacent to the N-terminus of the second fragment.

Two sequences will have 'sequence similarity to one another' if they have greater than 30%, 50%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% or 100% sequence identity.

The term 'identity' is known in the art. Identity is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described 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; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the Needle program, BLASTP, BLASTN (Altschul, S.F. et al., J. Molec. Biol. 215: 403-410 (1990)), and FASTA Pearson and Lipman Proc. Natl. Acad. Sci. USA 85; 2444-2448 (1988)). The BLAST family of programs is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215: 403-410 (1990)). The well known Smith Waterman algorithm may also be used to determine identity.

In one example, parameters for polypeptide sequence comparison include the following:

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

Comparison matrix: BLOSSUM62 from Henikoff and Henikoff,

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

Gap Penalty: 10

Gap extension penalty:0.5

A program useful with these parameters is publicly available as the 'Needle' program from EMBOSS package (Rice P.et al, Trends in Genetics 2000 col.16(6):276-277). The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

In one embodiment the first repeat portion and the second repeat portion have high structural similarity to one another. Two sequences can be considered to have high structural similarity when their sequence identity is higher than 40%, 45%, 50% or 60%. The presence of high sequence identity can be determined by using softwares that allow sequences alignment (EMBOSS, Blast, …). 3D model structure could be faithfully built, using the SwissModel web server, for a protein if it shares more than 40%, 45%, 50% or 60% sequence identity with a protein having a known 3D structure.

In one embodiment the polypeptide of the invention elicits antibodies that neutralise toxin A or toxin B. In a further embodiment the polypeptide elicits antibodies that neutralise toxin A. In a further embodiment the polypeptide elicits antibodies that neutralise toxin B. In a further embodiment the polypeptide elicits antibodies that neutralise toxin A and toxin B. The phrase 'elicits neutralising antibodies' means that the when the compositions are used to immunise a mammal, for example a mouse, a guinea pig or a human, the mammal generates neutralising antibodies.

Whether a polypeptide elicits neutralizing antibodies against a toxin can be measured by immunising mice with an immunogenic composition comprising the polypeptide, collecting sera and analysing the anti-toxin titres of the sera using by enzyme-linked immunosorbent assay (ELISA). The sera could be compared to a reference sample obtained from mice which have not been immunised. An example of this technique can be found in example 6. The polypeptide of the invention elicits antibodies that neutralise toxin A if the sera against the polypeptide gives an ELISA readout more than 10%, 20%, 30%, 50%, 70%, 80%, 90% or 100% higher than the reference sample.

In a further embodiment the polypeptide of the invention elicits a protective immune response in a mammalian host against strains of *C.difficile*. The phrase 'elicit a protective immune response'

means that when the immunogenic composition of the invention is used to immunise a mammal such as a mouse, guinea pig or human, the mammal generates antibodies capable of protecting the mammal from death caused by *C.difficile*. In one embodiment the mammalian host is selected from the group consisting of mouse, rabbit, guinea pig, monkey, non-human primate or human. In one embodiment the mammalian host is a mouse. In a further embodiment the mammalian host is a human.

Whether a polypeptide elicits a protective immune response in a mammalian host against strains of *C.difficile* can be determined using a challenge assay. In such an assay the mammalian host is vaccinated with the polypeptide and challenged by exposure to *C.difficile*. The time which the mammal survives after challenge is compared with the time which a reference mammal that has not been immunised with the polypeptide survives. A polypeptide elicits a protective immune response if a mammal immunised with the polypeptide survives at least 10%, 20%, 30%, 50%, 70%, 80%, 90%, or 100% longer after challenge with *C.difficile* than a reference mammal which has not been immunised with the polypeptide. In one embodiment the polypeptide of the invention elicits a protective immune response against *C.difficile* in a mammal selected from the group consisting of mouse, guinea pig, monkey or human. In one embodiment the mammal is a mouse, in a further embodiment the mammal is a human.

The native structure of the C-terminal (repeat) domains from toxins A and B consist of an extended β solenoid-like structure. This structure consists of primarily β sheet structures, with a minority of α helical structures as seen in Ho et al (PNAS 102:18373-18378 (2005)). The secondary structures present can be determined using circular dichroism (CD). For example, the secondary structure may be determined by measuring the shape and the magnitude of the CD spectra in the far-UV region (190-250nm) and comparing the results with those of known structures. This can be carried out using an optical path of 0.01cm from 178 to 250nm, with a 1nm resolution and bandwidth on a Jasco J-720 spectropolarimeter, for example as seen in example 5 below.

In one embodiment the first fragment comprises less than 28%, 25%, 23%, 20%, 18%, 15%, 10%, or 7% alpha helical secondary structure. In one embodiment the second fragment comprises less than 28%, 25%, 23%, 20%, 18%, 15%, 10%, or 7% alpha helical secondary structure. In a further embodiment both the first fragment and the second fragment comprise less than 28%, 25%, 23%, 20%, 18%, 15%, 10%, or 7% alpha helical secondary structure.

In one embodiment the first fragment comprises more than 20%, 25%, 28%, 30%, 33%, 35%, 38%, 40%, or 42% beta sheet structure. In one embodiment the second fragment comprises more than 20%, 25%, 28%, 30%, 33%, 35%, 38%, 40%, or 42% beta sheet structure. In a further embodiment both the first fragment and the second fragment comprises more than 20%, 25%, 28%, 30%, 33%, 35%, 38%, 40%, or 42% beta sheet structure.

In one embodiment the first proximal end is within repeat portion V (amino acids 2307-2440 of SEQ ID NO:1 or their equivalent in a different strain), repeat portion VI (amino acids 2441-2553 of SEQ ID NO:1 or their equivalent in a different strain), repeat portion VII (amino acids 2554-2644 of SEQ ID NO:1 or their equivalent in a different strain) or repeat portion VIII (amino acids 2645-2710 of SEQ ID NO:1 or their equivalent in a different strain) of toxin A. In a further embodiment the first proximal end is within repeat portion VII (amino acids 2554-2644 of SEQ ID NO:1 or their equivalent in a different strain) of toxin A. In a further embodiment the first proximal end repeat portion VIII (amino acids 2645-2710 of SEQ ID NO:1 or their equivalent in a different strain) of toxin A.

In one embodiment the second proximal end is within repeat portion I (amino acids 1834-1926 of SEQ ID NO:2 or their equivalent in a different strain), repeat portion II (amino acids 1927-2057 of SEQ ID NO:2 or their equivalent in a different strain), or repeat portion III (amino acids 2058-2189 of SEQ ID NO:2 or their equivalent in a different strain) of toxin B. In a further embodiment the second proximal end is within repeat portion II (amino acids 1927-2057 of SEQ ID NO:2 or their equivalent in a different strain) of toxin B. In a further embodiment the second proximal end is within repeat portion I (amino acids 1834-1926 of SEQ ID NO:2 or their equivalent in a different strain) of toxin B.

In one embodiment the first proximal end is within a long repeat. The first proximal end may be within long repeat V of toxin A (amino acids 2410-2440 of SEQ ID NO:1 or their equivalent in a different strain), or within long repeat VI of toxin A (amino acids 2523-2553 of SEQ ID NO:1 or their equivalent in a different strain), or within long repeat VII of toxin A (amino acids 2614-2644 of SEQ ID NO:1 or their equivalent in a different strain).

In one embodiment the second proximal end is within a long repeat. The second proximal end may be within long repeat I of toxin B (amino acids 1897-1926 of SEQ ID NO:2 or their equivalent in a different strain), or within long repeat II of toxin B (amino acids 2028-2057 of SEQ ID NO:2 or

their equivalent in a different strain), or within long repeat III of toxin B (amino acids 2160-2189 of SEQ ID NO:2 or their equivalent in a different strain).

In a further embodiment the first proximal end and the second proximal end are both within long repeats. In one embodiment the first proximal end is within long repeat V of toxin A (amino acids 2410-2440 of SEQ ID NO:1 or their equivalent in a different strain), or within long repeat VI of toxin A (amino acids 2523-2553 of SEQ ID NO:1 or their equivalent in a different strain), or within long repeat VII of toxin A (amino acids 2614-2644 of SEQ ID NO:1 or their equivalent in a different strain) and the second proximal end is within long repeatlof toxin B (amino acids 1897-1926 of SEQ ID NO:2 or their equivalent in a different strain), or within long repeat II of toxin B (amino acids 2028-2057 of SEQ ID NO:2 or their equivalent in a different strain), or within long repeat III of toxin B (amino acids 2160-2189 of SEQ ID NO:2 or their equivalent in a different strain). In one embodiment the first proximal end is within long repeat VII of toxin A (amino acids 2614-2644 of SEQ ID NO:1 or their equivalent in a different strain) and the second proximal end is within long repeat II of toxin B (amino acids 2028-2057 of SEQ ID NO:2 or their equivalent in a different strain).

In one embodiment the first proximal end is within amino acids 2620-2660 of toxin A. In one embodiment the second proximal end is within amino acids 2030-2050 of toxin B. In a further embodiment the first proximal end is within amino acids 2620-2660 of toxin A and the second proximal end is within amino acids 2030-2050 of toxin B.

In one embodiment the first fragment comprises at least 100, 150, 180, 200, 240, 250, 280, 300, 330, 350, 380, 400, 430, 450, 480, 500 or 530 amino acids. In one embodiment the second fragment comprises at least 100, 130, 150, 180, 200, 230, 250, 270, 300, 330, 350, 390, or 400 amino acids.

In one embodiment the polypeptide further comprises a linker. This linker may be between the first proximal end and the second proximal end, alternatively the linker may link the distal ends of the first fragment and/or the second fragment to a further sequence of amino acids.

A peptide linker sequence may be employed to separate the first fragment and second fragment. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following

factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first fragment and/or the second fragments; and (3) the lack of hydrophobic or charged residues that might react with the ToxA and/or ToxB functional epitopes. Peptide linker sequences may contain Glycine(Gly), Asparagine (Asn) and Serine (Ser) residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea *et al., Gene* 40:39-46 (1985); Murphy *et al., Proc. Natl. Acad. Sci. USA* 83:8258-8262 (1986); U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length.

In one embodiment the linker comprises 1-20, 1-15, 1-10, 1-5, 1-2, 5-20, 5-15, 5-15, 10-20, or 10-15 amino acids. In one embodiment the linker is a glycine linker, the linker may comprise multiple contiguous glycine residues (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, or 20), or alternatively the linker may comprise some glycine residues and some residues of other amino acids such as alanine. In a further embodiment the linker comprises a single glycine residue.

In an embodiment the polypeptide of the invention is part of a larger fusion protein. The fusion proteins may further comprise amino acids encoding an immunogenic portion of a further protein antigen. For example the fusion protein may further comprise an immunogenic portion of a protein antigen obtained or derived from a bacterium selected from the group consisting of *S.pneumoniae*, *H.influenzae*, *N.meningitidis*, *E.coli*, *M.cattarhalis*, *C.tetanii*, *C.diphtheriae*, *B.pertussis*, *S.epidermidis*, enterococci, *S.aureus,* and *Pseudomonas aeruginosa*. In this case the linker may be between t*he first frag*ment or the second fragment and a further immunogenic portion of a protein antigen.

The term "immunogenic portion thereof" or 'immunogenic fragment' refers to a fragment of a polypeptide wherein the fragment comprises an epitope that is recognized by cytotoxic T lymphocytes, helper T lymphocytes or B cells. Suitably, the immunogenic portion will comprise at least 30%, suitably at least 50%, especially at least 75% and in particular at least 90% (e.g. 95% or 98%) of the amino acids in the reference sequence. In one embodiment, the immunogenic portion will comprise all of the epitope regions of the reference sequence.

Immunogenic compositions according to the invention comprising the specific fusion proteins of *C. difficile* toxin A repeating domain fragment and a *C. difficile* toxin B repeating domain fragment

described herein are also expected to provide an improved immune response against *C.difficile* as compared to compositions comprising known *C. difficile* polypeptides or fragments.

ADJUVANTS

The immunogenic compositions disclosed herein include at least one C.*difficile* polypeptide in combination with an aluminum-free adjuvant. The polypeptides are formulated with an adjuvant that is free of aluminum or aluminum salts, that is, an aluminum-free adjuvant or adjuvant system.

In certain embodiments, the *C. difficile* polypeptide is formulated with an adjuvant comprising an immunologically active saponin fraction presented in the form of a liposome. The adjuvant may further comprise a lipopolysaccharide. The adjuvant may include QS21. For example, in one embodiment, the adjuvant contains QS21 in a liposomal formulation. In one embodiment, the adjuvant system includes 3D-MPL and QS21. For example, in one embodiment, the adjuvant contains 3D-MPL and QS21 in a liposomal formulation. Optionally, the adjuvant system also contains cholesterol. In one specific embodiment, the adjuvant includes QS21 and cholesterol. Optionally, the adjuvant system contains 1, 2-Dioleoyl-sn-Glycero-3-phosphocholine (DOPC). For example, in one specific adjuvant system contains cholesterol, DOPC, 3D-MPL and QS21.

In one specific example, the immunogenic composition includes an adjuvant formulated in a dose that includes: from about 0.1 to about 0.5 mg cholesterol; from about 0.25 to about 2 mg DOPC; from about 10 µg to about 100 µg 3D-MPL; and from about 10 µg to about 100 µg QS21. In a further specific example the immunogenic composition includes and adjuvant formulated in a dose that includes: from about 0.1 to about 0.5mg cholesterol, from about 0.25 to about 2mg DOPC, from about 10µg to about 100µg 3D-MPL, and from about 10µg to about 100µg QS21. In a further specific formulation the adjuvant is formulated in a dose comprising from about 0.1 to about 0.5mg cholesterol, from about 0.25 to about 2mg DOPC, from about 10µg to about 70µg 3D-MPL, and from about 10µg to about 70µg QS21. In one specific formulation, the adjuvant is formulated in a single dose that contains: about 0.25 mg cholesterol; about 1.0 mg DOPC; about 50 µg 3D-MPL; and about 50 µg QS21. In other embodiments, the immunogenic composition is formulated with a fractional dose (that is a dose, which is a fraction of the preceding single dose formulations, such as one half of the preceding quantity of components (cholesterol, DOPC, 3D-MPL and QS21), ¼ of the preceding quantity of components, or another fractional dose (e.g., 1/3, 1/6, etc.) of the preceding quantity of components.

In one embodiment, the immunogenic compositions according to the invention include an adjuvant containing combinations of lipopolysaccharide and Quillaja saponins that have been

disclosed previously, for example in EP0671948. This patent demonstrated a strong synergy when a lipopolysaccharide (3D-MPL) was combined with a Quillaja saponin (QS21).

The adjuvant may further comprise immunostimulatory oligonucleotides (for example, CpG) or a carrier.

A particularly suitable saponin for use in the present invention is Quil A and its derivatives. Quil A is a saponin preparation isolated from the South American tree *Quillaja Saponaria Molina* and was first described by Dalsgaard *et al*. in 1974 ("Saponin adjuvants", Archiv. für die gesamte Virusforschung, Vol. 44, Springer Verlag, Berlin, p243-254) to have adjuvant activity. Purified fragments of Quil A have been isolated by HPLC which retain adjuvant activity without the toxicity associated with Quil A (EP 0 362 278), for example QS7 and QS21 (also known as QA7 and QA21). QS21 is a natural saponin derived from the bark of *Quillaja saponaria* Molina, which induces CD8+ cytotoxic T cells (CTLs), Th1 cells and a predominant IgG2a antibody response and is a preferred saponin in the context of the present invention.

In a specific embodiment, QS21 is provided in its less reactogenic composition where it is quenched with an exogenous sterol, such as cholesterol for example. Several particular forms of less reactogenic compositions wherein QS21 is quenched with an exogenous cholesterol exist. In a specific embodiment, the saponin /sterol is in the form of a liposome structure (WO 96/33739, Example 1). In this embodiment the liposomes suitably contain a neutral lipid, for example phosphatidylcholine, which is suitably non-crystalline at room temperature, for example eggyolk phosphatidylcholine, dioleoyl phosphatidylcholine (DOPC) or dilauryl phosphatidylcholine. The liposomes may also contain a charged lipid which increases the stability of the lipsome-QS21 structure for liposomes composed of saturated lipids. In these cases the amount of charged lipid is suitably 1-20% w/w, preferably 5-10%. The ratio of sterol to phospholipid is 1-50% (mol/mol), suitably 20-25%.

Suitable sterols include β-sitosterol, stigmasterol, ergosterol, ergocalciferol and cholesterol. In one particular embodiment, the adjuvant composition comprises cholesterol as sterol. These sterols are well known in the art, for example cholesterol is disclosed in the Merck Index, 11th Edn., page 341, as a naturally occurring sterol found in animal fat.

Where the active saponin fraction is QS21, the ratio of QS21 : sterol will typically be in the order of 1:100 to 1:1 (w/w), suitably between 1:10 to 1:1 (w/w), and preferably 1:5 to 1:1 (w/w). Suitably excess sterol is present, the ratio of QS21:sterol being at least 1:2 (w/w). In one embodiment, the ratio of QS21:sterol is 1:5 (w/w). The sterol is suitably cholesterol.

In one embodiment, the invention provides a dose of an immunogenic composition comprising immunologically active saponin, preferably QS21, at a level of 60µg or less, for example between 1 and 60µg. In one embodiment, the dose of the immunogenic composition comprises QS21 at a level of approximately around 50 µg, for example between 45 and 55 µg, suitably between 46 – 54 µg or between 47 and 53 µg or between 48 and 52 µg or between 49 and 51 µg, or 50 µg per dose.

In another embodiment the dose of the immunogenic composition comprises QS21 at a level of around 25 µg, for example between 20 – 30 µg, suitably between 21 – 29 µg or between 22 and 28 µg or between 23 and 27 µg or between 24 and 26 µg, or 25 µg.

In another embodiment, the dose of the immunogenic composition comprises QS21 at a level of around 10 µg per, for example between 5 and 15 µg, suitably between 6 and 14 µg, for example between 7 and 13 µg or between 8 and 12 µg or between 9 and 11 µg, or 10µg.

