Clostridium difficile Antigens

The present invention relates to antigens for the prevention/ treatment/ suppression of *Clostridium difficile* infection (CDI). Also provided are methods for generating said antigens, methods for generating antibodies that bind to said antigens, and the use of said antibodies for the prevention/ treatment/ suppression of CDI.

Clostridium difficile infection (CDI) is now a major problem in hospitals worldwide. The bacterium causes nosocomial, antibiotic-associated disease which manifests itself in several forms ranging from mild self-limiting diarrhoea to potentially life-threatening, severe colitis. Elderly patients are most at risk from these potentially life-threatening diseases and incidents of CDI have increased dramatically over the last 10 years. In 2010 in the UK there were over 21,000 cases of CDI with over 2,700 associated deaths. CDI costs the UK National Health Service in excess of £500M per annum.

The various strains of *C. difficile* may be classified by a number of methods. One of the most commonly used is polymerase chain reaction (PCR) ribotyping in which PCR is used to amplify the 16S-23S rRNA gene intergenic spacer region of *C. difficile*. Reaction products from this provide characteristic band patterns identifying the bacterial ribotype of isolates. Toxinotyping is another typing method in which the restriction patterns derived from DNA coding for the *C. difficile* toxins are used to identify strain toxinotype. The differences in restriction patterns observed between toxin genes of different strains are also indicative of sequence variation within the *C. difficile* toxin family. For example, there is an approximate 13% sequence difference with the C-terminal 60kDa region of toxinotype 0 Toxin B compared to the same region in toxinotype III Toxin B.

Strains of *C. difficile* produce a variety of virulence factors, notable among which are several protein toxins: Toxin A, Toxin B and, in some strains, a binary toxin which is similar to *Clostridium perfringens* iota toxin. Toxin A is a large protein cytotoxin/ enterotoxin which plays a role in the pathology of infection and may influence the gut colonisation process. Outbreaks of CDI have been reported with Toxin A-

negative/Toxin B-positive strains, which indicate that Toxin B is also capable of playing a key role in the disease pathology.

The genetic sequences encoding Toxin A and Toxin B (Mw 308k and Mw 269k, respectively) are known - see, for example, Moncrief *et al.* (1997) Infect. Immun 63: 1105-1108. The two toxins have high sequence homology and are believed to have arisen from gene duplication. The toxins also share a common structure (see Figure 1), namely an N-terminal glucosyl transferase domain, a central hydrophobic region, four conserved cysteines, and a long series of C-terminal repeating units (RUs).

Both Toxins A and B exert their mechanisms of action via multi-step mechanisms, which include binding to receptors on the cell surface, internalisation followed by translocation and release of the effector domain into the cell cytosol, and finally intracellular action. Said mechanism of action involves the inactivation of small GTPases of the Rho family. In this regard, the toxins catalyse the transfer of a glucose moiety (from UDP-glucose) onto an amino residue of the Rho protein. Toxins A and B also contain a second enzyme activity in the form of a cysteine protease, which appears to play a role in the release of the effector domain into the cytosol after translocation. The *C. difficile* binary toxin modifies cell actin by a mechanism which involves the transfer of an ADP-ribose moiety from NAD onto its target protein.

Current therapies for the treatment of *C. difficile* infection rely on the use of antibiotics, notably metronidazole and vancomycin. However, these antibiotics are not effective in all cases and 20-30% of patients suffer relapse of the disease. Of major concern is the appearance in the UK of more virulent strains, which were first identified in Canada in 2002. These strains, which include those belonging to PCR ribotype 027 and toxinotype III, cause CDI with a directly attributable mortality more than 3-fold that observed previously.

New therapeutics are therefore required especially urgently since the efficacy of current antibiotics appears to be decreasing.

One approach is the use of antibodies which bind to and neutralise the activity of Toxin A and/ or Toxin B. This is based on the knowledge that strains of *C. difficile* that do not release these toxins, so called non-toxigenic strains, do not cause CDI. By way of example, animals can be immunised, their sera collected and the antibodies purified for administration to patients – this is defined as passive immunisation. In another approach patients with CDI or subjects at risk of developing such infections can be immunised with antigens which result in an increase in circulating and mucosal antibodies directed against Toxin A and/ or Toxin B - this is defined as active immunisation.

A critical requirement for both active and passive immunisation is the availability of suitable antigens with which to immunise the patient or animal respectively. These can comprise the natural toxins which can be purified from the media in which suitable toxigenic strains of *C. difficile* have been cultured. There are several disadvantages to this approach. Both Toxin A and Toxin B are present in culture medium in only small amounts and are difficult to purify without incurring significant losses. Thus, it is both costly and difficult to obtain the amounts necessary to meet world-wide needs. In addition, the natural toxins are unstable and toxic.

The above mentioned problems have resulted in there being few available *C. difficile* vaccine candidates. To-date, the only CDI vaccine in late-stage development is based on a mixture of native (i.e. naturally occurring) Toxins A and B, which have been extensively inactivated by chemical modification (Salnikova et al 2008, J Pharm Sci 97: 3735-3752).

One alternative to the use of natural toxins (and their toxoids) involves the design, development and use of recombinant fragments derived from Toxins A and B. Examples of existing antigens intended for use in treating/ preventing a *C. difficile* infection include peptides based on the C-terminal repeating units (RUs) of Toxin A or Toxin B – see, for example, WO 00/61762. A problem with such antigens, however, is that they are either poorly immunogenic (i.e. the antigens produce poor antibody titres), or, where higher antibody titres are produced, the antibodies demonstrate poor neutralising efficacy against *C. difficile* cytotoxic activity (i.e. insufficient neutralising antibodies are produced).

There is therefore a need in the art for new vaccines/ therapies/ therapeutics capable of specifically addressing *C. difficile* infection (CDI). This need is addressed by the present invention, which solves one or more of the above-mentioned problems.

In summary, the present invention provides antigens that are able to induce a potent toxin-neutralising response against C*. difficile* Toxin A and/ or B. The invention also provides methods for preparing recombinant antigens, and the use thereof as immunogens to enable the large-scale preparation of therapeutic antibodies. Said antibodies are able to induce a potent toxin-neutralising response against C*. difficile* Toxin A and/ or B and therefore have prophylactic and/ or therapeutic applications.

As mentioned above (see, for example, WO 00/61762), previous studies describe vaccine preparations based on the C-terminal, repeating units (RUs) of Toxin A and/ or Toxin B. Said RU fragments have a poor toxin-neutralising effect, and/ or are difficult to manufacture in large quantities.

In contrast, the present invention provides a C*. difficile* polypeptide antigen based on a Toxin A and/ or a Toxin B that does not contain or include one or more (e.g. all) of the repeating units (RUs) of Toxin A and/ or Toxin B. The polypeptide antigens of the invention consist of or comprise one or more domain from the central region of the toxins. Said antigens of the invention demonstrate good toxin-neutralising immune responses and/ or are readily manufactured in large quantities.

The present inventors have surprisingly identified that *C. difficile* antigen polypeptides which consist of or comprise one or more domain from the central region between the effector domain and the region of RUs (see Figure 1) provide a protective (toxin-neutralising) immune response that was greatly enhanced as compared to corresponding *C. difficile* toxin fragments comprising one or more of the RUs (see Tables 1 and 2). In one embodiment, the *C. difficile* antigen polypeptide comprises a Toxin B fragment.

Comparison of the data in Tables 1 and 2 confirms that the polypeptide antigens of the present invention elicit a considerably more potent toxin-neutralising immune

response than that of a corresponding polypeptide based that includes one or more of the C-terminal repeating units of *C. difficile* toxin (exemplified by the polypeptide designated TxB2). In more detail, after a 18-week immunisation period, the toxin-neutralising immune response provided by polypeptides of the present invention was more 60-fold higher than that provided by a corresponding RU-containing polypeptide. Thus, polypeptides of the present invention induce a potent toxin-neutralising immune response.

These findings are surprising for a number of reasons. Most importantly, a previous study in which animals were separately immunised with central domain fragments of *C. difficile* toxin (a fragment consisting of residues 510-1530, and a fragment consisting of residues 1530-1750) reported that these fragments failed to elicit the production of toxin-neutralising antibodies (Kink & Williams (1993) Infect. Immun. 66: 2018-2025). This study therefore suggests that domains within residues 510-1530 contribute no significant antibody-binding structural determinants. In addition, a further study has showed that antibodies raised against a whole *C. difficile*, while recognising a fragment consisting of the entire RU region alone, failed to recognise a fragment consisting of a central toxin region based on residues 901-1750 of the *C. difficile* same toxin (Genth *et al.*, (2000) Infect. Immun., 68: 1094-1101). This study therefore suggests that domains within residues 901-1750 contribute no significant antibody-binding structural determinants. Furthermore, while antibodies to the effector domain (residues 1-543) of *C. difficile* toxin have been shown to elicit a potent immune response (measured by simple enzyme immunoassay), said antibodies have no toxin-neutralising activity showing that antibody binding to the toxin does not correspond to toxin neutralisation (Roberts *et al.* (2012) Infect. Immun., 80: 875-882). Collectively, it is therefore extremely surprising that recombinant immunogens based on the central domains of the *C. difficile* toxins located between the effector domain and repeat regions are capable of inducing such a potent the toxin-neutralising immune response.

One aspect of the present invention provides a polypeptide containing, consisting of, or comprising an amino acid sequence that has at least 80% sequence identity with an amino acid sequence consisting of residues 1500-1700 (e.g. 1450-1750, or 1400-1800) of a *C. difficile* Toxin A sequence with the proviso that the polypeptide is not a

polypeptide comprising one or more of (e.g. all of) the RU units between amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or residues 1853-2366 of a *C. difficile* Toxin B. In one embodiment, said polypeptide lacks the sequence of amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or residues 1853-2366 of a *C. difficile* Toxin B.

Another aspect of the present invention provides a polypeptide containing, consisting of or comprising of an amino acid sequence that has at least 80% sequence identity with an amino acid sequence consisting of residues 542-1850 of a *C. difficile* Toxin A sequence with the proviso that the polypeptide does not comprise any of the RU units between residues 1851-2710 of *C. difficile* Toxin A or residues 1853-2366 of a *C. difficile* Toxin B. In one embodiment, said polypeptide lacks the sequence of amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or residues 1853-2366 of a *C. difficile* Toxin B.

Reference to a *C. difficile* Toxin A sequence means the amino acid sequence of a naturally-occurring *C. difficile* Toxin A (also referred to as a *C. difficile* Toxin A reference sequence). Examples of such sequences are readily understood by a skilled person, and simply for illustrative purposes some of the more common naturally-occurring Toxin A sequences are identified in the present specification (see, for example, SEQ ID NOs: 1 & 3) as well as throughout the literature.

Reference to 'at least 80% sequence identity' throughout this specification is considered synonymous with the phrase 'based on' and may embrace one or more of at least 85%, at least 90%, at least 93%, at least 95%, at least 97%, at least 99%, and 100% sequence identity. When assessing sequence identity, a reference sequence having a defined number of contiguous amino acid residues is aligned with an amino acid sequence (having the same number of contiguous amino acid residues) from the corresponding portion of a polypeptide of the present invention.

In one embodiment, the polypeptide amino acid sequence is based on (i.e. has at least 80% sequence identity with) amino acid residues 1500-1700or amino acid residues 1450-1750, or amino acid residues 1400-1800 of a *C. difficile* Toxin A. In another embodiment, the polypeptide amino acid sequence is based on amino acid

residues 544-1850 of a *C. difficile* Toxin A, such as amino acid residues 564-1850, amino acid residues 584-1850, amino acid residues 594-1850, amino acid residues 614-1850, amino acid residues 634-1850, amino acid residues 654-1850, amino acid residues 674-1850, amino acid residues 694-1850, amino acid residues 714-1850, amino acid residues 734-1850, amino acid residues 754-1850, amino acid residues 767-1850, amino acid residues 770-1850, amino acid residues 774-1850, amino acid residues 794-1850, amino acid residues 814-1850, amino acid residues 834-1850, amino acid residues 854-1850, amino acid residues 874-1850, amino acid residues 894-1850, amino acid residues 914-1850, amino acid residues 934-1850, amino acid residues 954-1850, amino acid residues 974-1850, amino acid residues 994-1850, amino acid residues 1014-1850, amino acid residues 1034-1850, amino acid residues 1054-1850, amino acid residues 1074-1850, amino acid residues 1094-1850, amino acid residues 1104-1850, amino acid residues 1124-1850, amino acid residues, amino acid residues 1131-1850, amino acid residues 1144-1850, amino acid residues 1164-1850, amino acid residues 1184-1850, amino acid residues 1204-1850, amino acid residues 1224-1850, amino acid residues 1244-1850, amino acid residues 1264-1850, amino acid residues 1284-1850, amino acid residues 1304-1850, amino acid residues 1324-1850, amino acid residues 1344-1850, amino acid residues 1450-1750 or amino acid residues 1550-1850; though with the proviso that the polypeptide does not include one or more of (e.g. all of) the RU units between residues 1851-2710 of *C. difficile* Toxin A or residues 1853-2366 of a *C. difficile* Toxin B. In one embodiment, said polypeptide lacks the sequence of amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or residues 1853-2366 of a *C. difficile* Toxin B. By way of example only, the above amino acid position numbering may refer to the *C. difficile* Toxin A sequences identified as SEQ ID NOs 1 and/ or 3.

In one embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 542-1850 of a Toxin A sequence (or a portion thereof). Examples are identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 5 and 6.

In another embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 542-1850 of a Toxin A sequence (or a portion thereof) that substantially lacks cysteine protease activity. An example is

identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 7.

In another embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 770-1850 of a Toxin A sequence (or a portion thereof). An example is identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 8.

In another embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 1130-1850 of a Toxin A sequence (or a portion thereof). An example is identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 9.

A related aspect of the present invention provides a polypeptide containing, consisting of, or comprising an amino acid sequence that has at least 80% sequence identity with an amino acid sequence consisting of residues 1500-1700 (e.g. 1450-1750, or 1400-1800) of a *C. difficile* Toxin B sequence with the proviso that the polypeptide does comprise one or more of (e.g. any of) the RU units between residues 1851-2710 of *C. difficile* Toxin A or residues 1853-2366 of a *C. difficile* Toxin B. In one embodiment, said polypeptide lacks the sequence of amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or residues 1853-2366 of a *C. difficile* Toxin B.

In another aspect of the present invention provides a polypeptide containing, consisting of, or comprising an amino acid sequence that has at least 80% sequence identity with an amino acid sequence consisting of residues 543-1852 of a *C. difficile* Toxin B sequence with the proviso that the polypeptide does not comprise one or more of (e.g. any of) the RU units between residues 1851-2710 of *C. difficile* Toxin A or residues 1853-2366 of a *C. difficile* Toxin B. In one embodiment, said polypeptide lacks the sequence of amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or residues 1853-2366 of a *C. difficile* Toxin B.

Reference to a *C. difficile* Toxin B sequence means the amino acid sequence of a naturally-occurring *C. difficile* Toxin B (also referred to as a *C. difficile* Toxin B

reference sequence). Examples of such sequences are readily understood by a skilled person, and simply for illustrative purposes some of the more common naturally-occurring Toxin B sequences are identified in the present specification (see, for example, SEQ ID NOs: 2 & 4) as well as throughout the literature.

In one embodiment, the polypeptide amino acid sequence is based on (i.e. has at least 80% sequence identity with) amino acid residues 1500-1700 or amino acid residues 1450-1750, or amino acid residues 1400-1800 of a *C. difficile* Toxin B. In another embodiment, the polypeptide amino acid sequence is based on amino acid residues 544-1852 of a *C. difficile* Toxin B, such as amino acid residues 564-1852, amino acid residues 584-1852, amino acid residues 594-1852, amino acid residues 614-1852, amino acid residues 634-1852, amino acid residues 654-1852, amino acid residues 674-1852, amino acid residues 694-1852, amino acid residues 714-1852, amino acid residues 734-1852, amino acid residues 754-1852, amino acid residues 767-1852, amino acid residues 770-1852, amino acid residues 774-1852, amino acid residues 794-1852, amino acid residues 814-1852, amino acid residues 834-1852, amino acid residues 854-1852, amino acid residues 874-1852, amino acid residues 894-1852, amino acid residues 914-1852, amino acid residues 934-1852, amino acid residues 954-1852, amino acid residues 974-1852, amino acid residues 994-1852, amino acid residues 1014-1852, amino acid residues 1034-1852, amino acid residues 1054-1852, amino acid residues 1074-1852, amino acid residues 1094-1852, amino acid residues 1104-1852, amino acid residues 1124-1852, amino acid residues 1131-1852, amino acid residues 1144-1852, amino acid residues 1164-1852, amino acid residues 1184-1852, amino acid residues 1204-1852, amino acid residues 1224-1852, amino acid residues 1244-1852, amino acid residues 1264-1852, amino acid residues 1284-1852, amino acid residues 1304-1852, amino acid residues 1324-1852, amino acid residues 1344-1852, amino acid residues 1450-1750 or amino acid residues 1550-1800, amino acid residues 1450-1750 or amino acid residues 1550-1850; though with the proviso that the polypeptide does not include one or more of (e.g. all of) the RU units between residues 1851-2710 of *C. difficile* Toxin A or residues 1853-2366 of a *C. difficile* Toxin B. In one embodiment, said polypeptide lacks the sequence of amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or residues 1853-2366 of a *C. difficile* Toxin B. By way of example only,

the above amino acid position numbering may refer to the *C. difficile* Toxin B sequences identified as SEQ ID NOs: 2 and/ or 4.