Specifically, a 0.5 ml vaccine dose volume contains 25 µg or 50 µg of QS21 per dose. Specifically, a 0.5 ml vaccine dose volume contains 50 µg of QS21 per dose.

The lipopolysaccharide may be a non-toxic derivative of lipid A, particularly monophosphoryl lipid A or more particularly 3-Deacylated monophoshoryl lipid A (3D – MPL).

3D-MPL is sold under the name MPL by GlaxoSmithKline Biologicals N.A. and is referred throughout the document as MPL or 3D-MPL. See, for example, US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094. 3D-MPL primarily promotes CD4+ T cell responses with an IFN-□ (Th1) phenotype. 3D-MPL can be produced according to the methods disclosed in GB 2 220 211 A. Chemically it is a mixture of 3-deacylated monophosphoryl lipid A with 3, 4, 5 or 6 acylated chains. Preferably in the compositions of the present invention small particle 3D-MPL is used. Small particle 3D-MPL has a particle size such that it may be sterile-filtered through a 0.22µm filter. Such preparations are described in WO 94/21292.

The invention therefore provides a dose of an immunogenic composition comprising lipopolysaccharide, preferably 3D-MPL, at a level of 75µg or less, for example between 1 and 60µg. In one embodiment the lipopolysaccharide is present at an amount of about 50µg per dose.

In one embodiment, the dose of the immunogenic composition comprises 3D-MPL at a level of around 50 µg, for example between 45 – 55 µg, suitably between 46 – 54 µg or between 47 and 53 µg or between 48 and 52 µg or between 49 and 51 µg, or 50 µg.

In one embodiment, the dose of the immunogenic composition comprises 3D-MPL at a level of around 25 µg, for example between 20 – 30 µg, suitably between 21 – 29 µg or between 22 and 28 µg or between 23 and 27 µg or between 24 and 26 µg, or 25 µg.

In another embodiment, the dose of the immunogenic composition comprises 3D-MPL at a level of around 10 µg, for example between 5 and 15 µg, suitably between 6 and 14 µg, for example between 7 and 13 µg or between 8 and 12 µg or between 9 and 11 µg, or 10µg.

In one embodiment, the volume of the dose is 0.5 ml. In a further embodiment, the immunogenic composition is in a volume suitable for a dose which volume is higher than 0.5 ml, for example 0.6, 0.7, 0.8, 0.9 or 1 ml. In a further embodiment, the human dose is between 1 ml and 1.5 ml.

Specifically, a 0.5 ml vaccine dose volume contains 25 µg or 50 µg of 3D-MPL per dose. Specifically, a 0.5 ml vaccine dose volume contains 50 µg of 3D-MPL per dose.

The dose of the immunogenic composition according to any aspect of the invention suitably refers to human dose. By the term “human dose” is meant a dose which is in a volume suitable for human use. Generally this is between 0.3 and 1.5 ml. In one embodiment, a human dose is 0.5 ml. In a further embodiment, a human dose is higher than 0.5 ml, for example 0.6, 0.7, 0.8, 0.9 or 1 ml. In a further embodiment, a human dose is between 1 ml and 1.5 ml.

Suitable compositions of the invention are those wherein liposomes are initially prepared without MPL (as described in WO 96/33739), and MPL is then added, suitably as small particles of below 100 nm particles or particles that are susceptible to sterile filtration through a 0.22 µm membrane. The MPL is therefore not contained within the vesicle membrane (known as MPL out). Compositions where the MPL is contained within the vesicle membrane (known as MPL in) also form an aspect of the invention. The polypeptide comprising a *C. difficile* toxin A fragment and / or a *C. difficile* toxin B fragment can be contained within the vesicle membrane or contained outside the vesicle membrane.

In a specific embodiment, QS21 and 3D-MPL are present in the same final concentration per dose of the immunogenic composition. In one aspect of this embodiment, a dose of

immunogenic composition comprises a final level of 25 µg of 3D-MPL and 25 µg of QS21 or 50 µg of 3D-MPL and 50 µg of QS21.

In one embodiment, the adjuvant includes an oil-in-water emulsion. For example, the oil-in-water emulsion can include an oil phase that incorporates a metabolisable oil, and an additional oil phase component, such as a tocol. The oil-in-water emulsion may also contain an aqueous component, such as a buffered saline solution (e.g., phosphate buffered saline). In addition, the oil-in-water emulsion typically contains an emulsifier. In one embodiment, the metabolizable oil is squalene. In one embodiment, the tocol is alpha-tocopherol. In one embodiment, the emulsifier is a nonionic surfactant emulsifier (such as polyoxyethethylene sorbitan monooleate, TWEEN80™). In exemplary embodiments, the oil-in-water emulsion contains squalene and alpha tocopherol in a ratio which is equal or less than 1 (w/w).

The metabolisable oil in the oil-in-water emulsion may be present in an amount of 0.5-10mg. The tocol in the oil-in-water emulsion may be present in an amount of 0.5 – 11 mg. The emulsifying agent in the oil-in-water emulsion may be present in an amount of 0.4 – 4 mg.

In order for any oil in water composition to be suitable for human administration, the oil phase of the emulsion system has to comprise a metabolisable oil. The meaning of the term metabolisable oil is well known in the art. Metabolisable can be defined as 'being capable of being transformed by metabolism' (Dorland's Illustrated Medical Dictionary, W.B. Sanders Company, 25th edition (1974)). The oil may be any vegetable oil, fish oil, animal oil or synthetic oil, which is not toxic to the recipient and is capable of being transformed by metabolism. Nuts, seeds, and grains are common sources of vegetable oils. Synthetic oils are also part of this invention and can include commercially available oils such as NEOBEE® (caprylic/capric triglycerides made using glycerol from vegetable oil sources and medium-chain fatty acids (MCTs) from coconut or palm kernel oils) and others. A particularly suitable metabolisable oil is squalene. Squalene (2,6,10,15,19,23-Hexamethyl-2,6,10,14,18,22-tetracosahexaene) is an unsaturated oil which is found in large quantities in shark-liver oil, and in lower quantities in olive oil, wheat germ oil, rice bran oil, and yeast, and is a particularly preferred oil for use in this invention. Squalene is a metabolisable oil by virtue of the fact that it is an intermediate in the biosynthesis of cholesterol (Merck index, 10th Edition, entry no.8619).

Suitably the metabolisable oil is present in the adjuvant composition in an amount of 0.5-10 mg, preferably 1-10, 2-10, 3-9, 4-8, 5-7, or 5-6 mg (e.g. 2-3, 5-6, or 9-10mg), specifically about 5.35 mg or around 2.14mg/dose.

Tocols are well known in the art and are described in EP0382271. Suitably the tocol is alpha-tocopherol or a derivative thereof such as alpha-tocopherol succinate (also known as vitamin E succinate). Said tocol is suitably present in in an amount of 0.5-11 mg, preferably 1-11, 2-10, 3-9, 4-8, 5-7, 5-6 mg (e.g. 10-11, 5-6, 2.5-3.5 or 1-3 mg). In a specific embodiment the tocol is present in an amount of about 5.94 mg or around 2.38mg per dose.

The oil in water emulsion further comprises an emulsifying agent. The emulsifying agent may suitably be polyoxyethylene sorbitan monooleate. In a particular embodiment the emulsifying agent may be Polysorbate® 80 (Polyoxyethylene (20) sorbitan monooleate) or Tween® 80.

Said emulsifying agent is suitably present in the adjuvant composition in an amount of 0.1-5, 0.2-5, 0.3-4, 0.4-3 or 2-3 mg (e.g. 0.4-1.2, 2-3 or 4-5 mg) emulsifying agent. In a specific embodiment the emulsifying agent is present in an amount of about 0.97 mg or about 2.425 mg.

In one embodiment, the amounts of specific components present in the composition are the amounts present in a 0.5 ml human dose. In a further embodiment, the immunogenic composition is in a volume suitable for a human dose which volume is higher than 0.5 ml, for example 0.6, 0.7, 0.8, 0.9 or 1 ml. In a further embodiment, the human dose is between 1 ml and 1.5 ml.

Where the adjuvant is in a liquid form and is to be combined with a liquid form of a polypeptide composition, the adjuvant composition in a human dose will be a fraction of the intended final volume of the human dose, for example approximately half of the intended final volume of the human dose, for example a 350 µl volume for an intended human dose of 0.7ml, or a 250 µl volume for an intended human dose of 0.5 ml. The adjuvant composition is diluted when combined with the polypeptide antigen composition to provide the final human dose of vaccine. The final volume of such dose will of course vary dependent on the initial volume of the adjuvant composition and the volume of polypeptide antigen composition added to the adjuvant composition. In an alternative embodiment, a liquid adjuvant is used to reconstitute a lyophilised polypeptide composition. In this embodiment, the human dose of the adjuvant composition is approximately equal to the final volume of the human dose. The liquid adjuvant composition is added to the vial containing the lyophilised polypeptide composition. The final human dose can vary between 0.5 and 1.5 ml.

The method of producing oil-in-water emulsions is well known to the person skilled in the art. Commonly, the method comprises mixing the tocol-containing oil phase with a surfactant such as a PBS/ polyoxyethylene sorbitan monooleate solution, followed by homogenisation using a

homogenizer. It would be clear to a man skilled in the art that a method comprising passing the mixture twice through a syringe needle would be suitable for homogenising small volumes of liquid. Equally, the emulsification process in microfluidiser (M110S Microfluidics machine, maximum of 50 passes, for a period of 2 minutes at maximum pressure input of 6 bar (output pressure of about 850 bar)) could be adapted by the man skilled in the art to produce smaller or larger volumes of emulsion. The adaptation could be achieved by routine experimentation comprising the measurement of the resultant emulsion until a preparation was achieved with oil droplets of the required diameter.

In an oil in water emulsion, the oil and emulsifier should be in an aqueous carrier. The aqueous carrier may be, for example, phosphate buffered saline.

Preferably the oil-in-water emulsion systems of the present invention have a small oil droplet size in the sub-micron range. Suitably the droplet sizes will be in the range 120 to 750 nm, more preferably sizes from 120 to 600 nm in diameter. Most preferably the oil-in water emulsion contains oil droplets of which at least 70% by intensity are less than 500 nm in diameter, more preferably at least 80% by intensity are less than 300 nm in diameter, more preferably at least 90% by intensity are in the range of 120 to 200 nm in diameter.

In one embodiment, the immunogenic composition is not 3µg or 10 µg of any of SEQ ID Nos. 1 to 7 combined with an adjuvant comprising an oil in water emulsion having 0.125 mL SB62 emulsion (Total volume), 5.35 mg squalene, 5.94 mg DL-α-tocopherol and 2.425 mg polysorbate 80 per 0.5 ml dose. In one embodiment, the immunogenic composition is not 3µg or 10 µg of any of SEQ ID Nos. 1 to 7 combined with an adjuvant comprising an oil in water emulsion 5.35 mg squalene, 5.94 mg DL-α-tocopherol and 2.425 mg polysorbate 80 per 0.5 ml dose. In one embodiment, the immunogenic composition does not contain an adjuvant comprising a oil in water emulsion having squalene, DL-α-tocopherol and polysorbate 80.

IMMUNOGENIC COMPOSITIONS AND VACCINES

In one embodiment the immunogenic composition further comprises additional antigens. In one embodiment the additional antigens are antigens derived from a bacterium selected from the group consisting of *Streptococcus pneumoniae (S. pneumoniae)*, *Haemophilus influenzae (H.influenzae)*, *Neisseria meningitidis (N.meningitidis)*, *Escherichia coli (E.coli)*, *Moraxella catarrhalis (M.cattarhalis)*, *Clostridium tetani (C. tetani)*, *Corynebacterium diptherieriae (C. diptherieriae)*, *Bordetella pertussis (B. pertussis)*, *Staphylococcus epidermidis* (*S.epidermidis)*, enterococci, and *Staphylococcus aureus (S.aureus)*. In a further embodiment the immunogenic

composition of the invention may comprise further antigens from *C.difficile* for example the S-layer proteins (WO01/73030). Optionally the immunogenic composition further comprises a saccharide from *C.difficile*.

There is further provided a vaccine comprising the immunogenic composition.This vaccine may further comprise a pharmaceutically acceptable excipient. In a further aspect of the invention there is provided a vaccine comprising the immunogenic composition of the invention and an adjuvant.

The vaccine preparations containing immunogenic compositions of the present invention may be used to protect a mammal susceptible to *C.difficile* infection or treat a mammal with a *C.difficile* infection, by means of administering said vaccine via a systemic or mucosal route. These administrations may include injection via the intramuscular, intraperitoneal, intradermal or subcutaneous routes; or via mucosal administration to the oral/alimentary, respiratory, genitourinary tracts. Although the vaccine of the invention may be administered as a single dose, components thereof may also be co-administered together at the same time or at different times (for instance pneumococcal saccharide conjugates could be administered separately, at the same time or 1-2 weeks after the administration of the any bacterial protein component of the vaccine for coordination of the immune responses with respect to each other). In addition to a single route of administration, 2 different routes of administration may be used. For example, saccharides or saccharide conjugates may be administered intramuscularly (IM) or intradermally (ID) and bacterial proteins may be administered intranasally (IN) or intradermally (ID). In addition, the vaccines of the invention may be administered IM for priming doses and IN for booster doses.

The content of toxins in the vaccine will typically be in the range 1-250μg, preferably 5-50μg, most typically in the range 5 - 25μg. Following an initial vaccination, subjects may receive one or several booster immunizations adequately spaced. Vaccine preparation is generally described in Vaccine Design ("The subunit and adjuvant approach" (eds Powell M.F. & Newman M.J.) (1995) Plenum Press New York). Encapsulation within liposomes is described by Fullerton, US Patent 4,235,877.

In one aspect of the invention is provided a vaccine kit, comprising a vial containing an immunogenic composition of the invention, optionally in lyophilised form, and further comprising a vial containing an adjuvant as described herein. It is envisioned that in this aspect of the invention, the adjuvant will be used to reconstitute the lyophilised immunogenic composition.

A further aspect of the invention is a method of preventing or treating *C.difficile* infection comprising administering to the subject an immunoprotective dose of the immunogenic composition or vaccine of the invention. In one embodiment there is provided a method of

preventing or treating primary and/or recurrence episodes of *C.difficile* infection comprising administering to the subject an immunoprotective dose of the immunogenic composition or vaccine of the invention. In one embodiment the subject is selected from the group consisting of mouse, rabbit, guinea pig, monkey, non-human primate or human. In one embodiment the subject is a mouse. In a further embodiment the subject is a human

A further aspect of the invention is an immunogenic composition or vaccine or kit of the invention for use in the treatment or prevention of infection or disease caused by *C.difficile*. In one embodiment there is provided an immunogenic composition or vaccine or kit of the invention for use in the treatment or prevention of primary and/or recurrence episodes of *C.difficile* disease.

A further aspect of the invention is use of the immunogenic composition or vaccine or kit of the invention in the manufacture of a medicament for the treatment or prevention of *C.difficile* disease. In one embodiment there is provided an immunogenic composition or vaccine or kit of the invention for use in the manufacture of a medicament for the treatment or prevention of primary and/or recurrence episodes of *C.difficile* disease.

"*C.difficile* disease" refers to any infection or disease caused by toxins released by *C. difficile*. Examples of *C.difficile* disease are antibiotic-associated diarrhea (AAD), pseudomembranous colitis and toxic megacolon, which can be life-threatening.

"Around" or "approximately" are defined as within 10% more or less of the given figure for the purposes of the invention.

The terms "comprising", "comprise" and "comprises" herein are intended by the inventors to be optionally substitutable with the terms "consisting of", "consist of" and "consists of", respectively, in every instance. The term "comprises" means "includes." Thus, unless the context requires otherwise, the word "comprises," and variations such as "comprise" and "comprising" will be understood to imply the inclusion of a stated compound or composition (*e.g.*, nucleic acid, polypeptide, antigen) or step, or group of compounds or steps, but not to the exclusion of any other compounds, composition, steps, or groups thereof. The abbreviation, "*e.g.*" is derived from the Latin exempli gratia, and is used herein to indicate a non-limiting example. Thus, the abbreviation "*e.g.*" is synonymous with the term "for example."

Embodiments herein relating to "vaccine" of the invention are also applicable to embodiments relating to "immunogenic compositions" of the invention, and vice versa.

Unless otherwise explained, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Definitions of common terms in molecular biology can be found in Benjamin Lewin, *Genes V*, published by 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); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

The singular terms "a," "an," and "the" include plural referents unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. The term "plurality" refers to two or more. It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are estimates, and are provided for description. Additionally, numerical limitations given with respect to concentrations or levels of a substance, such as a polypeptdie antigen, may be approximate.

All references or patent applications cited within this patent specification are incorporated by reference herein in their entirety.

These examples set forth below are for purposes of illustration only, and are not to be construed as limiting the scope of the invention in any manner.

It will be understood that particular aspects and embodiments described herein are shown by way of illustration and not as limitations of the invention. The principal features of this invention can be employed in various embodiments without departing from the scope of the invention. Those skilled in the art will recognize, or be able to ascertain using routine study, numerous equivalents to the specific procedures described herein. Such equivalents are considered to be within the scope of this invention and are covered by the claims.

All publications and patent applications mentioned in the specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent

applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference. All documents referred to herein are incorporated by reference to the fullest extent permissible.

The use of the word "a" or "an" when used in conjunction with the term "comprising" in the claims and/or the specification may mean "one," but it is also consistent with the meaning of "one or more," "at least one," and "one or more than one."

The invention will be further described by reference to the following, non-limiting, figures and examples.

EXAMPLES

Adjuvants used in the experiments:

Adjuvant A

Adjuvant having 50 µg QS21 presented in the form of a liposome, 50 µg 3D-MPL, 0.25 mg cholesterol and 1.0 mg DOPC per 0.5ml dose. A dose of 50µl suitable for immunizing hamsters contains 5µg QS21, 5µg 3D-MPL, 0.025mg cholesterol and 0.1mg DOPC.

Adjuvant B

Adjuvant consisting of an oil in water emulsion having 0.125 mL SB62 emulsion (Total volume), 5.35 mg squalene, 5.94 mg DL-α-tocopherol and 2.425 mg polysorbate 80 per 0.5 ml dose. A 50µl dose contains 0.0125ml SB62 emulsion (Total volme), 0.535mg squalene, 0.594mg DL-α-tocopherol and 0.2425 mg polysorbate 80

Adjuvant C

Adjuvant having 25 µg QS21 presented in the form of a liposome, 25 µg 3D-MPL, 0.25 mg cholesterol and 1.0 mg DOPC. A 50µl dose contains 2.5µg QS1`, 2.5µg 3D-MPL, 0.025mg cholesterol and 0.1mg DOPC.

GSK adjuvant AS04D

Adjuvant having 50µg MPL and 0.5mg (Al^{3+} total) per 0.5ml dose. A dose of 50µl contains 5µg MPL and 50µg (AL^{3+} total).

Alum

Adjuvant having Aluminium hydroxide, 500µg per 0.5ml dose. A dose of 50µl suitable of hamster immunisation contains 50µg per 50µg dose.

Protocols

Anti-ToxA and anti-ToxB ELISA response: Protocol for Examples 2 to 4, 7 and 8

ToxA or ToxB fragments (for the examples 2 and 4 -ToxA (2121-2686) and ToxB (1968-2366) for the examples 3, 5, 6, 7 and 8 -ToxA (2387-2706) and ToxB (1750-2360)) were coated at 2 µg/ml (for ToxA) or 1 µg/ml (for ToxB) in phosphate buffered saline (PBS) on high-binding microtitre plates (Nunc MAXISORP™), overnight at 4° C. The plates were blocked with PBS+1% Bovine serum albumin (BSA) for 30 min at RT with agitation. Hamster or mice sera are prediluted 1/100 (for post dose I or II samples) or 1/500 (for post dose III samples) in PBS-BSA0.2%-TWEEN™ 0.05% and then, 2 further twofold dilutions were made in microplates and incubated at room temperature (RT) for 30 min. After washing, bound hamster or mouse antibody was detected using Jackson ImmunoLaboratories Inc. peroxidase-conjugated Anti-Mouse (ref: 110-035-003) or Anti-Hamster (ref: 107-035-142) diluted 1:5000 in PBS-BSA0.2%-tween 0.05% ('detection antibody'). The detection antibodies were incubated for 30 min at 20°C with agitation. The color was developed using 4 mg O-phenylenediamine (OPD) + 5 µl H_2O_2 per 10 ml pH 4.5 0.1M citrate buffer for 15 minutes in the dark at room temperature. The reaction was stopped with 50 µl HCl, and the optical density (OD) was read at 490 nm relative to 620 nm.

The level of anti-ToxA or anti-ToxB antibodies present in each individual sera is determined by comparison to a reference serum (calibrated against Jackson Global IgG) added on each plate and expressed in µg/ml. A geometric mean titer GMT was calculated using Excel software for the samples in each treatment group for the hamster or mice immunizations.

Example 1

Design of five *C.difficile* ToxA-ToxB fusion proteins

Fusion proteins (also called 'fusions') containing fragments of the C-terminal repeating domains of ToxA and ToxB were designed. These fusion proteins contained a fragment of the C-terminal repeating domain of ToxA and a fragment of the C-terminal repeating domain of ToxB, and a junction between the C-terminal end of the ToxA fragment and the N terminal end of the ToxB fragment. Two strategies were devised, in the first strategy; the fusion was designed such that the long solenoid structure was maintained at the junction between the two fragments. In the second strategy, the two fragments of the fusions are separated by a linker to allow their independent correct folding.

The C-terminal part of ToxA and ToxB is composed of repeated sequences: short repeats (SR) and long repeats (LR) (PNAS 2005 vol 102 : 18373-18378).

The partial known 3D structure for the C-terminal domain of ToxA has been described (PNAS 2005 Greco et al., vol 102 : 18373-18378 ; Nature Structural & Molecular biology 2006 vol 13(5) : 460-461 ; PDB codes : 2F6E, 2G7C and 2QJ6).

The inventors predicted that there are two kinds of important interactions between residues of the C-terminal part of ToxA and ToxB. The first interaction occurs between residues contained in a LR and its preceding SR and is important to maintain the solenoid-like structure. The second type of interaction occurs between residues contained in a LR and the following SR and this interaction is mediating the carbohydrate-binding function of the toxin.

A new “structural-functional” repeat SR-LR-SR was defined. The structure of this repeat was maintained intact in the designed fusions.

The positions of the short (SR) and long repeats (LR) of ToxA and ToxB repeats are presented in Table 1.

A list of the “SR-LR-SR” boxes contained in the C-terminal domain of ToxA and ToxB is presented in Table 2 below.

Name	Start position	End position
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

Finally, the number of SRs between two LRs were maintained in the designed fusions to keep the long solenoid-like structure.

Before the design of junctions for the fusions, two working hypotheses were defined: first hypothesis, the shorter the fusions, the better the probability for the fusions to be stably over expressed; second hypothesis, according to the concept of “SR-LR-SR” boxes, the start position has to be chosen in order to ensure a correct folding of the first SR of this previously defined SR-

LR-SR box. Thus the fusions start at the beginning of the SR that precedes the SR-LR-SR box. Using these two hypotheses, three start positions were analysed: residue 2370, 2234 and 2121 of ToxA.

The start position 2370 was excluded. The start position 2234 was also excluded because one of the residues involved in interactions important for the protein structural stability is not conserved. So, it was decided that all the designed fusion will begin at residue 2121 of ToxA.

All fusions will end at the last residue of ToxB.

Four fusions (F1-4) were designed in order to maintain the entire fusion in a long solenoid-like structure between the two fusion fragments.

Fusions 1 (F1) and 2 (F2) were designed using the same hypothesis. All SR protein sequences of ToxA and ToxB had been compared using a multiple alignment software (ClustalW - Thompson JD *et al.* (1994) *Nucleic Acids Res.*, 22, 4673-4680). The more similar sequences were the third SR VIII of ToxA and the third SR II of ToxB and third SR III of ToxB. In order to make a choice between these two SR of ToxB, a structural homology modelling (using the SwissModel interface - Arnold K *et al.* (2006) *Bioinformatics*, 22, 195-201) was performed on the C-terminal part of ToxB using the known 3D structure of partial ToxA C-terminal domain (PDB code : 2QJ6). Using the third SR VIII of ToxA, the best local structural superposition (performed using SwissPDBViewer - Guex N *et al.* (1997), *Electrophoresis* 18, 2714-2723) was obtained with the third SR II of ToxB. So, two junctions were designed : the first one is between the third SR VIII of ToxA and the fourth SR II of ToxB (F1) and the second one is between the second SR VIII of ToxA and the third SR II of ToxB (F2).

To design fusion 3 (F3), a global structural superposition was performed between both the known structure of the partial C-terminal domain of ToxA and the predicted structure of C-terminal domain of ToxB (using SwissModel and SwissPDBViewer softwares). The best superposition was found between LR VII of ToxA and LR II of ToxB. So, it was decided to make a junction in this similar LR. The junction was performed firstly in a region where the sequence is conserved between ToxA and ToxB, after that in order to keep in the ToxA part of the fusion, the residues in interaction with the preceding SR and lastly, in order to keep in the ToxB part, the residues in interaction with the following SR.

For the design of fusion 4 (F4), the C-terminal domain of ToxB was divided in 4 fragments and a more precise homology modelling (SwissModel) was performed on them. The split was realised in order to keep intact the “SR-LR-SR” boxes (each domain finishes at the end of the SR that follows a LR). A structural superposition between the predicted structures of these fragments and the known 3D structure of ToxA was made and the best structural surperposition was obtained for

the third SR of ToxB (SR I) and the last SR of ToxA (third SR VIII). So, the junction was done between the second SR VIII of ToxA and the third SRIof ToxB.

The last fusion (F5) was designed in order to allow an independent correct folding of the two fragments of the fusion. A linker was added between the last residue of the ToxA protein sequence and the beginning of the fourth SR II of ToxB (always taking into account the importance of an intact "SR-LR-SR" box). Only one exogenous residue (Glycine) was added as a linker and located between two existing Glycines. Thus, the linker can also be described as composed of 3 Glycines surrounding by known (for ToxA) and predicted (for ToxB) beta-strand.

Example 2

Immunisation of hamsters with F2 fusion protein : Adjuvant A, Adjuvant A without QS21 or Adjuvant A without MPLand comparison with Adjuvant B, Alum or non adjuvanted formulations.

Groups of 16 hamsters (mix of male and female) were immunized IM at days 0, 14 and 28 with 1µg of the F2 fusion protein adjuvanted with 50µl of Adjuvant A, Adjuvant A without MPL, Adjuvant A without QS21, Adjuvant B , Alum or non adjuvanted.
Anti-ToxA and anti-ToxB ELISA titers were determined in individual sera collected at day 41/42 (Post III). The results are described in tables 3 and 4 and figures 1 and 2.

Table 3: ELISA anti-ToxA : concentrations (µg/ml) on individual day 42 sera

Groups	1	2	3	4	5	6
	F2 1µg/dose	F2 1µg/dose	F2 1µg/dose	F2 1µg/dose	F2 1µg/dose	F2 1µg/dose
ADJUVANT	A	A-QS21	A - MPL	B	Al(OH)3	non-adjuvanted
1	1472	29	935	457	577	3
2	1828	39	414	197	329	23
3	4928	35	747	375	401	64
4	2966	28	550	60	769	25
5	2025	19	388	408	652	3
6	1116	155	1498	173	449	37
7	2579	43	879	732	360	179
8	2463	33	780	196	1342	31
9	1425	52	1183	681	817	7
10	1984	307	696	190	534	199
11	1093	198	572	344	2665	106
12	2080	126	972	508	1944	22
13	3541	403	1823	193	1603	221

14	1635	152	813	1142	2209	680
15	4688	630	1046	494	855	359
16	3360	21	1119	163	664	100
Geomean	2213	77	830	312	815	49
CI-	1717	40	655	202	563	14
CI+	2802	124	1032	454	1129	104
St dev min	496	37	174	111	251	36
St dev max	589	47	202	141	314	55

CI=confidence interval

St Dev=standard deviation

Table 4: ELISA anti-ToxB : concentrations (µg/ml) on individual day 42 sera

Groups	1	2	3	4	5	6
	F2 1µg/dose	F2 1µg/dose	F2 1µg/dose	F2 1µg/dose	F2 1µg/dose	F2 1µg/dose
ADJUVANT	A	A-QS21	A - MPL	B	Al(OH)3	non-adjuvanted
1	811	33	817	65	63	5
2	1186	88	1144	107	177	5
3	4388	90	943	207	216	92
4	1158	66	1224	464	239	26
5	1992	53	1088	212	91	5
6	1518	305	1070	193	138	44
7	5437	54	1021	244	44	57
8	1493	74	899	266	313	28
9	1488	281	1506	189	135	30
10	2637	346	1588	214	318	256
11	1568	215	1125	251	863	95
12	1879	50	1890	549	534	93
13	1920	237	1724	147	355	57
14	1709	266	2450	606	543	57
15	1553	492	2033	301	251	178
16	1556	95	744	167	115	88
Geomean	1789	125	1254	225	207	41
CI-	1387	76	1039	162	128	16
CI+	2264	186	1498	301	308	76
St dev min	402	49	215	63	79	25
St dev max	475	61	244	75	101	34

5

Figures 1 and 2

Results for figure 1:

The F2 was injected with the full Adjuvant A, a similar formulation without QS21, or without MPL.

For the anti-ToxA response, both MPL and QS21 have an added value even if the QS21 seems the most effective.

For the anti-ToxB response, the effect of the MPL is less clear while the QS21 removal has a drastric effect on the IgG titers.

Results for figure 2:

Regarding the adjuvant comparison, the adjuvant A induced significant higher anti-ToxA and anti-ToxB antibodies compared to Alum, Adjuvant B or non adjuvanted formulations.Adjuvant B induced higher anti-ToxA and anti-ToxB antibodies compared to the non adjuvanted formulations.

Example 3

Pims 20110544 – Immunisation of male hamsters with a dose-range of F2 fusion protein formulated in Adjuvant B, AS04D, Adjuvant A or Adjuvant C.

Groups of 10 male hamsters were immunized IM at days 0, 14 and 28 with 0.3 µg , 1µg or 3 µg of the F2 fusion protein adjuvanted with 50µl doses of Adjuvant B, AS04D, Adjuvant A or Adjuvant C.

Anti-ToxA and anti-ToxB ELISA titers were measured from individual sera collected at day 13 (post I), day 27 (Post II) ,day 42 (Post III 14) and day 84 (Post III 56)

Pims 20110545 – Immunisation of female hamsters with a dose-range of F2 fusion protein formulated in Adjuvant B, AS04D, Adjuvant A or Adjuvant C.

Groups of 10 female hamsters were immunized IM at days 0, 14 and 28 with 0.3 µg , 1µg or 3 µg of the F2 fusion protein adjuvanted with Adjuvant B, AS04D, Adjuvant A or Adjuvant C.

Anti-ToxA and anti-ToxB ELISA titers were measured from individual sera collected at day 13 (post I), day 27 (Post II) ,day 42 (Post III 14) and day 84 (Post III 56)

Figures 3 to 7

Pims 20110544 and 20110545

Table 5: Post I Anti-ToxA ELISA titers (Geometric mean titre or GMT)

Group	1	2	3	4	5	6	7	8	9	10	11	12
Ag/ animal	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg
Adj.	C	C	C	A	A	A	B	B	B	AS04 D	AS04 D	AS04 D
544	28	17	6	84	17	6	44	27	3	68	74	33
545	50	36	12	109	111	26	85	20	16	194	107	36

Table 6: Post I Anti-ToxB ELISA titers (GMT)

Group	1	2	3	4	5	6	7	8	9	10	11	12
Ag/ animal	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg
Adj	C	C	C	A	A	A	B	B	B	AS04 D	AS04 D	AS04 D
544	115	82	52	412	121	79	189	119	34	7	5	5
545 (femal e)	178	257	150	391	383	224	196	164	69	27	10	3

Table 7: Post II Anti-ToxA ELISA titers (GMT)

Groups	1	2	3	4	5	6	7	8	9	10	11	12
Ag/animal	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg
Adjuvant	C	C	C	A	A	A	B	B	B	AS04D	AS04D	AS04D
544	475	416	191	974	530	277	312	332	138	487	597	401
545	762	834	428	1466	1424	688	818	423	650	803	844	849

Table 8: Post II Anti-ToxB ELISA titers (GMT)

Groups	1	2	3	4	5	6	7	8	9	10	11	12
Ag/animal	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg
Adjuvant	C	C	C	A	A	A	B	B	B	AS04D	AS04D	AS04D
544(male)	319	686	488	1181	1120	1049	435	401	340	64	56	28
545 (female)	974	1555	1419	1552	1871	2386	585	649	1486	215	109	91

Table 9: Post III day 42 Anti-ToxA ELISA titers (GMT)

Groups	1	2	3	4	5	6	7	8	9	10	11	12
Ag/animal	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg
Adjuvant	C	C	C	A	A	A	B	B	B	AS04D	AS04D	AS04D
544	449	674	471	1462	720	688	247	486	266	662	599	566
545	642	814	667	1654	1020	1560	800	539	888	923	1005	794

Table 10: Post III day 42 Anti-ToxB ELISA titers (GMT)

Groups	1	2	3	4	5	6	7	8	9	10	11	12
Ag/animal	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg
Adjuvant	C	C	C	A	A	A	B	B	B	AS04D	AS04D	AS04D
544	630	612	594	1259	1116	1172	335	426	512	67	86	51
545	568	831	550	1213	828	1273	556	695	944	182	145	134

Table 11: Post III day 84 Anti-ToxA ELISA titers (GMT)

Groups	1	2	3	4	5	6	7	8	9	10	11	12
Ag/animal	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg	1µg	0.3µg
Adjuvant	C	C	C	A	A	A	B	B	B	AS04D	AS04D	AS04D
544	213	239	169	357	356	307	106	70	154	182	257	266

545	311	309	401	595	578	717	311	251	417	353	414	337

Table 12: Post III day 84 Anti-ToxB ELISA titers (GMT)

Groups	1	2	3	4	5	6	7	8	9	10	11	12
Ag/animal	3µg	1µg	0.3µg	3µg	1µg	0.3µg	3µg l	1µg	0.3µg	3µg	1µg	0.3µg
Adjuvant	C	C	C	A	A	A	B	B	B	AS04D	AS04D	AS04D
544(male)	291	195	145	299	360	440	78	70	315	24	25	15
545 (female)	250	299	328	318	374	435	134	211	392	60	41	41

Example 4

Pims 20120011 and 20120008 – Immunisation of hamsters with F2 fusion protein or a mixture of Toxin A and Toxin B fragments, formulated in Alum, Adjuvant A or non adjuvanted.

Groups of 8 female and 8 male hamsters were immunized IM at days 0, 14 and 28 with 1 µg of the F2 fusion protein, an F2 fusion with a poly His tag (F2-His tagged), or a mix of C-terToxA (2387-2706) (0.6µg)+C-terToxB (1750-2360) (0.4µg) formulated in Alum, Adjuvant A or non adjuvanted.