In one embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 543-1852 of a Toxin B sequence (or a portion thereof). Examples are identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 10 or 11.

In another embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 543-1852 of a Toxin B sequence (or a portion thereof) that substantially lacks cysteine protease activity. An example is identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 12.

In another embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 767-1852 of a Toxin B sequence (or a portion thereof). An example is identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 13.

In another embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 1145-1852 of a Toxin B sequence (or a portion thereof). An example is identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 14.

In another embodiment a polypeptide is provided, which comprises or consists of a sequence based on amino acid residues 1350-1852 of a Toxin B sequence (or a portion thereof). An example is identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 15.

The antigen polypeptides of the invention may substantially lack cysteine protease activity. In another (or the same) embodiment, antigens substantially lack glucosyl transferase activity. For example, amino acid sequence(s) providing said activity (activities) may be absent (e.g. deleted) from the antigens of the present invention. Alternatively, key amino acid residues essential for providing such activities may be

either modified or deleted. Examples of amino acid modifications to substantially reduce the cysteine protease activity of Toxin A are cysteine 700 to alanine, histidine 655 to alanine, aspartic acid 589 to asparagine or a combination of more than one of these mutations. Examples of amino acid modifications to substantially reduce the cysteine protease activity of Toxin B are cysteine 698 to alanine, histidine 653 to alanine, aspartic acid 587 to asparagine or a combination of more than one of these mutations. Examples of amino acid modifications to substantially reduce the glucosyl transferase activity of Toxin A are aspartic acid 285 to alanine, aspartic acid 287 to alanine or a combination of both mutations. Examples of amino acid modifications to substantially reduce the glucosyl transferase activity of Toxin B are aspartic acid 286 to alanine, aspartic acid 288 to alanine or a combination of both mutations These enzymatic activities are present in native Toxin A and/ or Toxin B, and are associated with N-terminal domains of said Toxins (see Figure 1, and/ or SEQ ID s 1, 2, 3 and 4).

The antigen polypeptides of the invention may substantially lack the glucosyl transferase (effector) domain (amino acid residues 1-542 Toxin A; amino acid residues 1-543 Toxin B) of a native *C. difficile* Toxin. In one embodiment, the antigen polypeptide lacks a functional glucosyl transferase (effector) domain (amino acid residues 1-542 Toxin A; amino acid residues 1-543 Toxin B) of a native *C. difficile* Toxin. In another (or the same) embodiment, the antigen substantially lacks the cysteine protease domain (amino acid residues 543-770 Toxin A; 544-767 Toxin B) of a native *C. difficile* Toxin. In one embodiment, the antigen polypeptide lacks a functional cysteine protease domain (amino acid residues 543-770 Toxin A; amino acid residues 544-767 Toxin B) of a native *C. difficile* Toxin. In one embodiment, the antigen polypeptide comprises or consists of a functional translocation domain (amino acid residues 770-1850 Toxin A; amino acid residues 767-1852 Toxin B) of a native *C. difficile* Toxin. In one embodiment, the antigen polypeptide comprises or consists of a complete or full-length translocation domain (amino acid residues 770-1850 Toxin A; amino acid residues 767-1852 Toxin B) of a native *C. difficile* Toxin.
In one embodiment, the antigen polypeptide does not include an amino acid sequence consisting of residues 850-1330, or 1500-1800, or 660-1255, or 1256-1852, or 543-1692 of a *C. difficile* Toxin B sequence or an amino acid sequence consisting of residues 660-1100 or 1100-1610 of a *C. difficile* Toxin B sequence.

Said amino acid residue numbering refers to any Toxin A or Toxin B toxinotype, for example any one or more of the reference Toxin A and/ or Toxin B toxinotype SEQ ID NOs recited in the present specification. Accordingly, said amino acid residue numbering may refer to any specific Toxin A and/ or Toxin B reference SEQ ID NO recited in the present specification including an amino acid sequence variant having at least 80%, at least 85%, at least 90%, at least 93%, at least 95%, at least 97%, or at least 99% thereto.

Antigen polypeptides of the invention include chimeras in which one portion of the antigen is derived from the central domain(s) of Toxin A and/ or Toxin B and a second portion of the antigen is based on a domain of a bacterial surface layer protein component (SLP) of *C. difficile* (or a fragment thereof). Inclusion of said domain has been identified by the present inventors to confer several advantages. Such SLP domains facilitate the soluble expression of antigen polypeptides of the invention. In addition, since these SLP domains are of *C. difficile* origin, it is unnecessary to cleave and remove them from constructs prior to immunisation. Indeed, the present inventors believe that antibodies to such domains recognise the intact *C. difficile* bacterium and afford additional therapeutic benefits by preventing or slowing the process of bacterial colonisation. An example of such a *C. difficile* SLP domain is based a polypeptide comprising or consisting of the polypeptide product from *C. difficile* gene CD 2767 (or a fragment thereof). By way of specific example, reference is made to a polypeptide fragment based on amino acid residues 27-401 of the polypeptide product of *C. difficile* gene CD 2767 (or a portion thereof) – see, for example, a polypeptide consisting of or comprising the amino acid sequence SEQ ID 16. Such a domain may be, for example, positioned at the N-terminus and/ or C-terminus of a polypeptide of the invention.

For example, in one embodiment of the invention, a polypeptide antigen is provided which consists of or comprises a CD 2767 polypeptide (e.g. based on residues 27-401 thereof) and a *C. difficile* toxin polypeptide (e.g. a central domain) as hereinbefore defined.

In another embodiment a polypeptide antigen is provided which consists of or comprises a CD 2767 polypeptide (e.g. based on residues 27-401 thereof) and a *C.*

difficile toxin polypeptide based on an amino acid sequence consisting of or comprising amino acid residues 770-1850 of a Toxin A sequence (or a portion thereof).

In another embodiment a polypeptide antigen is provided which consists of or comprises a CD 2767 polypeptide (e.g. based on residues 27-401 thereof) and a *C. difficile* toxin polypeptide based on an amino acid sequence consisting of or comprising amino acid residues 542-1850 of a Toxin A sequence (or a portion thereof). An example is identified as a polypeptide comprising or consisting of the amino acid sequence SEQ ID 17.

In a related embodiment of the invention and antigen is provided which consists of a chimera of CD 2767 polypeptide residues 27-401 (or a portion thereof) with amino acid residues 767-1852 of a Toxin B sequence (or a portion thereof). See SEQ ID 18.

In another embodiment of the invention a polypeptide antigen is provided which consists of or comprises a CD 2767 polypeptide (e.g. based on residues 27-401 thereof) and a *C. difficile* toxin polypeptide based on an amino acid sequence consisting of or comprising amino acid residues 543-1852 of a Toxin B sequence (or a portion thereof).

Antigen polypeptides of the invention may additionally (or alternatively) to an SPL comprise other fusion protein partners to facilitate soluble expression. Fusion protein partners may be attached at the N- or C-terminus of the antigen construct but are usually placed at the N-terminal end. Examples of fusion partners are: NusA, thioredoxin, maltose-binding protein, small ubiquitin-like molecules (Sumo-tag). To facilitate removal of the fusion protein partner during purification, a unique protease site may be inserted between the fusion protein partner and the fusion protein *per se*. Such protease sites may include those for thrombin, factor Xa, enterokinase, PreScission™, Sumo™. Alternatively, removal of the fusion protein partner may be achieved via inclusion of an intein sequence between the fusion protein partner and the fusion protein *per se*. Inteins are self-cleaving proteins and in response to a stimulus (e.g. lowered pH) are capable of self-splicing at the junction between the

intein and the antigen construct thus eliminating the need for the addition of specific proteases. Examples of inteins include domains derived from *Mycobacterium tuberculosis* (RecA), and *Pyrococcus horikoshii* (RadA) (Fong et al. (2010) Trends Biotechnol. 28:272-279).