Anti-ToxA and anti-ToxB ELISA titers were determined in individual sera collected at day 13, day 27 and day 42 (Post III). **Figure 8**

Table 13

20120008 and 20120011 : Female and Male hamster immunization. Comparison of the immunogenicity of a Mix (ToxA+ToxB) with the F2 fusion protein formulated in Alum or Adjuvant A ELISA anti-ToxA : concentrations (µg/ml) on individual Post III sera						
Groups	1	2	3	4	5	6
	Mix ToxA (0.6µg)+ToxB(0.4µg)	F2 (1µg)	Mix ToxA (0.6µg)+ToxB(0.4µg)	F2 (1µg)	F2-His tagged (1µg)	F2 (1µg)
	Al(OH)3	Al(OH)3	Adj. A	Adj. A	Adj. A	non-adjuvanted
1	62	1156	41	3075	4006	195
2	77	871	105	1385	2463	127
3	104	833	239	3119	1202	31
4	66	506	143	3806	3903	574
5	179	736	340	3180	6852	116
6	641	1239	48	Dead	5571	32
7	128	328	345	5327	3174	10
8	48	3998	63	3542	4990	47
9	59	681	296	888	6769	13

10	352	491	164	978	1192	2
11	Dead	604	363	2913	1791	3
12	47	1524	87	1231	4362	68
13	88	486	77	3473	1647	123
14	23	1185	62	1341	4929	35
15	254	602	644	2794	3085	136
16	206	1094	24	2975	3354	90
Geomean	101	823	144	2310	3234	42
CI-	65	612	78	1727	2398	20
CI+	172	1148	213	3196	4381	99

Table 14

20120008 and 20120011 : Female and Male hamster immunization.
Comparison of the immunogenicity of a Mix (ToxA+ToxB) with the F2 fusion protein formulated in Alum or Adjuvant A
ELISA anti-ToxB : concentrations (µg/ml) on individual Post III sera

Groups	1	2	3	4	5	6
	Mix ToxA (0.6µg)+ToxB(0.4µg)	F2 (1µg)	Mix ToxA (0.6µg)+ToxB(0.4µg)	F2 (1µg)	F2-His tagged (1µg)	F2 (1µg)
	Al(OH)3	Al(OH)3	ADJUVANT A	ADJUVANT A	ADJUVANT A	non-adjuvanted
1	58	591	2510	2364	2647	304
2	172	302	4981	952	1021	302
3	14	167	1409	2630	804	125
4	53	186	4158	2261	1090	299
5	107	408	3046	2553	2682	296
6	925	373	1190	Dead	3808	193
7	40	65	1673	2938	2820	170
8	112	1112	1809	1404	3186	51
9	160	652	2701	1385	2837	128
10	24	75	1210	2252	1676	38
11	Dead	242	546	1749	2174	12
12	61	1081	5746	1623	2557	74
13	15	193	1975	1507	2020	130
14	56	344	3499	1018	4735	14
15	15	237	2897	2554	2673	80
	332	309	2489	2894	2148	127
Geomean	61	297	2228	1834	2210	100
CI-	35	194	1626	1542	1708	59
CI+	132	457	3094	2319	2850	176

Geomean=Geometric Mean Titre

Results:

Anti-ToxA and anti-ToxB antibodies were induced after immunization with the F2 fusion protein but also with the mix of ToxA and ToxB. Adjuvant A induces significantly higher ELISA titers compared to other formulations.

Example 5

Immunisation of mice with Tox A or Tox B fragments and ToxA-ToxB fusions

Figures 9 to 12

Balb/C mice were immunized with the constructs described in example 1.

Groups of 15 female Balb/c mice were immunized IM at days 0, 14 and 28 with 3µg or 10 µg of the separate fragments of ToxA and ToxB (ToxA C-term (2387-2706) (0.6µg)+C-terToxB (1750-2360) as well as with ToxA-ToxB fusions proteins adjuvanted with Adjuvant B. A control group of 10 mice was vaccinated with Adjuvant B alone.

Anti-ToxA and anti-ToxB ELISA titers were determined in individual sera collected at day 42 (post III).

Hemagglutination inhibition titers were determined in pooled Post III sera.

Anti-ToxA and anti-ToxB ELISA response: Protocol

Samples of the ToxA or ToxB fragments were coated at 1 µg/ml for ToxB fragments or 2µg/ml for ToxA fragments in phosphate buffered saline (PBS) on high-binding microtitre plates (Nunc MAXISORP™), overnight at 4° C. The plates were blocked with PBS-BSA 1% for 30 min at RT with agitation. The mice anti-sera are prediluted 1/1000 in PBS-BSA0.2%-TWEEN™ 0.05%. and then, further twofold dilutions were made in microplates and incubated at RT for 30 min with agitation. After washing, bound murine antibody was detected using Jackson ImmunoLaboratories Inc. peroxidase-conjugated affiniPure Goat Anti-Mouse IgG (H+L) (ref: 115-035-003) diluted 1:5000 in PBS-BSA0.2%-tween 0.05%. The detection antibodies were incubated for 30 min. at 20°C with agitation. The color was developed using 4 mg O-phenylenediamine (OPD) + 5 µl H_2O_2 per 10 ml pH 4.5 0.1M citrate buffer for 15 minutes in the dark at room temperature. The reaction was stopped with 50 µl HCl, and the optical density (OD) was read at 490 nm relative to 620 nm.

The level of anti-ToxA or anti-ToxB antibodies present in the sera was expressed in mid-point titers by comparison with a reference serum. A GMT was calculated for the 15 samples in each treatment group (10 for the control group).

Hemagglutination inhibition assay: Protocol

8 serial twofold dilutions of mice pooled antisera (25µl) were performed in phosphate buffered saline (PBS) in 96-well U-bottom microplates.

25 µl of native Toxin A (0,2 µg/well) were then added and the plates were incubated at room temperature for 30 minutes.

After incubation, 50 µl of purified rabbit erythrocytes diluted at 2% were added to each well. The plates were incubated at 37°C for 2 hours.

Plates were analysed visually, with hemagglutination presenting as diffuse red cells in the well and the inhibition of hemagglutination observed as a red point settled in the well.

The inhibition titers were defined as the reciprocal of the highest dilution of the serum inhibiting hemagglutination.

Cytotoxicity assay

IMR90 fibroblast cells were cultured at 37°C with 5% CO_2, in Eagle's Minimal Essential Medium EMEM + 10% fetal calf serum + 1% glutamine + 1% antibiotics (penicillin-streptomycin-amphotericin) and were seeded in 96-well tissue culture plates at a density of 5.10^4 cells/well.

After 24h, 100µl of the cell media was removed from the wells.

8 (for the experiment in example 5) or 12 (for the experiment in example 6) serial twofold dilutions of mice pooled antisera (50µl) were performed in cell media.

50 µl of native Toxin B (0.5ng/ml) is then added and the plates incubated at 37°C with 5% CO2 for 24 hours.

Cells were observed after 24 hours, and the proportion of rounded cells was determined.

The inhibition titers were defined as the reciprocal of the highest dilution of the serum inhibiting 50% cell rounding.

Results:

Anti-Tox A antibodies were induced after immunization with the ToxA fragment alone but also with each of the 5 fusions.

The functional properties of these antibodies were tested in the hemagglutination assay. This assay is only adapted for Tox A evaluation as no hemagglutination is observed with ToxB.

Haemaglutination inhibition was observed with the anti-Tox A fragment sera or sera directed against each of the ToxA-ToxB fusions.

An ELISA using ToxB antibodies was also performed. Anti-Tox B antibodies were induced after immunization with the ToxB fragment alone but also with the F2, F3 and F4 fusions.

Inhibition titers obtained using sera from mice immunised with the ToxB fragment or the ToxA-ToxB fusions were greater than that obtained using control sera.

Example 6

Immunisation of mice with Tox A-Tox B fusions

Groups of 25 female mice were immunized IM at days 0, 14 and 28 with 3 µg of the ToxA-ToxB fusion proteins formulated in Adjuvant B or 10 µg of a non-adjuvanted mix of ToxA(6µg)+ToxB(4µg). In addition to F2, four other fusion protein constructs were used, there were F54 Gly (SEQ ID NO:21), F54 New (SEQ ID NO:23), F5 ToxB (SEQ ID NO:25) and F52 New (SEQ ID NO:27).

An ELISA was carried out using the anti-ToxA and anti-ToxB ELISA response:protocol described in example 5 except here the samples of the toxA or toxB fragments were coated at 1µg/ml in phosphate buffered saline on high-binding microtitre plates. A hemagglutination inhibition assay was performed as described in example 5 except that the rabbit erythrocytes were diluted at 2.5% . A toxB cytotoxicity assay was performed as described in example 5 except that 50µl of Toxin B at a concentration of 0.015µg/ml was used. A further toxA cytotoxicity assay was performed as described below.

Anti-ToxA and anti-ToxB ELISA titers were determined in individual sera collected at day 42 (Post III 14) . **Figures 13 and 14**

Table 15

20110350 : Mice immunization with C. difficile fusion proteins formulatd in Adjuvant B ELISA a-ToxA : concentrations (µg/ml) on Post III sera						
Groups	1	2	3	4	5	6
	F2	Fusion A (F52new)	Fusion B (F54Gly)	Fusion C (F54new)	Fusion D (F5 ToxB)	Mix ToxA + ToxB
	3µg Ag/animal	3µg Ag/animal	3µg Ag/animal	3µg Ag/animal	3µg Ag/animal	10µg (6µg A + 4µg B) Ag/animal
	ADJUVANT B	ADJUVANT B	ADJUVANT B	ADJUVANT B	ADJUVANT B	non-adjuvanted
	IM	IM	IM	IM	IM	IM
1	518	681	735	436	586	328
2	761	1058	975	455	1252	22
3	557	718	977	684	1108	327
4	1415	992	862	577	555	152
5	782	457	1202	861	622	135
6	563	724	611	575	569	54

7	615	585	1330	824	807	147
8	994	878	611	576	1108	169
9	412	623	Dead	389	887	124
10	495	993	Dead	809	1146	219
11	366	1039	923	872	965	12
12	1075	1089	1454	748	950	122
13	882	754	1176	809	725	196
14	1028	1550	1164	1578	696	312
15	738	1268	1170	1352	678	188
16	519	721	1028	781	930	224
17	927	890	805	346	791	209
18	946	992	647	1255	410	32
19	715	1346	1142	670	634	92
20	546	Dead	1061	Dead	Dead	166
21	1107	1604	826	1038	791	112
22	1351	1289	1223	708	778	66
23	889	1319	61	567	759	102
24	976	824	597	414	456	281
25	1153	1212	995	868	758	242
Geomean	764	937	841	704	761	126
CI-	656	818	641	597	674	89
CI+	889	1074	1103	830	858	179

Table 16

20110350 : Mice immunization with C. difficile fusion proteins formulatd in Adjuvant B ELISA a-ToxB : concentrations (µg/ml) on Post III sera						
Groups	1	2	3	4	5	6
	F2	Fusion A (F52new)	Fusion B (F54Gly)	Fusion C (F54new)	Fusion D (F5 ToxB)	Mix ToxA + ToxB
	3µg Ag/animal	3µg Ag/animal	3µg Ag/animal	3µg Ag/animal	3µg Ag/animal	10µg (6µg A + 4µg B) Ag/animal
	ADJUVANT B	ADJUVANT B	ADJUVANT B	ADJUVANT B	ADJUVANT B	ADJUVANT B
1	916	749	736	454	581	1823
2	1090	551	491	517	872	575
3	687	1178	574	549	529	1386
4	940	868	501	536	828	854
5	837	522	653	420	866	1170
6	836	609	397	757	690	2132
7	309	702	405	792	732	1466
8	569	381	598	737	856	1539
9	382	450	Dead	914	632	1461
10	552	693	Dead	465	639	2267
11	679	832	1273	520	505	1037
12	1123	575	809	551	431	2273
13	764	659	532	1077	930	1405

14	944	970	909	1065	870	1978
15	812	628	222	727	1052	1414
16	502	552	736	560	1001	1269
17	983	639	551	431	728	1732
18	1133	628	604	561	1085	1211
19	938	1058	805	470	435	2017
20	519	Dead	857	Dead	Dead	2140
21	887	1025	319	670	310	2342
22	861	769	596	555	1025	2897
23	903	1161	552	694	260	2343
24	1327	536	431	737	313	3074
25	674	518	680	472	663	2697
Geomean	765	688	581	611	653	1661
CI-	663	607	495	545	550	1411
CI+	884	781	682	686	776	1957

Figures 15 to 17

A hemagglutination inhibition assay was performed as described in example 5.

A toxB cytotoxicity assay was performed as described in example 5 except that 12 serial twofold dilutions of mice pooled antisera were performed rather than 8, and the plates were incubated overnight rather than for 24 hours before cells were observed. A further toxA cytotoxicity assay was performed as described below.

ToxA cytotoxicity assay

HT29 cells were cultured at 37°C with 5%CO_2 in DMEM +10% fetal bovine serum +1% glutamine +1% antibiotics (penicillin-streptomycin-amphotericin) and were seeded in 96-well tissue culture plates at a density of 5.10^4 cells/well. After 24h, the cell media was removed from the wells. Serial twofold dilutions of mice pooled antisera (50µl) were performed in cell media.

50µl of native Toxin A (0.15ng/ml) is then added and the plates incubated at 37°C with 5% CO_2 for 48 hours. Cells were observed after 48 hours and the proportion of rounded cells were determined. The results of the anti-ToxA ELISA, anti-ToxB Elisa, Haemagglutination inhibition and cytotoxicity assays are described in the aforementioned Figures 13 to 17.

Example 7

Immunisation of mice with *C.difficile* ToxA-Cter, ToxB-Cter or fusion proteins in a non-adjuvanted formulation

Groups of 15 female mice were immunized IM at days 0, 14 28 and 120 with 1 or 3 µg of ToxA C-ter, ToxB C-ter or ToxA-ToxB fusion proteins. All these antigens were injected in a non-adjuvanted formulation. A control group (10 mice) was injected with the NaCl alone.

Anti-ToxA and anti-ToxB ELISA titers were measured from individual sera collected at day 28 (Post II), day 42 (Post III 14) , day 120 (Pre-Boost 87) and day 134 (Post IV 14). **Figures 18 and 19.**

5

Table 17

20100723 : Mice immunization with *C. difficile* ToxA-Cter, ToxB-Cter and fusion proteins non-adjuvanted ELISA a-ToxA : GMT concentrations (µg/ml)															
Bleeding	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	ToxA (aa 2387-2706)	ToxB (aa 1750-2360)	Fusion F1	Fusion F2	Fusion F3	Fusion F4	Fusion F5	ToxA (aa 2387-2706)	ToxB (aa 1750-2360)	Fusion F1	Fusion F2	Fusion F3	Fusion F4	Fusion F5	
Ag/ml	10µg	10µg	10µg	10µg	10µg	10µg	10µg	3µg	3µg	3µg	3µg	3µg	3µg	3µg Ag/ml	
Adjuvant	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	NaCl
	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM
Post II	178	3	16	60	74	44	269	84	3	10	49	26	25	60	3
Post III	222	2	40	94	192	109	366	137	1.25	51	110	88	67	137	1.25
Pre-Boost	128	4	28	70	120	58	172	165	4	48	62	63	44	107	3
Post IV	539	5	90	174	249	177	353	466	3	168	147	163	139	188	3

Table 18

20100723 : Mice immunization with *C. difficile* ToxA-Cter, ToxB-Cter and fusion proteins non-adjuvanted ELISA a-ToxB : GMT concentrations (µg/ml)															
Bleeding	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	ToxA (aa 2387-2706)	ToxB (aa 1750-2360)	Fusion F1	Fusion F2	Fusion F3	Fusion F4	Fusion F5	ToxA (aa 2387-2706)	ToxB (aa 1750-2360)	Fusion F1	Fusion F2	Fusion F3	Fusion F4	Fusion F5	
Ag/ml	10µg	10µg	10µg	10µg	10µg	10µg	10µg	3µg	3µg	3µg	3µg	3µg	3µg	3µg Ag/ml	
Adjuvant	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	Non adj.	NaCl
	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM
Post II	2	62	2	110	58	283	18	2	7	2	101	29	125	8	2
Post III	3	316	7	200	199	455	65	4	105	13	242	126	257	44	2
Pre-Boost	3	125	5	81	58	157	18	4	38	8	68	42	101	14	3
Post IV	5	856	27	258	227	380	125	6	632	79	232	175	293	126	4

Example 8

Immunisation of mice with *C.difficile* ToxA-Cter, ToxB-Cter or fusion proteins formulated in Adjuvant B.

Figures 20 and 21

Groups of 15 female mice were immunized IM at days 0, 14 and 28 and 120 with 1 or 3 µg of ToxA C-term (amino acids 2387-2706) and ToxB C-term (amino acids 1750-2360) or ToxA-ToxB fusion proteins. These proteins were injected in an Adjuvant B formulation.
A control group (10 mice) was injected with the Adjuvant B alone.

Anti-ToxA and anti-ToxB ELISA titers were determined in individual sera collected at day 28 (Post II), day 42 (Post III 14) , day 120 (Pre Boost III 87 or pre boost) and day 134 (Post IV 14).

Table 19

Mice immunization with *C. difficile* ToxA-Cter, ToxB-Cter and fusion proteins formulated in Adjuvant B
ELISA a-ToxA : concentrations (µg/ml) on Post II, III, Pre-Boost, Post IV sera

Bleeding	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	ToxA (aa 2387-2706)	ToxB (aa 1750-2360)	F1	F2	F3	F4	F5	ToxA (aa 2387-2706)	ToxB (aa 1750-2360)	F1	F2	F3	F4	F5 n	/
Ag/ml	10µg	10µg	10µg	10µg	10µg	10µg	10µg	3µg	3µg	3µg	3µg	3µg	3µg	3µg	/
ADJ UVA NT	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM
Post II	213	2.3	465	332	312	158	672	534	1.7	620	498	578	248	925	2.1
Post III	852	3	1144	859	695	644	1137	1376	4	1356	768	950	470	1537	2
Pre-Boost	738	4	762	484	417	394	674	1091	3	725	513	598	324	1197	3
Post IV	2764	9	1450	1423	1036	986	1626	301	26	990	1000	1099	732	1465	1

Table 20

20100798 : Mice immunization with *C. difficile* ToxA-Cter, ToxB-Cter and fusion proteins formulated in Adjuvant B
ELISA a-ToxB : concentrations (µg/ml) on Post II, III, Pre-Boost, Post IV sera

Bleeding	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	ToxA (aa 2387-2706)	ToxB (aa 1750-2360)	F1	F2	F3	F4	F5	ToxA (aa 2387-2706)	ToxB (aa 1750-2360)	F1	F2	F3	F4	F5	/
Ag/ml	10µg	10µg	10µg	10µg	10µg	10µg	10µg	3µg	3µg	3µg	3µg	3µg	3µg	3µg	/

ADJ UVA NT	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM	IM
Post II	2	246	10	424	284	794	25	3	338	17	619	471	753	93	2
Post III	2	1762	285	793	671	1204	253	4	1888	232	951	834	1026	365	2
Pre-Boost	3	989	180	512	396	761	129	4	984	108	479	441	581	205	3
Post IV	3	4212	911	1422	1353	1799	1046	78	657	738	1272	747	1572	812	2

Example 9

Cloning expression and purification of the fusion proteins

Expression plasmid and recombinant strain

Genes encoding the fusion proteins of partial C-terminal domains of ToxA and ToxB (SEQ ID NO:3, 4, 5, 6 and 7) and a His tag were cloned into the pET24b(+) expression vector (Novagen) using the NdeI/XhoI restriction sites using standard procedures. The final construct was generated by the transformation of *E. coli* strain BLR (DE3) with the recombinant expression vector according to standard method with CaCl2-treated cells (Hanahan D. « Plasmid transformation by Simanis. » In Glover, D. M. (Ed), DNA cloning. IRL Press London. (1985): p. 109-135.).