To facilitate purification, antigens of the invention may include one or more purification tags to enable specific chromatography steps (e.g. metal ion chelating, affinity chromatography) to be included in the purification processes. Such purification tags may, for example, include: repeat histidine residues (e.g. 6-10 histidine residues), maltose binding protein, glutathione S-transferase; and streptavidin. These tags may be attached at the N- and/ or C-terminus of the polypeptide antigens of the invention. To facilitate removal of such tags during purification, protease sites and/ or inteins (examples above) may be inserted between the polypeptide and the purification tag(s). Examples of expression constructs for Toxin A and B derived antigens of the invention are shown in SEQ IDs 17, 18, 19 and 20.

Thus, a typical antigen construct of the invention (starting from the N-terminus) may comprise:
- a first purification tag
- an optional fusion protein partner (to facilitate expression) and/ or an optional SLP
- a first (preferably specific) protease sequence or intein sequence
- the Toxin A and/ or B antigen sequence
- an optional second (preferably specific) protease sequence or intein sequence
- an optional second purification tag

The first and second purification tags may be the same or different. Similarly, the first and second protease/ intein sequence may be the same or different. The first and second options are preferably different to enable selective and controllable cleavage/ purification.

In one embodiment, the antigen of the invention is a chimera and consists of a portion or domain of a *C. difficile* surface protein in conjunction with a toxin antigen sequence based on the central domain(s) Toxins A and/ or B.

Accordingly, in one embodiment, a polypeptide of the invention may comprise (starting from the N-terminus):
- a first purification tag
- a first (preferably specific) protease sequence or intein sequence
- an antigen sequence which is a chimera of Toxin A or B and an SLP
- an optional second (preferably specific) protease sequence or intein sequence
- an optional second purification tag

Spacers may be introduced to distance the purification tag from the polypeptide – this may help to increase binding efficiency to affinity purification column media. The spacer may be placed (immediately) after the purification tag or between the fusion protein partner component and the remainder of the polypeptide *per se*. Similarly, spacers may be employed to distance the fusion protein partner and/ or SLP from the *C. difficile* toxin component. Typical spacer sequences may consist of between 10-40 amino acid residues to give either a linear or alpha-helical structure.

Accordingly, in one embodiment, a polypeptide of the invention may comprise (starting from the N-terminus):
- a first purification tag
- an optional first spacer sequence
- a fusion protein partner (to facilitate expression) and/ or an SLP
- an optional second spacer sequence
- a (preferably specific) protease sequence or intein sequence
- the Toxin A and/ or B derived antigen sequence
- an optional second (preferably specific) protease sequence or intein sequence
- an optional third spacer sequence
- an optional second purification tag

Genes encoding the constructs of the invention may be generated by PCR from *C. difficile* genomic DNA and sequenced by standard methods to ensure integrity. Alternatively and preferably genes may be synthesised providing the optimal codon bias for the expression host (e.g. *E. coli*, *Bacillus megaterium*). Thus, the present

invention provides corresponding nucleic acid sequences that encode the aforementioned polypeptides of the present invention.

Accordingly, a second aspect of the present invention provides a method for expressing one or more of the aforementioned polypeptide antigens of the invention, said method comprising:

1) providing a nucleic acid sequence that encodes one or more of said polypeptide antigens in a host cell, wherein said nucleic acid sequence is operably linked to a promoter; and

2) expressing said nucleic acid sequence in the host cell

Antigen polypeptides of the invention may be formulated as vaccines for human or animal use in a number of ways. For example, formulation may include treatment with an agent to introduce intra-molecular cross-links. One example of such an agent is formaldehyde, which may be incubated, for example, with antigen polypeptides of the invention for between 1-24 hours. Alternatively, longer incubation times of, for example, up to 2, 4, 6, 8 or 10 days may be employed. Following treatment with such an agent, antigens of the invention may be combined with a suitable adjuvant, which may differ depending on whether the antigen is intended for human or animal use.

A human or animal vaccine formulation may contain polypeptides of the present invention. Thus, in one embodiment, a vaccine formulation procedure of the present invention comprises the following steps:

- providing a recombinant polypeptide of the invention in suitable buffer system
- optionally (preferably) treating said mixture with a toxoiding component such as formaldehyde
- optionally transferring the polypeptide to a new buffer system
- combining the polypeptide with one or more suitable adjuvants and optionally other excipients

Accordingly, a third aspect of the present invention provides one or more of the aforementioned polypeptides of the invention, for use in the generation of antibodies

that bind to *C. difficile* Toxin A and/ or Toxin B. In one embodiment, said antibodies bind to and neutralise *C. difficile* Toxin A and/ or Toxin B.

For immunisation of animals, the *C. difficile* recombinant antigen polypeptides of the invention may be used as immunogens separately or in combination, either concurrently or sequentially, in order to produce antibodies specific for individual *C. difficile* toxins or combinations. For example, two or more recombinant antigens may be mixed together and used as a single immunogen. Alternatively a *C. difficile* toxin antigen (e.g. Toxin A-derived) may be used separately as a first immunogen on a first animal group, and another *C. difficile* toxin antigen (e.g. Toxin B-derived) may be used separately on a second animal group. The antibodies produced by separate immunisation may be combined to yield an antibody composition directed against *C. difficile* toxins. Non-limiting examples of suitable adjuvants for animal/veterinary use include Freund's (complete and incomplete forms), alum (aluminium phosphate or aluminium hydroxide), saponin and its purified component Quil A.

A fourth (vaccine) aspect of the present invention provides one or more of the aforementioned polypeptide antigens of the invention, for use in the prevention, treatment or suppression of CDI (e.g. in a mammal such as man). Put another way, the present invention provides a method for the prevention, treatment or suppression of CDI (e.g. in a mammal such as man), said method comprising administration of a therapeutically effective amount of one or more of the aforementioned polypeptides of the invention to a subject (e.g. a mammal such as man).

By way of example, a Toxin A-based antigen (any A toxinotype) may be employed alone or in combination with a Toxin B-based antigen (any B toxinotype). Similarly, a Toxin B-based antigen (any B toxinotype) may be employed alone or in combination with a Toxin A-based antigen (any A toxinotype). Said antigens may be administered in a sequential or simultaneous manner. Vaccine applications of the present invention may further include the combined use (e.g. prior, sequential or subsequent administration) of one or more antigens such as a *C. difficile* antigen (e.g. a non-Toxin antigen; or a *C. difficile* bacterium such as one that has been inactivated or attenuated), and optionally one or more nosocomial infection antigens (e.g. an antigen, notably a surface antigen, from a bacterium that causes nosocomial

infection; and/ or a bacterium that causes a nosocomial infection such as one that has been inactivated or attenuated). Examples of bacteria that cause nosocomial infection include one or more of: *E. coli*, *Klebsiella pneumonae*, *Staphylococcus aureus* such as MRSA, Legionella, *Pseudomonas aeruginosa*, *Serratia marcescens*, Enterobacter spp, Citrobacter spp, *Stenotrophomonas maltophilia*, Acinetobacter spp such as *Acinetobacter baumannii*, *Burkholderia cepacia*, and Enterococcus such as vancomycin-resistant Enterococcus (VRE).

In one embodiment, said vaccine application may be employed prophylactically, for example to treat a patient before said patient enters a hospital (or similar treatment facility) to help prevent hospital-acquired infection. Alternatively, said vaccine application may be administered to vulnerable patients as a matter of routine.

A related vaccine aspect of the invention provides one or more antibodies (comprising or consisting whole IgG and/or Fab and/or $F(ab')_2$ fragments) that binds to the one or more aforementioned polypeptides of the invention, for use in the prevention, treatment or suppression of CDI (e.g. in a mammal such as man). Put another way, the present invention provides a method for the prevention, treatment or suppression of CDI (e.g. in a mammal such as man), said method comprising administration of a therapeutically effective amount of said antibody (or antibodies) to a subject (e.g. a mammal such as man).

By way of example, an anti-Toxin A-based antigen (any A toxinotype) antibody may be employed alone or in combination with an anti-Toxin B-based antigen (any B toxinotype) antibody. Similarly, an anti-Toxin B-based antigen (any B toxinotype) antibody may be employed alone or in combination with an anti-Toxin A-based antigen (any A toxinotype) antibody. Said antibodies may be administered in a sequential or simultaneous manner. Vaccine applications of the present invention may further include the combined use (e.g. prior, sequential or subsequent administration) of one or more antibodies that bind to antigens such as a *C. difficile* antigen (e.g. a non-Toxin antigen; or a *C. difficile* bacterium), and optionally one or more antibodies that bind to one or more nosocomial infection antigens (e.g. an antigen, notably a surface antigen, from a bacterium that causes nosocomial infection; and/ or a bacterium that causes a nosocomial infection). Examples of

bacteria that cause nosocomial infection include one or more of: *E. coli*, *Klebsiella pneumonae*, *Staphylococcus aureus* such as MRSA, Legionella, Pseudomonas aeruginosa, Serratia marcescens, Enterobacter spp, Citrobacter spp, Stenotrophomonas maltophilia, Acinetobacter spp such as *Acinetobacter baumannii*, *Burkholderia cepacia*, and Enterococcus such as vancomycin-resistant Enterococcus (VRE).

In one embodiment, said vaccine application may be employed prophylactically, for example once a patient has entered hospital (or similar treatment facility). Alternatively, said vaccine application may be administered to patients in combination with one or more antibiotics.

In one embodiment, said antibodies have been generated by immunisation of an animal (eg. a mammal such as man, or a non-human animal such as goat or sheep) with one or more of the aforementioned antigens of the present invention.

In one embodiment, the antibodies of the present invention do not (substantially) bind to the repeat regions of *C. difficile* Toxin A and/ or Toxin B.

For the preparation of vaccines for human (or non-human animal) use, the active immunogenic ingredients (whether these be antigenics of the present invention and/ or corresponding antibodies of the invention that bind thereto) may be mixed with carriers or excipients, which are pharmaceutically acceptable and compatible with the active ingredient. Suitable carriers and excipients include, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the vaccine may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, and/or adjuvants which enhance the effectiveness of the vaccine.

The vaccine may further comprise one or more adjuvants. One non-limiting example of an adjuvant with the scope of the invention is aluminium hydroxide. Other non-limiting examples of adjuvants include but are not limited to: N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-

alanine-2- (1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria, monophosphoryl lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2 % squalene/ Tween 80 emulsion.

Typically, the vaccines are prepared as injectables, either as liquid solutions or suspensions. Of course, solid forms suitable for solution in, or suspension in, liquid prior to injection may also be prepared. The preparation may also be emulsified, or the peptide encapsulated in liposomes or microcapsules.

Vaccine administration is generally by conventional routes *e.g.* intravenous, subcutaneous, intraperitoneal, or mucosal routes. The administration may be by parenteral injection, for example, a subcutaneous or intramuscular injection.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be prophylactically and/or therapeutically effective. The quantity to be administered, which is generally in the range of 5 micrograms to 250 micrograms of antigen per dose, depends on the subject to be treated, capacity of the subject's immune system to synthesize antibodies, and the degree of protection desired. Precise amounts of active ingredient required to be administered may depend on the judgment of the practitioner and may be particular to each subject.

The vaccine may be given in a single dose schedule, or optionally in a multiple dose schedule. A multiple dose schedule is one in which a primary course of vaccination may be with 1-6 separate doses, followed by other doses given at subsequent time intervals required to maintain and /or reinforce the immune response, for example, at 1-4 months for a second dose, and if needed, a subsequent dose(s) after several months. The dosage regimen will also, at least in part, be determined by the need of the individual and be dependent upon the judgment of the practitioner.

In one embodiment, a volume X (e.g. 1-6 ml) of buffer solution containing 10-500 µg of a polypeptide of the invention is mixed with an equivalent volume X (i.e. 1-6 ml) of adjuvant (e.g. Freund's complete adjuvant) to form an emulsion. Mixing with the adjuvant is carried out for several minutes to ensure a stable emulsion. A primary

immunisation is then performed (e.g. i.m. injection) with said emulsion (e.g. 1-10 ml). In parallel, a volume X (e.g. 1-6 ml) of buffer solution containing 10-500 µg of a polypeptide of the invention is mixed with an equivalent volume X (i.e. 1-6 ml) of adjuvant (e.g. Freund's incomplete adjuvant) to form an emulsion. Mixing with the adjuvant is carried out for several minutes to ensure a stable emulsion. Subsequent immunisations (e.g. 2, 3, 4, 5 or 6) are then performed (e.g. i.m. injection) with said emulsion (e.g. 1-10 ml) on a monthly basis. Antibody titre is typically tested by sampling at a time period of approximately 2 weeks after each of said monthly immunisations, and antibody harvesting is performed when optimal antibody titre has been achieved.

In addition, the vaccine containing the immunogenic antigen(s) may be administered in conjunction with other immunoregulatory agents, for example, immunoglobulins, antibiotics, interleukins (*e.g.*, IL-2, IL-12), and/or cytokines (*e.g.*, IFN gamma).

Additional formulations suitable for use with the present invention include microcapsules, suppositories and, in some cases, oral formulations or formulations suitable for distribution as aerosols. For suppositories, traditional binders and carriers may include, for example, polyalkylene glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of about 0.5 % to 10 %, including for instance, about 1 %-2 %.

Antigens of the invention may also have uses as ligands for use in affinity chromatography procedures. In such procedures, antigens of the invention may be covalently immobilised onto a matrix, such as Sepharose, e.g. using cyanogen bromide-activated Sepharose. Such affinity columns may then be used to purify antibody from antisera or partially purified solutions of immunoglobulins by passing them through the column and then eluting the bound IgG fraction (e.g. by low pH). Almost all of the antibody in the eluted fraction will be directed against the antigen of the invention, with non-specific antibodies and other proteins having been removed. These affinity purified IgG fractions have applications both as immunotherapeutics and as reagents in diagnostics. For immunotherapeutics, affinity purified antibodies enable a lower dose to be administered making adverse side effects less likely. For

diagnostics, affinity purified agents often give improved specificity and fewer false positive results.

DEFINITIONS

Clostridium difficile is a species of Gram-positive bacterium of the genus Clostridium.

Clostridium difficile infection (CDI) means a bacterial infection which affects humans and animals and which results in a range of symptoms from mild self-limiting diarrhoea to life-threatening conditions such as pseudomembranous colitis and cytotoxic megacolon. In this disease, *C. difficile* replaces some of the normal gut flora and starts to produce cytotoxins which attack and damage the gut epithelium. Primary risk factors for human CDI include: receiving broad-spectrum antibiotics, being over 65 years old and being hospitalised.

Clostridium difficile Toxin A is a family of protein cytotoxins/ enterotoxins of approximately 300 kDa in size. Toxin A has an enzyme activity within the N-terminal region which acts to disrupt the cytoskeleton of the mammalian cell causing cell death. There a number of naturally occurring variants of Toxin A within the strains of *Clostridium difficile* which are called 'toxinotypes'. The various toxinotypes of Toxin A have variations within their primary sequence of usually <10% overall. Examples of suitable Toxin A sequences include SEQ ID NOs: 1 and 3.

Clostridium difficile Toxin B is a family of protein cytotoxins of approximately 270 kDa in size which are similar to Toxin A but significantly more cytotoxic. Like Toxin A, Toxin B has an enzyme activity within the N-terminal region which acts to disrupt the cytoskeleton of the mammalian cell causing cell death. There are a number of naturally occurring variants of Toxin B within the strains of *C. difficile* which are called 'toxinotypes'. The various toxinotypes of Toxin B have variations within their primary sequence of up to 15% overall. Examples of suitable Toxin B sequences include SEQ ID NOs: 2 and 4.

C. difficile repeat units are regions within the C-terminus of Toxin A and B that contain repeating motifs which were first identified by von Eichel-Streiber and Sauerborn (1990; Gene 30: 107-113). In the case of Toxin A there are 31 short

repeats and 7 long repeats with each repeat consisting of a β-hairpin followed by a loop. Toxin B consists of a similar structure but with fewer repeats. The repeat units of Toxin A are contained within residues 1850-2710 and those for Toxin B within residues 1852-2366. The repeat regions play a role in receptor binding. The receptor binding regions (i.e. that define the toxin's structural binding pockets) appear to be clustered around the long repeat regions to form 'binding modules'.

Central domains of Toxin A and B are believed to play a role in translocation of the toxins into mammalian cells. The central domains of Toxin A are based on residues 542-1849 and those for Toxin B are based on residues 543-1851. Of the central domain regions of Toxins A and B, the first domain is a cysteine protease, which plays a role in the internalisation of the toxin's effector domain (which contains the glucosyl transferase activity).

Toxinotypes are often used to classify strains of *C. difficile*. Toxinotyping is based on a method which characterises the restriction patterns obtained with the toxin genes. Toxinotypes of Toxins A and B represent variants, by primary amino acid sequence, of these protein toxins. In one embodiment, the *C. difficile* toxin is selected from one of toxinotypes 0 to XV. Preferred Toxinotypes (plus example Ribotypes and Strains) are listed in the Table immediately below. The listed Toxinotypes are purely illustrative and are not intended to be limiting to the present invention.