Host strain:

BLR(DE3). BLR is a recA derivative of BL21. Strains having the designation (DE3) are lysogenic for a λ prophage that contains an IPTG inducible T7 RNA polymerase. λ DE3 lysogens are designed for protein expression from pET vectors This strain is also deficient in the *lon* and *ompT* proteases.

Genotype : *E.coli* BLR::DE3 strain, F^- *ompT* $hsdS_B(r_B^- m_B^-)$ *gal dcm* (*DE3*) Δ(srl-recA)306::Tn*10* (Tet^R)

Expression of the recombinant proteins:

An *E.coli* transformant was stripped from agar plate and used to inoculate 200 ml of LBT broth ± 1% (w/v) glucose + kanamycin (50 µg/ml) to obtain O.D.600nm between 0.1 -0.2. Cultures were incubated overnight at 37 °C, 250 RPM.

This overnight culture was diluted to 1:20 in 500 ml of LBT medium containing kanamycin (50 µg/ml) and grown at 37°C at a stirring speed of 250 rpm until O.D.620 reached 0.5/0.6.

At O.D.600nm around 0.6, the culture was cooled down before inducing the expression of the recombinant protein by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; EMD Chemicals Inc., catalogue number: 5815) and incubated overnight at 23 °C, 250 RPM.

After overnight induction (around 16 hours), $O.D._{600nm}$ was evaluated after induction and culture was centrifuged at 14 000 RPM for 15 minutes and pellets were frozen at -20°C separately.

Purification:

The bacterial pellet was resuspended in 20 mM bicine buffer (pH 8.0) containing 500 mM NaCl and a mixture of protease inhibitor (Complete, Roche). Bacteria were lysed using a French Press system 20 000 PSI. Soluble (supernatant) and insoluble (pellet) components were separated by centrifugation for example at 20 000g for 30 min at 4°C.

The 6-His tagged-protein was purified under native conditions on IMAC. The soluble components were loaded on a GE column (15 ml for example) (Ni loaded) preequilibrated with the same buffer used to bacterial resuspension. After loading on the column, the column was washed with the same buffer. Elution was performed using a 20mM bicine buffer (pH 8.0) containing 500 mM NaCl and different concentrations of imidazole (5-600 mM). After gel analysis, more pure fractions were selected, concentrated and loaded on SEC chromatography for further purification step.

Fractions containing the fusion proteins were selected on the basis of purity by SDS-PAGE and dialyzed against bicine buffer (20mM Bicine, 150 mM NaCl,with or without 5mM EDTA pH8.0), Protein concentration was determined using DC Protein Assay of BioRad. Proteins were thus pooled, sterile-filtered on 0.22 µm, stored at -80°C.

Alternatively, IMAC purification was preceded by a DEAE purification step using 2mM bicine buffer (pH 8.0) for loading and washing, and eluted using a gradient with the same buffer but with 1M NaCl added.

Example 10

Cloning expression and purification of the separate *C.difficile* Tox A and Tox B fragments

Expression plasmid and recombinant strain.

Genes encoding the protein fragments of ToxA and ToxB (SEQ ID NO:8 and SEQ ID NO:9) and a His tag were cloned into the pET24b(+) expression vector (Novagen) using the NdeI/XhoI restriction sites using standard procedures. The final construct was generated by the transformation of *E. coli* strain BLR (DE3) with the recombinant expression vector according to standard method with CaCl2-treated cells (Hanahan D. « Plasmid transformation by Simanis. » In Glover, D. M. (Ed), DNA cloning. IRL Press London. (1985): p. 109-135.).

Host strain:

BLR(DE3). BLR is a recA derivative of BL21. Strains having the designation (DE3) are lysogenic for a λ prophage that contains an IPTG inducible T7 RNA polymerase. λ DE3 lysogens are designed for protein expression from pET vectors This strain is also deficient in the *lon* and *ompT* proteases.

Genotype : *E.coli* BLR::DE3 strain, F^- *ompT* $hsdS_B(r_B^- m_B^-)$ *gal dcm* (*DE3*) Δ(srl-recA)306::Tn*10* (Tet^R)

Expression of the recombinant proteins:

A *E.coli* transformant was stripped from agar plate and used to inoculate 200 ml of LBT broth ± 1% (w/v) glucose + kanamycin (50 µg/ml) to obtain $O.D._{600nm}$ between 0.1 -0.2. Cultures were incubated overnight at 37 °C, 250 RPM.

This overnight culture was diluted to 1:20 in 500 ml of LBT medium containing kanamycin (50 µg/ml) and grown at 37°C at a stirring speed of 250 rpm until $O.D._{620}$ reached 0.5/0.6.

At an O.D.at 600nm of around 0.6, the culture was cooled down before inducing the expression of the recombinant protein by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; EMD Chemicals Inc., catalogue number: 5815) and incubated overnight at 23 °C, 250 RPM.

After the overnight induction (around 16 hours), O.D. at 600nm was evaluated after induction and culture was centrifuged at 14 000 RPM for 15 minutes and pellets were frozen at -20°C separately.

Purification:

The bacterial pellet was resuspended in 20 mM bicine buffer (pH 8.0) containing 500 mM NaCl supplemented by a mixture of protease inhibitor (Complete without EDTA, Roche cat 11873580001) and benzonase. (Roche cat 1.01695.0001). Bacteria were lysed using a French Press system 2 X 20 000 PSI. Soluble (supernatant) and insoluble (pellet) components were separated by centrifugation at 34 000g or 48 000g for 25-30 min at 4°C. Supernatant was harvested and filtrated on 0.22 µm filter.

The 6-His tagged-protein was purified under native conditions on IMAC. The soluble components were loaded on a GE column (for example 15ml) (Ni loaded) pre-equilibrated with the same buffer used to bacterial resuspension. After loading, the column was washed with the same buffer.

For ToxA:

Elution was performed using a 20mM bicine buffer (pH 8.0) containing 500 mM NaCl and different concentrations of imidazole (5-100 mM). After gel analysis, more pure fractions were selected, concentrated and loaded on SEC chromatography (SUPERDEX^{TM} 75) for further purification step in the same buffer without imidazole.

For ToxB:

A second wash was performed with 20mM bicine buffer (pH 8.0) containing 500 mM NaCl and 0.5 % deoxycholate or same buffer with 150 mM NaCl. Elution was performed using a 20mM bicine buffer (pH 8.0) containing 500 mM NaCl and different concentrations of imidazole (10-500 mM). After gel analysis, more pure fractions were selected, supplemented with 5 mM EDTA and loaded on SEC chromatography (SUPERDEX^{TM} 200) for further purification step in same buffer with 5 mM EDTA.

Fractions containing ToxA or ToxB fragments were selected on the basis of purity by SDS-PAGE and dialyzed against bicine buffer (20mM Bicine, 150 mM NaCl, pH8.0), protein concentration was determined using RCDC Protein Assay of BioRad. Proteins were thus pooled, sterile-filtered on 0.22 µm, stored at -80°C.

Example 11

Molecular weight evaluation of the five *C.difficile* ToxA-ToxB fusions

Analytical ultracentrifugation is used to determine the homogeneity and size distribution in solution of the different species within a protein sample by measuring the rate at which molecules move in response to a centrifugal force. This is based on the calculation of the coefficients of sedimentation of the different species that are obtained by sedimentation velocity experiment, which depend on their molecular shape and mass.

1. Protein samples are spun in a Beckman-Coulter PROTEOMELAB™ XL-1 analytical ultracentrifuge at 42 000RPM after the AN-60Ti rotor had been equilibrated to 15°C.

   a. F1 fusion protein, 500µg/ml, 20mM Bicine, 150mM NaCl, pH8,0
   b. F2 fusion protein, 500µg/ml, 20mM Bicine, 150mM NaCl, pH8,0
   c. F3 fusion protein, 500µg/ml, 20mM Bicine, 150mM NaCl, pH8,0
   d. F4 fusion protein, 500µg/ml, 20mM Bicine, 150mM NaCl, pH8,0
   e. F5 fusion protein, 500µg/ml, 20mM Bicine, 150mM NaCl, pH8,0

2. For data collection, 160 scans were recorded at 280nm every 5 minutes.

3. Data analysis was performed using the program SEDFIT for determination of the C(S) distribution. Determination of the partial specific volume of the proteins was performed with the SEDNTERP software from their amino acid sequence. SEDNTERP was also used to determine the viscosity and the density of the buffer.

4. The molecular weight of the different species was determined from the C(S) distribution plot (concentration vs sedimentation coefficient), considering that it's a better representation of the raw data than the C(M) distribution (concentration vs molecular weight) to characterize the size distribution of a mixture. The molecular weight of the major species detected from the C(S) distribution of all five ToxA-ToxB fusion proteins corresponds to their monomeric form. The best fit frictional ratios determined for the five fusions are all between 2 and 2,2. This may indicate that the proteins are present in solution as an elongated form, which would be consistent with the protein structure.

Example 12

Evaluation of secondary and tertiary structures of *C.difficile* ToxA-ToxB fusions by circular dichroism and fluorescence spectroscopy

Circular dichroism is used to determine the secondary structure composition of a protein by measuring the difference in the absorption of left-handed polarized light versus right-handed polarized light which is due to structural asymmetry. The shape and the magnitude of the CD spectra in the far-UV region (190-250nm) are different whether a protein exhibits a beta-sheet, alpha-helix or random coil structure. The relative abundance of each secondary structure type in a given protein sample can be calculated by comparison to reference spectra.

The tertiary structure of a protein sample can be assessed by the evaluation of the immobilisation of the aromatic amino acids. The observation of a CD signal in the near-UV region (250-50nm) may be attributable to the polarization of phenylalanine, tyrosine and tryptophane residues and is a good indication that the protein is folded into a well defined structure.

The following protocol was used:

1. Far UV spectra are measured using an optical path of 0,01cm from 178 to 250nm, with a 1nm resolution and bandwidth on a Jasco J-720 spectropolarimeter. Temperature of the cell is maintained at 23°C by a Peltier thermostated RTE-111 cell block. A nitrogen flow of 10L/min is maintained during the measurements.

2. Near-UV spectra are measured using an optical path of 0,01cm from 250 to 300nm, with a 1nm resolution and bandwidth on a Jasco J-720 spectropolarimeter. Temperature of the cell is maintained at 23°C by a Peltier thermostated RTE-111 cell block. A nitrogen flow of 6L/min is maintained during the measurements.

The observation of the far-UV spectra (figure 22) for all five ToxA-ToxB fusion proteins suggests a weak content of alpha helix structures and a high content of beta sheet structures. Also, all proteins exhibited a maximum at 230nm, which is unusual for soluble globular proteins. This particularity has been well characterized in the literature and is associated with a small group of proteins known for their absence of alpha helix and their high content in beta sheet and aromatic amino acids (Zsila, Analytical Biochemistry,391(2009) 154-156). Those particularities are coherent with the structure that is expected for the ToxA-ToxB fusion proteins. Crystal structures of 13 proteins exhibiting the characteristic CD spectra with a positive signal at 230nm were compared (Protein Data Bank). The average secondary structure content of those proteins is 42% beta sheet ±9% and 7% alpha helix ±6%. This strongly indicates that the spectral signature of the ToxA-ToxB fusion proteins is diagnostic of a high beta sheet and low alpha helix containing protein.

The observation of the shape of the near-UV spectra (figure 23) for all five fusion proteins indicates that at least some of the aromatic amino acids are immobilised, which is a strong indication of a compact and specific tertiary structure. Moreover, the treatment of the protein with a denaturing concentration of urea caused the disappearance of the near-UV signal, which is an additional indication that this characteristic spectra was due to protein folding.

Example 13

Design, cloning, expression and purification of 4 further fusion proteins

Four further fusion proteins were designed using the design principles described in example 1, these were named F54 Gly (SEQ ID NO:21), F54 New (SEQ ID NO:23), F5 ToxB (SEQ ID NO:25) and F52 New (SEQ ID NO:27).

These fusion proteins were expressed according to the procedure described in example 9.

Example 14

Molecular weight evaluation of the *C.difficile* ToxA-ToxB fusions described in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, and SEQ ID NO: 27

The molecular weight of the fusions described in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, and SEQ ID NO:27 were determined as described in example 11.

The molecular weight of the main species determined from the C(S) distribution of all four protein fusions described in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, and SEQ ID NO:27 corresponds to their monomeric form and all proteins exhibit sedimentation properties similar to F1 to F5 fusions.

Example 15

Evaluation of secondary and tertiary structures of *C.difficile* ToxA-ToxB fusions described in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:1525 and SEQ ID NO:27

The secondary and tertiary structures of the fusions described in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, and SEQ ID NO:27 were assessed according to the method described in example 5. The far UV CD for these fusion proteins can be found in figure 24, and the near UV spectra for these fusions can be found in figure 25.

Analysis of the near and far UV CD spectra of the proteins described in SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, and SEQ ID NO:27 shows that all four have the same high beta sheet structure than F1 to F5 fusions. In addition, observation of the near UV spectra show no significant difference in the position of the aromatic amino acids in the tertiary structure compared to F1 to F5 fusions.

SEQUENCES

SEQ ID NO:1 – sequence of toxin A

MSLISKEELIKLAYSIRPRENEYKTILTNLDEYNKLTTNNNENKYLQLKKLNESIDVFMN
KYKTSSRNRALSNLKKDILKEVILIKNSNTSPVEKNLHFVWIGGEVSDIALEYIKQWADI
NAEYNIKLWYDSEAFLVNTLKKAIVESSTTEALQLLEEEIQNPQFDNMKFYKKRMEFIYD
RQKRFINYYKSQINKPTVPTIDDIIKSHLVSEYNRDETVLESYRTNSLRKINSNHGIDIR
ANSLFTEQELLNIYSQELLNRGNLAAASDIVRLLALKNFGGVYLDVDMLPGIHSDLFKTI
SRPSSIGLDRWEMIKLEAIMKYKKYINNYTSENFDKLDQQLKDNFKLIIESKSEKSEIFS
KLENLNVSDLEIKIAFALGSVINQALISKQGSYLTNLVIEQVKNRYQFLNQHLNPAIESD
NNFTDTTKIFHDSLFNSATAENSMFLTKIAPYLQVGFMPEARSTISLSGPGAYASAYYDF
INLQENTIEKTLKASDLIEFKFPENNLSQLTEQEINSLWSFDQASAKYQFEKYVRDYTGG
SLSEDNGVDFNKNTALDKNYLLNNKIPSNNVEEAGSKNYVHYIIQLQGDDISYEATCNLF
SKNPKNSIIIQRNMNESAKSYFLSDDGESILELNKYRIPERLKNKEKVKVTFIGHGKDEF
NTSEFARLSVDSLSNEISSFLDTIKLDISPKNVEVNLLGCNMFSYDFNVEETYPGKLLLS
IMDKITSTLPDVNKNSITIGANQYEVRINSEGRKELLAHSGKWINKEEAIMSDLSSKEYI
FFDSIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIESSIGDYIYY
EKLEPVKNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSV
RFINKSNGESVYVETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNT
LNAAFFIQSLIDYSSNKDVLNDLSTSVKVQLYAQLFSTGLNTIYDSIQLVNLISNAVNDT
INVLPTITEGIPIVSTILDGINLGAAIKELLDEHDPLLKKELEAKVGVLAINMSLSIAAT
VASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKATSVVNYFNHLSESKKYGPLKT
EDDKILVPIDDLVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIP
SLSIYSAIGIETENLDFSKKIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLY
PGKFYWRFYAFFDYAITTLKPVYEDTNIKIKLDKDTRNFIMPTITTNEIRNKLSYSFDGA
GGTYSLLLSSYPISTNINLSKDDLWIFNIDNEVREISIENGTIKKGKLIKDVLSKIDINK
NKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVAKSYSLLLSGDKNYLISN
LSNTIEKINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFYNDST
LEFNSKDFIAEDINVFMKDDINTITGKYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNES
VYSSYLDFVKNSDGHHNTSNFMNLFLDNISFWKLFGFENINFVIDKYFTLVGKTNLGYVE
FICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTGEDISTSLDFSYEPLYG
IDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINLDSSSFEYKWS
TEGSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSF
NSENELDRDHLGFKIIDNKTYYYDEDSKLVKGLININNSLFYFDPIEFNLVTGWQTINGK
KYYFDINTGAALTSYKIINGKHFYFNNDGVMQLGVFKGPDGFEYFAPANTQNNNIEGQAI
VYQSKFLTLNGKKYYFDNNSKAVTGWRIINNEKYYFNPNNAIAAVGLQVIDNNKYYFNPD
TAIISKGWQTVNGSRYYFDTDTAIAFNGYKTIDGKHFYFDSDCVVKIGVFSTSNGFEYFA
PANTYNNNIEGQAIVYQSKFLTLNGKKYYFDNNSKAVTGLQTIDSKKYYFNTNTAEAATG
WQTIDGKKYYFNTNTAEAATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQ
IGVFKGPNGFEYFAPANTDANNIEGQAILYQNEFLTLNGKKYYFGSDSKAVTGWRIINNK
KYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANNESKMVTGVFKG
PNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNL
NTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFY
FNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVT
GLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIM
QIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDG
NRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAI

RYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGV
DGVKAPGIYG

SEQ ID NO:2 – sequence of toxin B

MSLVNRKQLEKMANVRFRTQEDEYVAILDALEEYHNMSENTVVEKYLKLKDINSLTDIYI
DTYKKSGRNKALKKFKEYLVTEVLELKNNNLTPVEKNLHFVWIGGQINDTAINYINQWKD
VNSDYNVNVFYDSNAFLINTLKKTVVESAINDTLESFRENLNDPRFDYNKFFRKRMEIIY
DKQKNFINYYKAQREENPELIIDDIVKTYLSNEYSKEIDELNTYIEESLNKITQNSGNDV
RNFEEFKNGESFNLYEQELVERWNLAAASDILRISALKEIGGMYLDVDMLPGIQPDLFES
IEKPSSVTVDFWEMTKLEAIMKYKEYIPEYTSEHFDMLDEEVQSSFESVLASKSDKSEIF
SSLGDMEASPLEVKIAFNSKGIINQGLISVKDSYCSNLIVKQIENRYKILNNSLNPAISE
DNDFNTTTNTFIDSIMAEANADNGRFMMELGKYLRVGFFPDVKTTINLSGPEAYAAAYQD
LLMFKEGSMNIHLIEADLRNFEISKTNISQSTEQEMASLWSFDDARAKAQFEEYKRNYFE
GSLGEDDNLDFSQNIVVDKEYLLEKISSLARSSERGYIHYIVQLQGDKISYEAACNLFAK
TPYDSVLFQKNIEDSEIAYYYNPGDGEIQEIDKYKIPSIISDRPKIKLTFIGHGKDEFNT
DIFAGFDVDSLSTEIEAAIDLAKEDISPKSIEINLLGCNMFSYSINVEETYPGKLLLKVK
DKISELMPSISQDSIIVSANQYEVRINSEGRRELLDHSGEWINKEESIIKDISSKEYISF
NPKENKITVKSKNLPELSTLLQEIRNNSNSSDIELEEKVMLTECEINVISNIDTQIVEER
IEEAKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSHFISFEDISETDEGFSIRF
INKETGESIFVETEKTIFSEYANHITEEISKIKGTIFDTVNGKLVKKVNLDTTHEVNTLN
AAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQLFSTGLNTITDAAKVVELVSTALDETID
LLPTLSEGLPIIATIIDGVSLGAAIKELSETSDPLLRQEIEAKIGIMAVNLTTATTAIIT
SSLGIASGFSILLVPLAGISAGIPSLVNNELVLRDKATKVVDYFKHVSLVETEGVFTLLD
DKIMMPQDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHL
SIYDVLEVQKEELDLSKDLMVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEG
EFYWRYFAFIADALITTLKPRYEDTNIRINLDSNTRSFIVPIITTEYIREKLSYSFYGSG
GTYALSLSQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKKGDLIEGILSTLSIEEN
KIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVDLLSKSYKLLISGELKILMLNS
NHIQQKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVYMD
DSKPSFGYYSNNLKDVKVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDES
GVAEILKFMNRKGNTNTSDSLMSFLESMNIKSIFVNFLQSNIKFILDANFIISGTTSIGQ
FEFICDENDNIQPYFIKFNTLETNYTLYVGNRQNMIVEPNYDLDDSGDISSTVINFSQKY
LYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYINEKINVNINDLSIRY
VWSNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEIILSFT
PSYYEDGLIGYDLGLVSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNLITGFVTVG
DDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEG
EAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFN
SDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFA
HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIG
LSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIG
VFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESD
KYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIED
KMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYI
AATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:3 – sequence of Fusion 1

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY

QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVI
YFFGVDGVKAPGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGN
EEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVG
FVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLV
RVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYY
FNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFT
DEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:4 – sequence of Fusion 2

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLNQIGDYK
YYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGN
EEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVG
FVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLV
RVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYY
FNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFT
DEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:5 – sequence of Fusion 3

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQY
YFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNI
YGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRT
GLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGW
LDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:6 – sequence of Fusion 4

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGETIIDDKN
YYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHY
FSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGV
FNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGL
SLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAP
ANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYF
DEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEG
ESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:7 – sequence of Fusion 5

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVI
YFFGVDGVKAPGIYGGGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHN
EDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDG
IMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVE
YSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFE
NNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEK
RYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:8 sequence of individual toxin A fragment

MASTGYTSINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSK
AVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPD
GFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDN
KNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKV
YYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAP

SEQ ID NO:9 – sequence of individual toxin B fragment

MILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEIILSFTPSYYEDGLIGYDLGLVSLYN
EKFYINNFGMMVSGLIYINDSLYYFKPPVNNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSG
VLQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGK
AFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGF
KYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQ
YYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDN
IYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMR
TGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTG
WLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTA

SEQ ID NO:10 – sequence of toxin A fragment from fusion 1

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVI
YFFGVDGVKAP

SEQ ID NO:11 – sequence of toxin A fragment from fusion 2

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAG

SEQ ID NO:12 – sequence of toxin A fragment from fusion 3

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
Y

SEQ ID NO:13 – sequence of toxin A fragment from fusion 4

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAG

SEQ ID NO:14 – sequence of toxin A fragment from fusion 5

MGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQ
NEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNF
YFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKK
YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIM
QIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVA
VTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRY
QNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFE
YFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVI
YFFGVDGVKAPGIYG

SEQ ID NO:15 – sequence of toxin B fragment from fusion 1

GFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGI
LNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYF
SDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGE
TYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGY
INIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVI
IDGEEYYFDPDTAQLVISE

SEQ ID NO:16 – sequence of toxin B fragment from fusion 2

GLNQIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYF
AHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYF
NDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYG
QAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGL
ISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLD
LDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:17 – sequence of toxin B fragment from fusion 3

FAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYY
FNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIY
GQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTG

LISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWL
DLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:18 – sequence of toxin B fragment from fusion 4

GETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEW
KELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFA
ENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDE
DTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDT
SDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGIN
LIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQ
NTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:19 – sequence of toxin B fragment from fusion 5

GFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGI
LNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYF
SDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGE
TYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGY
INIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVI
IDGEEYYFDPDTAQLVISE

SEQ ID NO :20 – nucleotide sequence of F54 Gly fusion protein

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTTAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTATCAAC
GGCAAACACTTTTATTTTAACACCGACGGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACGGCTTTGAATACTTTGCACCGGCAAAT
ACCGATGCCAATAATATTGAAGGCCAGGCCATTCTGTATCAGAATGAATTTCTGACCCTGAACGGCAAAAAATACTACTTTGGCAGCGAT
AGCAAAGCAGTTACCGGTTGGCGCATCATCAACAATAAGAAATATTACTTCAACCCGAATAATGCAATTGCAGCAATTCATCTGTGCACC
ATTAACAACGACAAATATTATTTCAGCTATGACGGTATTCTGCAGAATGGCTACATTACCATCGAACGCAACAACTTTTATTTCGATGCC
AACAACGAAAGCAAAATGGTGACCGGTGTTTTCAAAGGCCCTAATGGTTTTGAGTATTTCGCTCCGGCAAACACCCATAATAACAACATT
GAAGGTCAGGCGATCGTTTATCAGAACAAATTCCTGACGCTGAATGGTAAGAAATACTATTTCGATAATGACAGCAAAGCCGTGACCGGC
TGGCAGACAATTGACGGGAAGAAATATTACTTTAATCTGAATACCGCAGAAGCAGCAACCGGTTGGCAAACGATCGACGGTAAAAAGTAC
TACTTCAACCTGAACACAGCCGAAGCAGCCACAGGATGGCAGACTATTGATGGAAAAAAATACTATTTCAACACCAACACCTTTATTGCA
TCTACCGGTTATACCAGCATTAACGGTAAACATTTCTACTTCAACACCGATGGTATCATGCAGATCGGCGTTTTCAAAGGTCCAAATGGT
TTCGAATACTTTGCCCCTGCCAATACAGATGCAAATAACATCGAGGGTCAGGCAATCCTGTACCAAAACAAATTTCTGACCCTGAATGGG
AAAAAATATTACTTTGGTAGCGATTCTAAAGCCGTTACCGGTCTGCGTACCATTGATGGTAAAAAATACTACTTTAATACGAATACAGCC
GTTGCGGTTACAGGCTGGCAGACCATTAACGGGAAAAAATACTATTTTAACACAAATACCAGCATTGCCTCAACGGGTTATACCATTATT
TCGGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAATCGGAGTCTTTAAAGGACCTGATGGGTTCGAATATTTTGCGCCTGCG
AACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAATCGCTTTCTGTATCTGCACGACAACATCTATTATTTTGGCAAC
AATTCAAAAGCAGCCACCGGCTGGGTTACAATTGATGGCAACCGCTACTATTTCGAACCGAATACCGCAATGGGTGCAAATGGCTACAAA
ACCATCGATAATAAAAATTTCTATTTTCGCAACGGTCTGCCGCAGATCGGGGTATTTAAAGGTAGCAACGGCTTCGAATACTTCGCTCCA
GCGAATACGGACGCGAACAATATTGAGGGTCAAGCGATTCGTTATCAAAACCGTTTTCTGCATCTGCTGGGCAAAATCTACTACTTTGGC
AATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGGATACCGCCATGGCAGCAGCCGGTGGT
CTGTTTGAAATTGATGGCGTGATCTATTTTTTTGGTGTGGATGGTGTTAAAGCACCGGGGAATATACGGTGGTACCGGCTTTGTGACCGTG
GGTGATGATAAATACTATTTCAATCCGATTAACGGTGGTGCAGCGAGCATTGGCGAAACCATCATCGATGACAAAAACTATTATTTCAAC
CAGAGCGGTGTGCTGCAGACCGGTGTGTTTAGCACCGAAGATGGCTTTAAATATTTTGCGCCAGCGAACACCCTGGATGAAAACCTGGAA
GGCGAAGCGATTGATTTTACCGGCAAACTGATCATCGATGAAAACATCTATTACTTCGATGATAACTATCGTGGTGCGGTGGAATGGAAA
GAACTGGATGGCGAAATGCATTATTTTTCTCCGGAAACCGGTAAAGCGTTTAAAGGCCTGAACCAGATCGGCGATTACAAATACTACTTC

AACAGCGATGGCGTGATGCAGAAAGGCTTTGTGAGCATCAACGATAACAAACACTATTTCGATGATAGCGGTGTGATGAAAGTGGGCTAT
ACCGAAATTGATGGCAAACATTTCTACTTCGCGGAAAACGGCGAAATGCAGATTGGCGTGTTCAATACCGAAGATGGTTTCAAATACTTC
GCGCACCATAACGAAGATCTGGGTAACGAAGAAGGCGAAGAAATTAGCTATAGCGGCATCCTGAACTTCAACAACAAAATCTACTACTTT
GATGATAGCTTTACCGCGGTGGTGGGCTGGAAAGATCTGGAAGATGGCAGCAAATATTATTTCGATGAAGATACCGCGGAAGCGTATATT
GGCCTGAGCCTGATTAACGATGGCCAGTACTATTTTAACGATGATGGCATTATGCAGGTGGGTTTCGTGACCATTAATGATAAAGTGTTC
TATTTCAGCGATAGCGGCATTATTGAAAGCGGCGTGCAGAACATTGATGATAACTACTTCTACATCGATGATAACGGCATTGTGCAGATC
GGCGTTTTTGATACCAGCGATGGCTACAAATATTTCGCACCGGCCAATACCGTGAACGATAACATTTATGGCCAGGCGGTGGAATATAGC
GGTCTGGTGCGTGTGGGCGAAGATGTGTATTATTTCGGCGAAACCTATACCATCGAAACCGGCTGGATTTATGATATGGAAAACGAAAGC
GATAAATATTACTTTAATCCGGAAACGAAAAAAGCGTGCAAAGGCATTAACCTGATCGATGATATCAAATACTATTTTGATGAAAAAGGC
ATTATGCGTACCGGTCTGATTAGCTTCGAAAACAACAACTATTACTTCAACGAAAACGGTGAAATGCAGTTCGGCTACATCAACATCGAA
GATAAAATGTTCTACTTCGGCGAAGATGGTGTTATGCAGATTGGTGTTTTTAACACCCCGGATGGCTTCAAATACTTTGCCCATCAGAAT
ACCCTGGATGAAAATTTCGAAGGTGAAAGCATTAACTATACCGGCTGGCTGGATCTGGATGAAAAACGCTACTACTTCACCGATGAATAC
ATTGCGGCGACCGGCAGCGTGATTATTGATGGCGAAGAATACTACTTCGATCCGGATACCGCGCAGCTGGTGATTAGCGAACATCATCAT
CATCACCAT

SEQ ID NO :21 – amino acid of F54Gly fusion protein

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEFLTLN
GKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANNESKMVTGVF
KGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYY
FNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILY
QNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDG
IMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKT
IDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMP
DTAMAAAGGLFEIDGVIYFFGVDGVKAPGIYGGTGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFS
TEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFN
SDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGIL
NFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESG
VQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESD
KYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPD
GFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISEHHHHHH

SEQ ID NO :22 – nucleotide sequence of F54 New fusion protein

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTTAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTATCAAC
GGCAAACACTTTTATTTTAACACCGACGGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACGGCTTTGAATACTTTGCACCGGCAAAT
ACCGATGCCAATAATATTGAAGGCCAGGCCATTCTGTATCAGAATGAATTTCTGACCCTGAACGGCAAAAAATACTACTTTGGCAGCGAT
AGCAAAGCAGTTACCGGTTGGCGCATCATCAACAATAAGAAATATTACTTCAACCCGAATAATGCAATTGCAGCAATTCATCTGTGCACC
ATTAACAACGACAAATATTATTTCAGCTATGACGGTATTCTGCAGAATGGCTACATTACCATCGAACGCAACAACTTTTATTTCGATGCC
AACAACGAAAGCAAAATGGTGACCGGTGTTTTCAAAGGCCCTAATGGTTTTGAGTATTTCGCTCCGGCAAACACCCATAATAACAACATT
GAAGGTCAGGCGATCGTTTATCAGAACAAATTCCTGACGCTGAATGGTAAGAAATACTATTTCGATAATGACAGCAAAGCCGTGACCGGC
TGGCAGACAATTGACGGGAAGAAATATTACTTTAATCTGAATACCGCAGAAGCAGCAACCGGTTGGCAAACGATCGACGGTAAAAAGTAC
TACTTCAACCTGAACACAGCCGAAGCAGCCACAGGATGGCAGACTATTGATGGAAAAAAATACTATTTCAACACCAACACCTTTATTGCA
TCTACCGGTTATACCAGCATTAACGGTAAACATTTCTACTTCAACACCGATGGTATCATGCAGATCGGCGTTTTCAAAGGTCCAAATGGT
TTCGAATACTTTGCCCCTGCCAATACAGATGCAAATAACATCGAGGGTCAGGCAATCCTGTACCAAAACAAATTTCTGACCCTGAATGGG
AAAAAATATTACTTTGGTAGCGATTCTAAAGCCGTTACCGGTCTGCGTACCATTGATGGTAAAAAATACTACTTTAATACGAATACAGCC
GTTGCGGTTACAGGCTGGCAGACCATTAACGGGAAAAAATACTATTTTAACACAAATACCAGCATTGCCTCAACGGGTTATACCATTATT
TCGGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAATCGGAGTCTTTAAAGGACCTGATGGGTTCGAATATTTTGCGCCTGCG
AACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAATCGCTTTCTGTATCTGCACGACAACATCTATTATTTTGGCAAC
AATTCAAAAGCAGCCACCGGCTGGGTTACAATTGATGGCAACCGCTACTATTTCGAACCGAATACCGCAATGGGTGCAAATGGCTACAAA

ACCATCGATAATAAAAATTTCTATTTTCGCAACGGTCTGCCGCAGATCGGGGTATTTAAAGGTAGCAACGGCTTCGAATACTTCGCTCCA
GCGAATACGGACGCGAACAATATTGAGGGTCAAGCGATTCGTTATCAAAACCGTTTTCTGCATCTGCTGGGCAAAATCTACTACTTTGGC
AATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGGATACCGCCATGGCAGCAGCCGGTGGT
CTGTTTGAAATTGATGGCGTGATCTATTTTTTTGGTGTGGATGGTGTTAAAGCAGTTACCGGCTTTGTGACCGTGGGTGATGATAAATAC
TATTTCAATCCGATTAACGGTGGTGCAGCGAGCATTGGCGAAACCATCATCGATGACAAAAACTATTATTTCAACCAGAGCGGTGTGCTG
CAGACCGGTGTGTTTAGCACCGAAGATGGCTTTAAATATTTTGCGCCAGCGAACACCCTGGATGAAAACCTGGAAGGCGAAGCGATTGAT
TTTACCGGCAAACTGATCATCGATGAAAACATCTATTACTTCGATGATAACTATCGTGGTGCGGTGGAATGGAAAGAACTGGATGGCGAA
ATGCATTATTTTTCTCCGGAAACCGGTAAAGCGTTTAAAGGCCTGAACCAGATCGGCGATTACAAATACTACTTCAACAGCGATGGCGTG
ATGCAGAAAGGCTTTGTGAGCATCAACGATAACAAACACTATTTCGATGATAGCGGTGTGATGAAAGTGGGCTATACCGAAATTGATGGC
AAACATTTCTACTTCGCGGAAAACGGCGAAATGCAGATTGGCGTGTTCAATACCGAAGATGGTTTCAAATACTTCGCGCACCATAACGAA
GATCTGGGTAACGAAGAAGGCGAAGAAATTAGCTATAGCGGCATCCTGAACTTCAACAACAAAATCTACTACTTTGATGATAGCTTTACC
GCGGTGGTGGGCTGGAAAGATCTGGAAGATGGCAGCAAATATTATTTCGATGAAGATACCGCGGAAGCGTATATTGGCCTGAGCCTGATT
AACGATGGCCAGTACTATTTTAACGATGATGGCATTATGCAGGTGGGTTTCGTGACCATTAATGATAAAGTGTTCTATTTCAGCGATAGC
GGCATTATTGAAAGCGGCGTGCAGAACATTGATGATAACTACTTCTACATCGATGATAACGGCATTGTGCAGATCGGCGTTTTTGATACC
AGCGATGGCTACAAATATTTCGCACCGGCCAATACCGTGAACGATAACATTTATGGCCAGGCGGTGGAATATAGCGGTCTGGTGCGTGTG
GGCGAAGATGTGTATTATTTCGGCGAAACCTATACCATCGAAACCGGCTGGATTTATGATATGGAAAACGAAAGCGATAAATATTACTTT
AATCCGGAAACGAAAAAAGCGTGCAAAGGCATTAACCTGATCGATGATATCAAATACTATTTTGATGAAAAAGGCATTATGCGTACCGGT
CTGATTAGCTTCGAAAACAACAACTATTACTTCAACGAAAACGGTGAAATGCAGTTCGGCTACATCAACATCGAAGATAAAATGTTCTAC
TTCGGCGAAGATGGTGTTATGCAGATTGGTGTTTTTAACACCCCGGATGGCTTCAAATACTTTGCCCATCAGAATACCCTGGATGAAAAT
TTCGAAGGTGAAAGCATTAACTATACCGGCTGGCTGGATCTGGATGAAAAACGCTACTACTTCACCGATGAATACATTGCGGCGACCGGC
AGCGTGATTATTGATGGCGAAGAATACTACTTCGATCCGGATACCGCGCAGCTGGTGATTAGCGAACATCATCATCATCACCAT