Toxinotype	Example Ribotypes	Example Strains	Reference
0	001, 106	VPI10463	Rupnik et al. (1998) J. Clinical Microbiol.
1	003, 012, 102	EX623	
2	103	AC008	
3	027, 034, 075, 080	R20291, QCD-32g58	
4	023, 034, 075, 080	55767	
5	066, 078	SE881	
6	045, 063, 066	51377	
7	063	57267	
8	017, 047	1470	

9	019	51680	36: 2240-2247
10	036	8864	
11	033	IS58, R11402	Rupnik et al. (2001) Microbiology 147: 439-447
12	056	IS25	
13	070	R9367	
14	111	R10870	
15	122	R9385	

An "antibody" is used in the broadest sense and specifically covers polyclonal antibodies and antibody fragments so long as they exhibit the desired biological activity. For example, an antibody is a protein including at least one or two, heavy (H) chain variable regions (abbreviated herein as VHC), and at least one or two light (L) chain variable regions (abbreviated herein as VLC). The VHC and VLC regions can be further subdivided into regions of hypervariability, termed "complementarity determining regions" ("CDR"), interspersed with regions that are more conserved, termed "framework regions" (FR). The extent of the framework region and CDRs has been precisely defined (see, Kabat, E.A., et al. Sequences of Proteins of Immunological Interest, Fifth Edition, U.S. Department of Health and Human Services, NIH Publication No. 91-3242, 1991, and Chothia, C. et al, J. Mol. Biol. 196:901-917, 1987, which are incorporated herein by reference). Preferably, each VHC and VLC is composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FRI, CDRI, FR2, CDR2, FR3, CDR3, FR4.

The VHC or VLC chain of the antibody can further include all or part of a heavy or light chain constant region. In one embodiment, the antibody is a tetramer of two heavy immunoglobulin chains and two light immunoglobulin chains, wherein the heavy and light immunoglobulin chains are inter-connected by, e.g., disulfide bonds. The heavy chain constant region includes three domains, CHI, CH2 and CH3. The light chain constant region is comprised of one domain, CL. The variable region of the heavy and light chains contains a binding domain that interacts with an antigen.

The constant regions of the antibodies typically mediate the binding of the antibody to host tissues or factors, including various cells of the immune system (e.g., effector cells) and the first component (CIq) of the classical complement system. The term "antibody" includes intact immunoglobulins of types IgA, IgG, IgE, IgD, IgM (as well as subtypes thereof), wherein the light chains of the immunoglobulin may be of types kappa or lambda.

The term antibody, as used herein, also refers to a portion of an antibody that binds to a toxin of *C. difficile* (e.g. Toxin A or B), e.g., a molecule in which one or more immunoglobulin chains is not full length, but which binds to a toxin. Examples of binding portions encompassed within the term antibody include (i) a Fab fragment, a monovalent fragment consisting of the VLC, VHC, CL and CHI domains; (ii) a $F(ab')_2$ fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fc fragment consisting of the VHC and CHI domains; (iv) a Fv fragment consisting of the VLC and VHC domains of a single arm of an antibody, (v) a dAb fragment (Ward et al, Nature 341:544-546, 1989), which consists of a VHC domain; and (vi) an isolated complementarity determining region (CDR) having sufficient framework to bind, e.g. an antigen binding portion of a variable region. An antigen binding portion of a light chain variable region and an antigen binding portion of a heavy chain variable region, e.g., the two domains of the Fv fragment, VLC and VHC, can be joined, using recombinant methods, by a synthetic linker that enables them to be made as a single protein chain in which the VLC and VHC regions pair to form monovalent molecules (known as single chain Fv (scFv); see e.g., Bird et al. (1988) Science lAl-ATi-Alβ; and Huston et al. (1988) Proc. Natl. Acad. ScL USA 85:5879-5883). Such single-chain antibodies (as well as camelids) are also encompassed within the term antibody. These are obtained using conventional techniques known to those with skill in the art, and the portions are screened for utility in the same manner as are intact antibodies.

The term “fragment” means a peptide typically having at least 70, 90, 110, 130, 150, 170, 190, 210, 230, 250, 270, 290, 310, 330, 350 contiguous amino acid residues of (based on) the corresponding reference sequence.

The term "variant" means a peptide or peptide fragment having at least eighty, preferably at least eighty five, more preferably at least ninety percent amino acid sequence homology with a reference polypeptide sequence (e.g., a *C. difficile* toxin polypeptide amino acid sequence, and / or a fusion protein partner amino acid sequence, and / or an SLP amino acid reference sequence). For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences may be compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequent coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percentage sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

Any of a variety of sequence alignment methods can be used to determine percent identity, including, without limitation, global methods, local methods and hybrid methods, such as, e.g., segment approach methods. Protocols to determine percent identity are routine procedures within the scope of one skilled in the art. Global methods align sequences from the beginning to the end of the molecule and determine the best alignment by adding up scores of individual residue pairs and by imposing gap penalties. Non-limiting methods include, e.g., CLUSTAL W, see, e.g., Julie D. Thompson et al., CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position- Specific Gap Penalties and Weight Matrix Choice, 22(22) Nucleic Acids Research 4673-4680 (1994); and iterative refinement, see, e.g., Osamu Gotoh, Significant Improvement in Accuracy of Multiple Protein. Sequence Alignments by Iterative Refinement as Assessed by Reference to Structural Alignments, 264(4) J. Mol. Biol. 823-838 (1996). Local methods align sequences by identifying one or more conserved motifs shared by all of the input sequences. Non-limiting methods include, e.g., Match-box, see, e.g., Eric Depiereux and Ernest Feytmans, Match-Box: A Fundamentally New Algorithm for the Simultaneous Alignment of Several Protein Sequences, 8(5) CABIOS 501 -509 (1992); Gibbs sampling, see, e.g., C. E. Lawrence et al., Detecting Subtle Sequence Signals: A Gibbs Sampling Strategy for Multiple Alignment, 262(5131) Science 208-214 (1993); Align-M, see, e.g., Ivo Van Walle et

al., Align-M - A New Algorithm for Multiple Alignment of Highly Divergent Sequences, 20(9) Bioinformatics:1428-1435 (2004).

Thus, percent sequence identity is determined by conventional methods. See, for example, Altschul et al., Bull. Math. Bio. 48: 603-16, 1986 and Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA 89:10915-19, 1992. Briefly, two amino acid sequences are aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "blosum 62" scoring matrix of Henikoff and Henikoff (ibid.) as shown below (amino acids are indicated by the standard one-letter codes).

Alignment scores for determining sequence identity

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4																			
R	-1	5																		
N	-2	0	6																	
D	-2	-2	1	6																
C	0	-3	-3	-3	9															
Q	-1	1	0	0	-3	5														
E	-1	0	0	2	-4	2	5													
G	0	-2	0	-1	-3	-2	-2	6												
H	-2	0	1	-1	-3	0	0	-2	8											
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7					
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4				
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5			
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11		
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4

The percent identity is then calculated as:

$$\frac{\text{Total number of identical matches}}{\text{[length of the longer sequence plus the number of gaps introduced into the longer sequence in order to align the two sequences]}} \times 100$$

Substantially homologous polypeptides are characterized as having one or more amino acid substitutions, deletions or additions. These changes are preferably of a minor nature, that is conservative amino acid substitutions (see below) and other substitutions that do not significantly affect the folding or activity of the polypeptide; small deletions, typically of one to about 30 amino acids; and small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue, a small linker peptide of up to about 20-25 residues, or an affinity tag.

Conservative amino acid substitutions

Basic:	arginine
	lysine
	histidine
Acidic:	glutamic acid
	aspartic acid
Polar:	glutamine
	asparagine
Hydrophobic:	leucine
	isoleucine
	valine
Aromatic:	phenylalanine
	tryptophan
	tyrosine
Small:	glycine
	alanine

serine
threonine
methionine

In addition to the 20 standard amino acids, non-standard amino acids (such as 4-hydroxyproline, 6-*N*-methyl lysine, 2-aminoisobutyric acid, isovaline and α -methyl serine) may be substituted for amino acid residues of the polypeptides of the present invention. A limited number of non-conservative amino acids, amino acids that are not encoded by the genetic code, and unnatural amino acids may be substituted for clostridial polypeptide amino acid residues. The polypeptides of the present invention can also comprise non-naturally occurring amino acid residues.

Non-naturally occurring amino acids include, without limitation, trans-3-methylproline, 2,4-methano-proline, cis-4-hydroxyproline, trans-4-hydroxy-proline, N-methylglycine, allo-threonine, methyl-threonine, hydroxy-ethylcysteine, hydroxyethylhomo-cysteine, nitro-glutamine, homoglutamine, pipecolic acid, tert-leucine, norvaline, 2-azaphenylalanine, 3-azaphenyl-alanine, 4-azaphenyl-alanine, and 4-fluorophenylalanine. Several methods are known in the art for incorporating non-naturally occurring amino acid residues into proteins. For example, an *in vitro* system can be employed wherein nonsense mutations are suppressed using chemically aminoacylated suppressor tRNAs. Methods for synthesizing amino acids and aminoacylating tRNA are known in the art. Transcription and translation of plasmids containing nonsense mutations is carried out in a cell free system comprising an *E. coli* S30 extract and commercially available enzymes and other reagents. Proteins are purified by chromatography. See, for example, Robertson et al., J. Am. Chem. Soc. 113:2722, 1991; Ellman et al., Methods Enzymol. 202:301, 1991; Chung et al., Science 259:806-9, 1993; and Chung et al., Proc. Natl. Acad. Sci. USA 90:10145-9, 1993). In a second method, translation is carried out in Xenopus oocytes by microinjection of mutated mRNA and chemically aminoacylated suppressor tRNAs (Turcatti et al., J. Biol. Chem. 271:19991-8, 1996). Within a third method, *E. coli* cells are cultured in the absence of a natural amino acid that is to be replaced (e.g., phenylalanine) and in the presence of the desired non-naturally occurring amino acid(s) (e.g., 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, or 4-fluorophenylalanine). The non-naturally occurring amino acid

is incorporated into the polypeptdie in place of its natural counterpart. See, Koide et al., Biochem. 33:7470-6, 1994. Naturally occurring amino acid residues can be converted to non-naturally occurring species by *in vitro* chemical modification. Chemical modification can be combined with site-directed mutagenesis to further expand the range of substitutions (Wynn and Richards, Protein Sci. 2:395-403, 1993).

A limited number of non-conservative amino acids, amino acids that are not encoded by the genetic code, non-naturally occurring amino acids, and unnatural amino acids may be substituted for amino acid residues of polypeptides of the present invention.

Essential amino acids in the polypeptides of the present invention can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, Science 244: 1081-5, 1989). Sites of biological interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction or photoaffinity labeling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos et al., Science 255:306-12, 1992; Smith et al., J. Mol. Biol. 224:899-904, 1992; Wlodaver et al., FEBS Lett. 309:59-64, 1992. The identities of essential amino acids can also be inferred from analysis of homologies with related components (e.g. the translocation or protease components) of the polypeptides of the present invention.

Multiple amino acid substitutions can be made and tested using known methods of mutagenesis and screening, such as those disclosed by Reidhaar-Olson and Sauer (Science 241:53-7, 1988) or Bowie and Sauer (Proc. Natl. Acad. Sci. USA 86:2152-6, 1989). Briefly, these authors disclose methods for simultaneously randomizing two or more positions in a polypeptide, selecting for functional polypeptide, and then sequencing the mutagenised polypeptides to determine the spectrum of allowable substitutions at each position. Other methods that can be used include phage display (e.g., Lowman et al., Biochem. 30:10832-7, 1991; Ladner et al., U.S. Patent No. 5,223,409; Huse, WIPO Publication WO 92/06204) and region-directed mutagenesis (Derbyshire et al., Gene 46:145, 1986; Ner et al., DNA 7:127, 1988).

Toxin-neutralising means the capacity of a substance to prevent the cytotoxic action of either Toxin A or B on a mammalian cell. In assays for toxin-neutralising activity, a fixed amount of toxin is mixed with various concentrations of a neutralising substance (e.g. an antibody) and the mixture applied to and incubated with a mammalian cell line (e.g. Vero cells) for a fixed time. The neutralising titre may be measured by several methods:

(a) The dilution of the substance (serum, antibody, purified IgG) that completely protects the cells from the cytotoxic effects of either Toxin A or B. These cytotoxic effects are evident by cell rounding and endpoint may be quantified by microscopy methods.

(b) The dilution of the substance (serum, antibody, purified IgG) that protects 50% of the cells (ED_{50} titre) from the cytotoxic effects of either Toxin A or B. The ED_{50} titre may be assessed by the use of dyes (e.g. crystal violet) which give a measure of cell integrity. Fitting titration data to either 4- or 5-parameter logistic curves provides an accurate estimation of the ED_{50} titre. ED_{50} estimations, which are generally more accurate than microscopy based methods, provide a quantitative estimation of the toxin-neutralising capacity of serum and purified antibodies.

Toxin-neutralising titres are measured in the presence of a fixed concentration of Toxin A or Toxin B which is set at a multiple of that required to induce cell death over a 24 incubation period. Typically, final concentrations of Toxin A may be set at 50 ng/ml and Toxin B between 0.5 - 2 ng/ml. The difference in the concentrations between Toxin A and Toxin B reflect the significantly higher specific cytotoxic activity of Toxin B. Thus, an ED_{50} titre for an antibody of serum solution of 1000 units/ml indicates that at a 1000-fold dilution, the antibody solution is capable of neutralising 50% of the Toxin A or B cytotoxic activity. With respect to titres in serum, a toxin-neutralising titre $\geq$1000 unit/ml may be regarded as potent neutralising activity.

For highly purified IgG solutions, neutralising activity may also be expressed as the concentration of IgG (µg/ml) required to neutralise 50% of the Toxin A or B cytotoxic activity. In this case, a titre value $\leq$ 10 µg/ml IgG may be regarded as potent neutralising activity.

C. difficile surface proteins (SLPs) means those proteins that are associated with the bacterial cell wall. Examples of 29 *C. difficile* surface (cell wall proteins) are given in Table 1 of Fagan *et al.* (2011) J. Medical Microbiol. 60:1225-1228, which is hereby incorporated in its entirety by reference thereto.

FIGURES

Figure 1 illustrates to structures of *C. difficile* Toxins A and B showing amino acid residues at the various domain boundaries.

Figure 2 illustrates the purification of Toxin B recombinant fragment residues 767-1852 as a fusion protein with thioredoxin. The left-hand Figure shows a 4-12 % SDS-PAGE analysis of TxB3. Columns C2-C8 show lanes with partially purified fragment.

Figure 3 illustrates the purification of Toxin A recombinant fragment residues 542-1850 as a fusion protein with thioredoxin.

Figure 4. Expression and purification of CD2767 (residues 27-401). Purification of CD2767 (residues 27-401) by immobilised metal ion affinity chromatography. Key: column load, L; flow through fraction, FT; eluted purified CD2767 polypeptide, E1. The intense band of the CD2767 polypeptide is illustrative of its high solubility.

Figure 5. Expression and purification of a fusion protein consisting of CD2767 (residues 27-401) with Toxin A (residues 543-1851). The fusion protein was expressed as a soluble polypeptide (top band in lanes C1 and C2) at approximately 2% of the total protein. C1 and C2 represent duplicate of the soluble fraction of the total expressed protein.

Figure 6. Antibody-mediated neutralisation of Toxin A and Toxin B as measured by the Vero cell ED_{50} assay. The capacity of various dilutions of antiserum to Toxin A (residues 543-1851) (●) to prevent the cytotoxic effects of purified Toxin A (50 ng/ml) was assessed using crystal violet staining to measure cell viability. Reduced absorbance indicates a lack of cell integrity. Neutralisation by Toxin B (residues

767-1852) antiserum (▼) was assessed using Toxin B at 2 ng/ml. Each antiserum was a pool obtained from the immunisation of 3 sheep.

EXAMPLES

Example 1 - Expression and purification of Toxin B fragment recombinant fragment residues 767-1852 as a fusion protein with thioredoxin (His_6TrxTxBcentral)

Expression

L-broth (100 ml) supplemented with 200 µg/ml ampicillin and 0.4 % glucose was inoculated with a scrape from a glycerol freeze (BL21 (DE3) *E. coli* harbouring plasmid pET59His_6TrxTxBcentral) and maintained overnight at 30°C and 180 rpm. The overnight culture was used as a 2.5 % inoculum for Terrific Broth (4 x 1L in 2.5L unbaffled flasks) supplemented with 200 µg/ml ampicillin and 0.2 % glucose. Cultures were maintained at 37°C with orbital shaking (180 rpm) to an absorbance at 600 nm of 0.6. The temperature of the cultures was reduced to 16°C and protein expression induced with the addition of 1 mM IPTG. The culture was maintained overnight at 16°C with orbital shaking as before. Cell paste (60 g) was harvested by centrifugation (Sorvall RC3BP centrifuge, H6000A rotor, 4000 g for 20 minutes).