SEQ ID NO :23 amino acid sequence of F54 New fusion protein

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQN
EFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANNESK
MVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAEAATGWQTI
DGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIE
GQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHF
YFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMG
ANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGK
VYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAVTGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVF
STEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYF
NSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGI
LNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIES
GVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENES
DKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTP
DGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISEHHHHHH

SEQ ID NO :24 nucleotide sequence of F5 ToxB fusion protein

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTTAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTATCAAC
GGCAAACACTTTTATTTTAACACCGACGGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACGGCTTTGAATACTTTGCACCGGCAAAT
ACCGATGCCAATAATATTGAAGGCCAGGCCATTCTGTATCAGAATGAATTTCTGACCCTGAACGGCAAAAAATACTACTTTGGCAGCGAT
AGCAAAGCAGTTACCGGTTGGCGCATCATCAACAATAAGAAATATTACTTCAACCCGAATAATGCAATTGCAGCAATTCATCTGTGCACC
ATTAACAACGACAAATATTATTTCAGCTATGACGGTATTCTGCAGAATGGCTACATTACCATCGAACGCAACAACTTTTATTTCGATGCC
AACAACGAAAGCAAAATGGTGACCGGTGTTTTCAAAGGCCCTAATGGTTTTGAGTATTTCGCTCCGGCAAACACCCATAATAACAACATT
GAAGGTCAGGCGATCGTTTATCAGAACAAATTCCTGACGCTGAATGGTAAGAAATACTATTTCGATAATGACAGCAAAGCCGTGACCGGC
TGGCAGACAATTGACGGGAAGAAATATTACTTTAATCTGAATACCGCAGAAGCAGCAACCGGTTGGCAAACGATCGACGGTAAAAAGTAC
TACTTCAACCTGAACACAGCCGAAGCAGCCACAGGATGGCAGACTATTGATGGAAAAAAATACTATTTCAACACCAACACCTTTATTGCA

TCTACCGGTTATACCAGCATTAACGGTAAACATTTCTACTTCAACACCGATGGTATCATGCAGATCGGCGTTTTCAAAGGTCCAAATGGT
TTCGAATACTTTGCCCCTGCCAATACAGATGCAAATAACATCGAGGGTCAGGCAATCCTGTACCAAAACAAATTTCTGACCCTGAATGGG
AAAAAATATTACTTTGGTAGCGATTCTAAAGCCGTTACCGGTCTGCGTACCATTGATGGTAAAAAATACTACTTTAATACGAATACAGCC
GTTGCGGTTACAGGCTGGCAGACCATTAACGGGAAAAAATACTATTTTAACACAAATACCAGCATTGCCTCAACGGGTTATACCATTATT
TCGGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAATCGGAGTCTTTAAAGGACCTGATGGGTTCGAATATTTTGCGCCTGCG
AACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAATCGCTTTCTGTATCTGCACGACAACATCTATTATTTTGGCAAC
AATTCAAAAGCAGCCACCGGCTGGGTTACAATTGATGGCAACCGCTACTATTTCGAACCGAATACCGCAATGGGTGCAAATGGCTACAAA
ACCATCGATAATAAAAATTTCTATTTTCGCAACGGTCTGCCGCAGATCGGGGTATTTAAAGGTAGCAACGGCTTCGAATACTTCGCTCCA
GCGAATACGGACGCGAACAATATTGAGGGTCAAGCGATTCGTTATCAAAACCGTTTTCTGCATCTGCTGGGCAAAATCTACTACTTTGGC
AATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGGATACCGCCATGGCAGCAGCCGGTGGT
CTGTTTGAAATTGATGGCGTGATCTATTTTTTTGGTGTGGATGGTGTTAAAGCAGTGAGCGGTCTGATTTATATTAACGATAGCCTGTAT
TACTTTAAACCACCGGTGAATAACCTGATTACCGGCTTTGTGACCGTGGGTGATGATAAATACTATTTCAATCCGATTAACGGTGGTGCA
GCGAGCATTGGCGAAACCATCATCGATGACAAAAACTATTATTTCAACCAGAGCGGTGTGCTGCAGACCGGTGTGTTTAGCACCGAAGAT
GGCTTTAAATATTTTGCGCCAGCGAACACCCTGGATGAAAACCTGGAAGGCGAAGCGATTGATTTTACCGGCAAACTGATCATCGATGAA
AACATCTATTACTTCGATGATAACTATCGTGGTGCGGTGGAATGGAAAGAACTGGATGGCGAAATGCATTATTTTTCTCCGGAAACCGGT
AAAGCGTTTAAAGGCCTGAACCAGATCGGCGATTACAAATACTACTTCAACAGCGATGGCGTGATGCAGAAAGGCTTTGTGAGCATCAAC
GATAACAAACACTATTTCGATGATAGCGGTGTGATGAAAGTGGGCTATACCGAAATTGATGGCAAACATTTCTACTTCGCGGAAAACGGC
GAAATGCAGATTGGCGTGTTCAATACCGAAGATGGTTTCAAATACTTCGCGCACCATAACGAAGATCTGGGTAACGAAGAAGGCGAAGAA
ATTAGCTATAGCGGCATCCTGAACTTCAACAACAAAATCTACTACTTTGATGATAGCTTTACCGCGGTGGTGGGCTGGAAAGATCTGGAA
GATGGCAGCAAATATTATTTCGATGAAGATACCGCGGAAGCGTATATTGGCCTGAGCCTGATTAACGATGGCCAGTACTATTTTAACGAT
GATGGCATTATGCAGGTGGGTTTCGTGACCATTAATGATAAAGTGTTCTATTTCAGCGATAGCGGCATTATTGAAAGCGGCGTGCAGAAC
ATTGATGATAACTACTTCTACATCGATGATAACGGCATTGTGCAGATCGGCGTTTTTGATACCAGCGATGGCTACAAATATTTCGCACCG
GCCAATACCGTGAACGATAACATTTATGGCCAGGCGGTGGAATATAGCGGTCTGGTGCGTGTGGGCGAAGATGTGTATTATTTCGGCGAA
ACCTATACCATCGAAACCGGCTGGATTTATGATATGGAAAACGAAAGCGATAAATATTACTTTAATCCGGAAACGAAAAAAGCGTGCAAA
GGCATTAACCTGATCGATGATATCAAATACTATTTTGATGAAAAAGGCATTATGCGTACCGGTCTGATTAGCTTCGAAAACAACAACTAT
TACTTCAACGAAAACGGTGAAATGCAGTTCGGCTACATCAACATCGAAGATAAAATGTTCTACTTCGGCGAAGATGGTGTTATGCAGATT
GGTGTTTTTAACACCCCGGATGGCTTCAAATACTTTGCCCATCAGAATACCCTGGATGAAAATTTCGAAGGTGAAAGCATTAACTATACC
GGCTGGCTGGATCTGGATGAAAAACGCTACTACTTCACCGATGAATACATTGCGGCGACCGGCAGCGTGATTATTGATGGCGAAGAATAC
TACTTCGATCCGGATACCGCGCAGCTGGTGATTAGCGAACATCATCATCATCACCAT

SEQ ID NO :25 amino acid sequence of F5 ToxB fusion protein

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQN
EFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANNESK
MVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAEAATGWQTI
DGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIE
GQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHF
YFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMG
ANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGK
VYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAVSGLIYINDSLYYFKPPVNNLITGFVTVGDDKYYFNPINGGAASIGE
TIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYF
SPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKY
FAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQ
VGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDV
YYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINI
EDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPD
TAQLVISEHHHHHH

SEQ ID NO :26 - nucleotide sequence of F52 new fusion protein

ATGGCAACCGGTTGGCAGACCATCGATGGCAAAAAATATTATTTTAATACCAACACCGCAATTGCAAGCACCGGCTATACCATTATCAAC
GGCAAACACTTTTATTTTAACACCGACGGCATTATGCAGATTGGTGTGTTTAAAGGTCCGAACGGCTTTGAATACTTTGCACCGGCAAAT
ACCGATGCCAATAATATTGAAGGCCAGGCCATTCTGTATCAGAATGAATTTCTGACCCTGAACGGCAAAAAATACTACTTTGGCAGCGAT
AGCAAAGCAGTTACCGGTTGGCGCATCATCAACAATAAGAAATATTACTTCAACCCGAATAATGCAATTGCAGCAATTCATCTGTGCACC
ATTAACAACGACAAATATTATTTCAGCTATGACGGTATTCTGCAGAATGGCTACATTACCATCGAACGCAACAACTTTTATTTCGATGCC
AACAACGAAAGCAAAATGGTGACCGGTGTTTTCAAAGGCCCTAATGGTTTTGAGTATTTCGCTCCGGCAAACACCCATAATAACAACATT
GAAGGTCAGGCGATCGTTTATCAGAACAAATTCCTGACGCTGAATGGTAAGAAATACTATTTCGATAATGACAGCAAAGCCGTGACCGGC
TGGCAGACAATTGACGGGAAGAAATATTACTTTAATCTGAATACCGCAGAAGCAGCAACCGGTTGGCAAACGATCGACGGTAAAAAGTAC
TACTTCAACCTGAACACAGCCGAAGCAGCCACAGGATGGCAGACTATTGATGGAAAAAAATACTATTTCAACACCAACACCTTTATTGCA
TCTACCGGTTATACCAGCATTAACGGTAAACATTTCTACTTCAACACCGATGGTATCATGCAGATCGGCGTTTTCAAAGGTCCAAATGGT
TTCGAATACTTTGCCCCTGCCAATACAGATGCAAATAACATCGAGGGTCAGGCAATCCTGTACCAAAACAAATTTCTGACCCTGAATGGG
AAAAAATATTACTTTGGTAGCGATTCTAAAGCCGTTACCGGTCTGCGTACCATTGATGGTAAAAAATACTACTTTAATACGAATACAGCC
GTTGCGGTTACAGGCTGGCAGACCATTAACGGGAAAAAATACTATTTTAACACAAATACCAGCATTGCCTCAACGGGTTATACCATTATT
TCGGGTAAACACTTCTACTTTAATACCGATGGTATTATGCAAATCGGAGTCTTTAAAGGACCTGATGGGTTCGAATATTTTGCGCCTGCG
AACACTGATGCGAACAATATCGAAGGACAGGCAATCCGCTATCAGAATCGCTTTCTGTATCTGCACGACAACATCTATTATTTTGGCAAC
AATTCAAAAGCAGCCACCGGCTGGGTTACAATTGATGGCAACCGCTACTATTTCGAACCGAATACCGCAATGGGTGCAAATGGCTACAAA
ACCATCGATAATAAAAATTTCTATTTTCGCAACGGTCTGCCGCAGATCGGGGTATTTAAAGGTAGCAACGGCTTCGAATACTTCGCTCCA
GCGAATACGGACGCGAACAATATTGAGGGTCAAGCGATTCGTTATCAAAACCGTTTTCTGCATCTGCTGGGCAAAATCTACTACTTTGGC
AATAACAGTAAAGCAGTTACTGGATGGCAGACAATCAATGGTAAAGTGTACTATTTTATGCCGGATACCGCCATGGCAGCAGCCGGTGGT
CTGTTTGAAATTGATGGCGTGATCTATTTTTTTGGTGTGGATGGTGTTAAAGCAGTGAAAGGCCTGAACCAGATCGGCGATTACAAATAC
TACTTCAACAGCGATGGCGTGATGCAGAAAGGCTTTGTGAGCATCAACGATAACAAACACTATTTCGATGATAGCGGTGTGATGAAAGTG
GGCTATACCGAAATTGATGGCAAACATTTCTACTTCGCGGAAAACGGCGAAATGCAGATTGGCGTGTTCAATACCGAAGATGGTTTCAAA
TACTTCGCGCACCATAACGAAGATCTGGGTAACGAAGAAGGCGAAGAAATTAGCTATAGCGGCATCCTGAACTTCAACAACAAAATCTAC
TACTTTGATGATAGCTTTACCGCGGTGGTGGGCTGGAAAGATCTGGAAGATGGCAGCAAATATTATTTCGATGAAGATACCGCGGAAGCG
TATATTGGCCTGAGCCTGATTAACGATGGCCAGTACTATTTTAACGATGATGGCATTATGCAGGTGGGTTTCGTGACCATTAATGATAAA
GTGTTCTATTTCAGCGATAGCGGCATTATTGAAAGCGGCGTGCAGAACATTGATGATAACTACTTCTACATCGATGATAACGGCATTGTG
CAGATCGGCGTTTTTGATACCAGCGATGGCTACAAATATTTCGCACCGGCCAATACCGTGAACGATAACATTTATGGCCAGGCGGTGGAA
TATAGCGGTCTGGTGCGTGTGGGCGAAGATGTGTATTATTTCGGCGAAACCTATACCATCGAAACCGGCTGGATTTATGATATGGAAAAC
GAAAGCGATAAATATTACTTTAATCCGGAAACGAAAAAAGCGTGCAAAGGCATTAACCTGATCGATGATATCAAATACTATTTTGATGAA
AAAGGCATTATGCGTACCGGTCTGATTAGCTTCGAAAACAACAACTATTACTTCAACGAAAACGGTGAAATGCAGTTCGGCTACATCAAC
ATCGAAGATAAAATGTTCTACTTCGGCGAAGATGGTGTTATGCAGATTGGTGTTTTTAACACCCCGGATGGCTTCAAATACTTTGCCCAT
CAGAATACCCTGGATGAAAATTTCGAAGGTGAAAGCATTAACTATACCGGCTGGCTGGATCTGGATGAAAAACGCTACTACTTCACCGAT
GAATACATTGCGGCGACCGGCAGCGTGATTATTGATGGCGAAGAATACTACTTCGATCCGGATACCGCGCAGCTGGTGATTAGCGAACAT
CATCATCATCACCAT

SEQ ID NO :27 – amino acid sequence of F52 New fusion protein

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEFLTLN
GKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANNESKMVTGVF
KGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYY
FNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILY
QNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDG
IMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKT
IDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMP
DTAMAAAGGLFEIDGVIYFFGVDGVKAVKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFY
FAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAY
IGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVN
DNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISF
ENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDE

YIAATGSVIIDGEEYYFDPDTAQLVISEHHHHHH

SEQ ID NO:28 – sequence of toxin A fragment of F54 Gly

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAIL
YQNEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERN
NFYFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDG
KKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDG
IMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTA
VAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAI
RYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG
FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDG
VIYFFGVDGVKAPGIYG

SEQ ID NO:29 – sequence of toxin A fragment of F54 New

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAIL
YQNEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERN
NFYFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDG
KKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDG
IMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTA
VAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAI
RYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG
FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDG
VIYFFGVDGVKAV

SEQ ID NO:30 – sequence of toxin A fragment of F5 ToxB

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAIL
YQNEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERN
NFYFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDG
KKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDG
IMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTA
VAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAI
RYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG
FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDG
VIYFFGVDGVKAV

SEQ ID NO:31 – sequence of toxin A fragment of F52 New

MATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAIL
YQNEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERN
NFYFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDG
KKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDG
IMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTA

VAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAI
RYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNG
FEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDG
VIYFFGVDGVKAV

SEQ ID NO:32 – sequence of toxin B fragment of F54Gly

TGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEGEAIDF
TGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKH
YFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYF
DDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGV
QNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIY
DMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYF
GEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDP
DTAQLVISE

SEQ ID NO:33 – sequence of toxin B fragment of F54 New

TGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEGEAIDF
TGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKH
YFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYF
DDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGV
QNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIY
DMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYF
GEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDP
DTAQLVISE

SEQ ID NO:34 – sequence of toxin B fragment of F5 ToxB

SGLIYINDSLYYFKPPVNNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDG
FKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYK
YYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGN
EEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVG
FVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLV
RVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYY
FNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFT
DEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:35 – sequence of toxin B fragment of F52 New

KGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKY
FAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYY
FNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIY
GQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTG

LISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWL
DLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:36 – C-TAB.G5 fusion protein

MVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAE
AATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPN
GFEYFAPANTHNNNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTING
KKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHD
NIYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDA
NNIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVK
APGIYGRSMHNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPAN
TLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGV
MQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISY
SGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKV
FYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYY
FGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNEN
GEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYI
AATGSVIIDGEEYYFDPDTAQLVISE

SEQ ID NO:37 C-TAB.G5.1 fusion protein

VTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAEA
ATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPNG
FEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGK
KYYFNTNTSIASTGYTIISGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDN
IYYFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDAN
NIEGQAIRYQNRFLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKA
PGIYGRSMHNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANT
LDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVM
QKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYS
GILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVF
YFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYF
GETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQF
GYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGS
VIIDGEEYYFDPDTAQLVISE

Claims

1. An Immunogenic composition comprising:
   a) a polypeptide comprising an isolated *Clostridium difficile* toxin A fragment and / or an isolated *C. difficile* toxin B fragment; and
   b) an adjuvant comprising an immunologically active saponin fraction presented in the form of a liposome.

2. An immunogenic composition according to claim 2 in which said adjuvant comprises a lipopolysaccharide.

3. An immunogenic composition according to claim 1 or claim 2 wherein said adjuvant comprises a sterol, wherein the composition comprises the sterol at a ratio of saponin: sterol of from 1:1 to 1:100 w/w.

4. An immunogenic composition according to claim 3 wherein the ratio of saponin : sterol is from 1:1 to 1:10 w/w.

5. An immunogenic composition according to claim 4 wherein the ratio of saponin : sterol is from 1:1 to 1:5 w/w.

6. An immunogenic composition according to any of claims 3 to 5 wherein said sterol is cholesterol.

7. An immunogenic composition according to any of claims 1 to 6 wherein said immunologically active saponin fraction is QS21.

8. An immunogenic composition according to any one of claims 2 to 7 wherein said lipopolysaccharide is a lipid A derivative.

9. An immunogenic composition according to claim 8 wherein said lipid A derivative is 3D-MPL.

10. An immunogenic composition according to claim 9 wherein said immunologically active saponin fraction is QS21 and the ratio QS21 : 3D-MPL is 1 : 1 (w/w).

11. An immunogenic composition according to any one of claims 2 to 10 wherein said lipopolysaccharide is present at an amount of about 1 – about 70 µg per dose.

12. An immunogenic composition according to claim 11 wherein said lipopolysaccharide is present at an amount of about 50 µg.

13. An immunogenic composition according to any preceding claim wherein said saponin is present at an amount of about 1 – about 70 µg per dose.

14. An immunogenic composition according to claim 13 wherein said saponin is present at an amount of about 50 µg.

15. The immunogenic composition of any of claims 1 to 14, wherein said adjuvant comprises 1, 2-Dioleoyl-sn-Glycero-3-phosphocholine (DOPC).

16. The immunogenic composition of any of claims 1 to 15, wherein said adjuvant comprises cholesterol, DOPC, 3D-MPL and QS21.

17. The immunogenic composition of claim 16, wherein the adjuvant is formulated in a dose comprising:

from about 0.1 to about 0.5 mg cholesterol,

from about 0.25 to about 2 mg DOPC,

from about 10 µg to about 70 µg 3D-MPL, and

from about 10 µg to about 70 µg QS21.

18. The immunogenic composition of claim 17, wherein the adjuvant is formulated in a dose comprising:

about 0.25 mg cholesterol,

about 1.0 mg DOPC,

about 50 µg 3D-MPL, and

about 50 µg QS21.

19. An immunogenic composition comprising:
a) A polypeptide comprising an isolated *Clostridium difficile* toxin A fragment and / or an isolated *C. difficile* toxin B fragment; and
b) an adjuvant comprising an oil in water emulsion, wherein said oil in water emulsion comprises a metabolisable oil, a tocol and an emulsifier.

20. An immunogenic composition as claimed in claim 19 wherein the immunogenic composition is not 3µg or 10 µg of any of SEQ ID Nos. 1 to 7 combined with an adjuvant comprising or consisting of 0.125 mL SB62 emulsion (Total volume), 5.35 mg squalene, 5.94 mg DL-α-tocopherol and 2.425 mg polysorbate 80 per 0.5 ml dose.

21. An immunogenic composition according to claim 19 or claim 20 wherein the oil in water emulsion comprises 1-10, 2-10, 3-9, 4-8. 5-7, or 5-6 mg metabolisable oil, per dose.

22. An immunogenic composition according to any one of claims 19 to 21 wherein the oil in water emulsion comprises 0.5-11, 1-11, 2-10, 3-9, 4-8, 5-7, 5-6 mg tocol per dose.

23. An immunogenic composition according to any one of claims 19 to 22 wherein the oil in water emulsion comprises 0.1-5, 0.2-5, 0.3-4, 0.4-3 or 2-3 mg emulsifying agent per dose.

24. An immunogenic composition according to any one of claims 19 to 23 wherein composition comprises metabolisable oil at an amount of about 5.35 mg per dose.

25. An immunogenic composition according to any one of claims 19 to 23 wherein the amount of metabolisable oil is about 2.14 mg per dose.

26. An immunogenic composition according to any one of claims 19 to 25 wherein the amount of tocol is about 5.94 mg per dose.

27. An immunogenic composition according to any one of claims 19 to 25 wherein the amount of tocol is about 2.38 mg per dose.

28. An immunogenic composition according to any one of claims 19 to 27 wherein the amount of emulsifying agent is about 2.425 mg per dose.

29. An immunogenic composition according to any one of claims 19 to 28 wherein the amount of emulsifying agent is 0.97 mg per dose.

30. An immunogenic composition according to any one of claims 19 to 29 wherein the metabolisable oil is squalene.

31. An immunogenic composition as claimed in one of claims 19 to 30 wherein the tocol is alpha-tocopherol.

32. An immunogenic composition as claimed in any one of claims 19 to 31 wherein the emulsifying agent is polyoxyethylene sorbitan monooleate.

33. An immunogenic composition as claimed in claim 32 wherein the polyoxyethylene sorbitan monooleate is selected from the group comprising: Polysorbate® 80 or Tween® 80.

34. An immunogenic composition according to any of the preceding claims wherein the immunogenic composition has a volume of 0.5 to 1.5 ml.

35. An immunogenic composition according to any of the preceding claims wherein the immunogenic composition has a volume of 0.5 ml.

36. An immunogenic composition according to any one of claims 19 to 35 wherein the immunogenic composition has a volume of 0.7 ml.

37. An immunogenic composition according to any one of claims 19 to 35 wherein the the immunogenic composition has a volume of 1.0 ml.

38. An immunogenic composition as claimed in any one of the preceding claims wherein the polypeptide comprises a *Clostridium difficile* toxin A repeating domain fragment and / or a *C. difficile* toxin B repeating domain fragment.

39. An immunogenic composition as claimed in any one of the preceding claims wherein the polypeptide comprises a fragment of the N-terminal domain of toxin A and / or the N-terminal domain of toxin B of *Clostridium difficile*.

40. An immunogenic composition as claimed in any one of the preceding claims wherein the polypeptide in the immunogenic composition has been chemically detoxified.

41. An immunogenic composition as claimed in claim 40 wherein the chemical detoxification is by incubation in formaldehyde.

42. An immunogenic composition according to any preceding claim wherein the polypeptide comprises a toxin A repeating domain fragment and a toxin B repeating domain fragment.

43. An immunogenic composition according to any preceding claim wherein the polypeptide elicits antibodies that neutralize toxin A or toxin B or both.

44. An immunogenic composition according to any preceding claim wherein the polypeptide is a polypeptide comprising a first fragment and a second fragment, wherein

(i) the first fragment is a toxin A repeating domain fragment;
(ii) the second fragment is a toxin B repeating domain fragment;
(iii) the first fragment has a first proximal end;
(iv) the second fragment has a second proximal end; and

wherein the first fragment and the second fragment are adjacent to one another.

45. An immunogenic composition according to claim 44 wherein the first proximal end is within a short repeat.

46. An immunogenic composition according to claim 44 wherein the second proximal end is within a short repeat.

47. An immunogenic composition according to any one of claims 44 to 46 wherein the first proximal end does not disrupt a short repeat-long repeat-short repeat (SR-LR-SR) portion.

48. An immunogenic composition according to any one of claims 44 to 47 wherein the second proximal end does not disrupt a short repeat-long repeat-short repeat (SR-LR-SR) portion.

49. An immunogenic composition according to any one of claims 44 to 46 wherein the first proximal end and the second proximal end do not disrupt short repeat-long repeat-short repeat (SR-LR-SR) portions.

50. An immunogenic composition according to any one of claims 44 to 45 wherein the first proximal end is not within amino acids 1878-1940, 2012-2074, 2146-2208, 2258-2322, 2394-2456, 2507-2569 or 2598-2660 of toxin A (SEQ ID NO. 1).

51. An immunogenic composition according to any one of claims 44 to 46 wherein the second proximal end is not within amino acids 1881-1942, 2012-2074, 2144-2208 or 2278-2339 of toxin B (SEQ ID NO. 2).

52. An immunogenic composition according to any one of claims 44 to 51 wherein the first proximal end is within repeat portion VIII (amino acids 2645-2710) of toxin A.

53. An immunogenic composition according to any one of claims 44 to 52 wherein the first proximal end is within amino acids 2700-2710 or 2680-2690 of toxin A.

54. An immunogenic composition according to any one of claims 44 to 53 wherein the second proximal end is within repeat portion I (amino acids 1834-1926) of toxin B.

55. An immunogenic composition according to any one of claims 44 to 54 wherein the second proximal end is within amino acids 1860-1878 or 1854-1876 of toxin B.

56. An immunogenic composition according to any one of claims 44 to 53 wherein the second proximal end is within repeat portion II (amino acids 1927-2057) of toxin B.

57. An immunogenic composition according to any one of claims 44 to 53 or 56 wherein the second proximal end is within amino acids 1960-1970, 1988-1998 or 1867-1877 of toxin B.

58. An immunogenic composition according to any one of claims 44 to 51 and 54 to 57 wherein the first fragment consists of an entire toxin A repeating domain (amino acids 1832-2710).

59. An immunogenic composition according to any one of claims 44 to 53 and 58 wherein the second fragment consists of an entire toxin B repeating domain (amino acids 1833-2366).

60. An immunogenic composition according to any one of claims 52, 53, 54 or 55 wherein the first proximal end is within repeat portion VIII (amino acids 2645-2710) of toxin A and wherein the second proximal end is within repeat portion I (amino acids 1834-1926) of toxin B.

61. An immunogenic composition according to any one of claims 52, 53, 56 or 57 wherein the first proximal end is within repeat portion VIII (amino acids 2645-2710) of toxin A and wherein the second proximal end is within repeat portion II (amino acids 1927-2057) of toxin B.

62. An immunogenic composition according to any one of claims 52, 53, 56 or 57 wherein the first proximal end is within short repeat 3 of repeat portion VIII of toxin A (amino acids

2687-2710) and the second proximal end is within short repeat 4 of repeat portion II of toxin B (amino acids 1988-2007).

63. An immunogenic composition according to any one of claims 52, 53, 56 or 57 wherein the first proximal end is within short repeat 2 of repeat portion VIII (amino acids 2665-2686) of toxin A and the second proximal end is within short repeat 3 of repeat portion II of toxin B (amino acids 1968-1987).

64. An immunogenic composition according to any one of claims 52, 53, 54 or 55 wherein the first proximal is within short repeat 2 of repeat portion VIII of toxin A (amino acids 2665-2686) and the second proximal end is within short repeat 3 of repeat portion I of toxin B (1877-1896).

65. An immunogenic composition according to any one of claims 44 to 51, 52, 53, 54, 55 and 60 wherein the first proximal end is within short repeat 3 of repeat portion VIII of toxin A (amino acids 2645-2686) and the second proximal end is within short repeat 1 of repeat portion I of toxin B (amino acids 1834-1854).

66. An immunogenic composition according to any one of claims 52, 53, 56 or 57 wherein the first proximal end is within short repeat 3 of repeat portion VIII of toxin A (amino acids 2687-2710) and the second proximal end is within short repeat 4 of repeat portion II of toxin B (amino acids 1988-2007).

67. An immunogenic composition according to claim 61 wherein the first proximal end is within amino acids 2700-2710 of toxin A and the second proximal end is within amino acids 1960-1970 of toxin B.

68. An immunogenic composition according to claim 61 wherein the first proximal end is within amino acids 2680-2690 of toxin A and the second proximal end is within amino acids 1960-1970 of toxin B.

69. An immunogenic composition according to claim 61wherein the first proximal end is within amino acids 2700-2710 of toxin A and the second proximal end is within amino acids 1988-1998 of toxin B.

70. An immunogenic composition according to claim 61 wherein the first proximal end is within amino acids 2680-2690 and the second proximal end is within amino acids 1860-1878.

71. An immunogenic composition according to any preceding claim wherein the polypeptide comprises an immunogenic fragment of 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 or SEQ ID NO:27.

72. An immunogenic composition according to any preceding claim wherein the polypeptide comprises an immunogenic fragment of at least 500, 600, 700, 750, 800, 850 or 900 amino acids of 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 or SEQ ID NO:27.

73. An immunogenic composition according to any preceding claim wherein the polypeptide comprises

(i) 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 or SEQ ID NO:27; or

a variant having at least 90%, 95%, 98%, 99% or 100% similarity to 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 or SEQ ID NO:27.

74. An immunogenic composition according to any one of claims 1 to 37 wherein the polypeptide comprises:
(i) 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 or SEQ ID NO:34 or SEQ ID NO:35; or

(ii) A variant having at least 90%, 95%, 98%, 99% or 100% similarity to 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 or SEQ ID NO:34 or SEQ ID NO:35; or

(iii) A fragment of at least 250, 280, 300, 350, 380, 400, 430, 450, 480, 500, 530, 550, 580, or 600 amino acids of 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 or SEQ ID NO:34 or SEQ ID NO:35.

75. An immunogenic composition according to any one of claims 1 to 37 wherein the polypeptide is a polypeptide comprising a first fragment and a second fragment, wherein
(i) the first fragment is a toxin A repeating domain fragment;
(ii) the second fragment is a toxin B repeating domain fragment;
(iii) the first fragment comprises a first proximal end within a first repeat portion;
(iv) the second fragment comprises a second proximal end within a second repeat portion; and

wherein the first fragment and the second fragment are adjacent to one another and wherein the first repeat portion and the second repeat portion have structural similarity to one another.

76. An immunogenic composition according to claim 75 wherein the first proximal end is within repeat portion VII of toxin A.

77. An immunogenic composition according to claim 75 or claim 76 wherein the second proximal end is within repeat portion II of toxin B.

78. An immunogenic composition according to any one of claims 75 to 77 wherein the first proximal end is within a long repeat.

79. An immunogenic composition according to any one of claims 75 to 78 wherein the second proximal end is within a long repeat.

80. An immunogenic composition according to any one of claims 75 to 79 wherein the first proximal end is within long repeat VII of toxin A (amino acids 2614-2644).

81. An immunogenic composition according to claim 80 wherein the first proximal end is within amino acids 2620 and 2640 of toxin A.

82. An immunogenic composition according to any one of claims 75 to 81 wherein the second proximal end is within long repeat II of toxin B (amino acids 2028-2057).

83. An immunogenic composition according to claim 82 wherein the first proximal end is within amino acids 2030 and 2050 of toxin B.

84. An immunogenic composition according to any one of the preceding claims wherein the polypeptide further comprises a linker.

85. An immunogenic composition according to claim 84 wherein the linker comprises 1-20 amino acids.

86. An immunogenic composition according to any one of claims 84 or 85 wherein the linker is a glycine linker.

87. An immunogenic composition according to any one of claims 84 to 86 wherein the linker is between the proximal end of the first fragment and the proximal end of the second fragment.

88. An immunogenic composition according to any preceding claim wherein the polypeptide is part of a larger fusion protein.

89. An immunogenic composition according to any preceding claim wherein the polypeptide comprises more than 450, 475, 500, 525, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825 or 850 amino acids from toxin A.

90. An immunogenic composition according to any preceding claim wherein the polypeptide comprises less than 850, 825, 800, 775, 750, 725, 700, 675, 650, 625, or 600 amino acids from toxin A.

91. An immunogenic composition according to any preceding claim wherein the polypeptide comprises more than 350, 375, 400, 425, 450, 475, 500 or 525 amino acids from toxin B.

92. An immunogenic composition according to any preceding claim wherein the polypeptide comprises less than 525, 500, 475, or 450 amino acids from toxin B.

93. An immunogenic composition according to any one of the preceding claims wherein the polypeptide elicits a protective immune response in a mammalian host against strains of *C.difficile*.

94. The immunogenic composition of any one of the preceding claims wherein the polypeptide comprises less than 25%, 20%, 18% or 15% alpha helical structure.

95. The immunogenic composition of any one of the preceding claims wherein the polypeptide comprises more than 25%, 30%, 35%, 38% or 40% beta sheet structure.

96. The immunogenic composition of any one of the preceding claims further comprising additional antigens.

97. The immunogenic composition of claim 96 wherein the additional antigens are antigens derived from a bacterium selected from the group consisting of *Streptococcus pneumonia, Haemophilus influenzae*, *Neisseria meningitidis*, *Escherichia coli*, *Moraxella catarrhalis*, *Clostridium tetani, Corynebacterium diptherieriae, Bordetella pertussis, Staphylococcus epidermidis*, enterococci, and *Staphylococcus aureus* .

98. The immunogenic composition of any one of the preceding claims further comprising a saccharide from *C.difficile*.

99. A vaccine comprising the immunogenic composition of any one of the preceding claims and a pharmaceutically acceptable excipient.

100. A use of the immunogenic composition of any one of the preceding claims or the vaccine of claim 99 in the treatment or prevention of *C.difficile* disease.

101. The immunogenic composition of any one of the preceding claims or the vaccine of claim 99 for use in the treatment or prevention of *C.difficile* disease.

102. A use of the immunogenic composition of any one of the preceding claims or the vaccine of claim 99 in the preparation of a medicament for the prevention or treatment of *C.difficile* disease.

103. A method of preventing or treating *C.difficile* disease comprising administering the immunogenic composition of any one of the preceding claims or the vaccine of claim 99 to a mammalian subject.