Immobilised nickel affinity purification of His_6TrxTxBcentral

Cells (60 g) were resuspended in buffer (pH 8, 20 mM Tris, 50 mM NaCl) and subjected to lysis using sonication. The lysate was cleared by centrifugation (Sorvall RC5C centrifuge, SS-34 rotor, 20,000 g, 20 minutes) and made up to 1 M ammonium sulphate with a saturated solution. The solution was stored on ice for 1 hour and the resultant precipitate collected by centrifugation (Heraeus Multifuge X3R centrifuge, 4000g, 4°C). The precipitate was resuspended in 250 ml of low imidazole buffer (pH 7.5, 50 mM Hepes, 0.5 M NaCl, 20 mM imidazole) and applied to a 30 ml nickel column (Ø 26 mm) at a flow rate of 3 ml/min. The column was washed with low imidazole buffer and bound protein eluted using a gradient from 0-100 % high imidazole buffer (pH7.5, 50 mM Hepes, 0.5 M NaCl, 0.5 M imidazole). Fractions were analysed on 4-12% NuPAGE Bis-Tris polyacrylamide gels with coomassie staining. SDS PAGE of partially purified fractions are shown in Figure 2.

Example 2 - Expression and purification of Toxin A fragment recombinant fragment residues 542-1850 (TxACPD)

L-broth (100 ml) supplemented with 100 µg/ml ampicillin and 0.2 % glucose was inoculated with a glycerol freeze (BL21 (DE3) *E. coli* harbouring plasmid pET59TxACPDcentral). The culture was maintained (37°C, 180 rpm) to an absorbance at 600 nm of 0.6. The 100 ml culture was used as a 2 % inoculum for Terrific Broth (4 x 0.75 L) supplemented with 200 µg/ml ampicillin and 0.2 % glucose. Cultures were maintained at 37°C with orbital shaking (180 rpm) to an absorbance at 600 nm of 0.6. The temperature of the cultures was reduced to 16°C and protein expression induced 1 hour later with the addition of 1 mM IPTG. The culture was maintained overnight at 16°C with orbital shaking as before. Cell paste (37 g) was harvested by centrifugation (Sorvall RC3BP centrifuge, H6000A rotor, 4000 g, 20 minutes) and stored at -80°C.

Cells (37g) were resuspended with 260 ml buffer (20 mM Tris, 50 mM NaCl, pH 8) and subjected to lysis using sonication. The lysate was cleared by centrifugation (30,000 g, 20 minutes) and the clarified lysate stirred gently over ice whilst 130 ml of saturated ammonium sulphate solution (pH 8) was added to bring the mixture to 33 % saturation. The mixture was left on ice for 15-20 minutes to allow a cloudy white precipitate to form. The precipitate was harvested by centrifugation (30,000 g, 15 minutes) and resuspended in 70 ml 'low imidazole' buffer (pH7.5, 50 mM Hepes, 0.5 M NaCl, 20 mM imidazole, 5 % glycerol).The solution was applied to a 30 ml nickel column (Ø 26 mm) at a flow rate of 2 ml/min. The column was washed at 2 ml/min until the UV absorbance of the flow through returned to near baseline levels. Bound material was eluted from the column with a 160 ml gradient (2 ml/min) to 100% 'high imidazole' buffer (pH7.5, 50 mM Hepes, 0.5 M NaCl, 0.5 M imidazole, 15 % glycerol). Fractions were analysed on 4-12% NuPAGE Bis-Tris polyacrylamide gels and those containing the highest amount of the expression construct pooled.

To cleavage of the His_6Thioredoxin tag, protein solution containing the constructs (18 ml, 1.5 mg/ml) from the first immobilised nickel column was thawed at room temperature and restriction grade thrombin (30 U) was added and the mixture which was incubated at room temperature ~ 20°C for 16-18 hours.

The protein mix was loaded at 1 ml/min onto the column consisting of 30 ml (Ø 26 mm) nickel charged chelating Sepharose column equilibrated with buffer A (50 mM HEPES, 0.5 M NaCl, 20 % glycerol, pH 7.5). The column was washed with 30 ml buffer A. Bound protein was then eluted using 4% (50 ml), 8% (50 ml) and finally a gradient (120 ml) to 100 % buffer B (50 mM HEPES, 0.5 M NaCl, 0.5 M imidazole, 20 % glycerol, pH 7.5). The fraction containing the purified fractions were analysed by SDS PAGE and fraction conyaing the purest construct pooled and dialysed into storage buffer (pH 7.5, 50 mM Hepes, 0.5 M NaCl, 20% glycerol).

SDS PAGE analysis and Western blot analysis of purified TxACPD constructs

Purified protein solution was mixed 1:1 with 4x LDS-PAGE loading buffer supplemented with 5 mM DTT. The sample was heated at 95°C for five minutes and loaded in duplicate (5 and 10 µl) onto a 4-12 % NuPAGE Bis-Tris polyacrylamide gel. The gel was run in MES running buffer at 200 V for 45 minutes. One part of the gel was subjected to coomassie staining and the other blotted onto a nitrocellulose membrane at 40 V for 1 hour in transfer buffer. The membrane was blocked with 5 % skimmed milk in tris buffered saline supplemented with 0.1 % Tween 20 (TBST) for 40 minutes. The membrane was incubated for 40 minutes with sheep anti-Toxin A antibody diluted 1:25,000 in 1% skimmed milk TBST. The stock antibody concentration was 50 mg/ml. The membrane was washed for 4 x 15 minutes in TBST. A donkey anti-sheep antibody alkaline phosphatase conjugate was applied to the membrane at a dilution of 1:10,000 in 1 % skimmed milk TBST. The solution was left on the membrane for 40 minutes with gentle agitation as before. The membrane was washed as before with TBST and the blot developed using NBT/BCIP one step reagent. SDS PAGE and Western blot are shown in Figure 3.

Example 3 – Expression and purification of residues 27-401 of *C. difficile* protein CD2767

A synthetic gene which encodes residues 27-401 of *C. difficile* protein CD2767 was synthesised commercially with its codon bias optimised for expression in a host such as *E. coli.* The gene was inserted into a pET28a expression vector and transformed into a BL21 *E. coli* expression strain using standard molecular biology procedures. The *E. coli* expression strain was grown and protein expression induced with IPTG

essentially as described in Example 1 except kanamycin was used in place of ampicillin. Cell pellets were either used directly or frozen at -20°C.

For protein extraction, cells were thawed and resuspended in 50 mM TrisHCl pH 8.0 buffer containing 0.5M NaCl and 20mM imidazole, sonicated (6 x 30sec with 30sec cooling after each) and then centrifuged at 47000 x g for 20 min. The His6-tagged residue 27-401 CD2767 polypeptide was then purified from the supernatant fluid by using immobilised metal ion (Ni) affinity chromatography. Application of the sample to the column and washing was in the above Tris/NaCl/imidazole buffer. The purified construct was then eluted with a gradient to 0.5M imidazole in the same buffer.

The CD2767 (residues 27-401) polypeptide was obtained as >90% pure protein by the single purification step and appeared as an intense band of approx 47kDa on SDS PAGE (Figure 4). The protein fragment could be concentrated to >120 mg/ml as measured by absorbance at 280 nm or to >167mg/ml as measured by the Bradford protein assay (bovine serum albumin as a standard). Both these assays illustrate the extremely high solubility of the CD2767 (residues 27-401) polypeptide and its potential usefulness as a solubility enhancing component within recombinant fusion proteins.

Example 4 - Expression and purification of either Toxin A or B recombinant fragments as a fusion protein with residues 27-401 of *C. difficile* protein CD2767

A synthetic gene which encodes a fusion protein in which the N-terminus consists of residues 27-401 of *C. difficile* protein CD2767 and the C-terminus consists of Toxin B fragment recombinant fragment residues 767-1852 may be synthesised commercially with its codon bias optimised for expression in a host such as *E. coli*. A synthetic gene which encodes a fusion protein in which the N-terminus consists of residues 27-401 of *C. difficile* protein CD2767 and the C-terminus consists of Toxin A fragment recombinant fragment residues 770-1850 may be similarly obtained. These and other fusion proteins mat be incorporated with expression vectors with various purification tags (6 histidine) incorporated to facilitate purification. An example of such an expression construct is shown in SEQ ID 19 which consists of CD2767 (residues 27-401) and Toxin A (residues 542-1850).

Expression and purification of the above constructs may be undertaken by similar methods as those outlined in Examples 1 and 2 and expression of a construct consisting of CD2767 (residues 27-401) as an N-terminal fusion to Toxin A (residues 543-1851) is shown in Figure 5. Here, addition of the CD2767 domain to the Toxin A fragment renders it soluble and expressible as 2% of the total soluble protein. After expression in *E. coli*, purification of the construct is effected by immobilised metal ion affinity chromatography and other chromatography methods such as ion exchange chromatography.

Example 5 – Formulation of antigens of the invention for immunisation of animals

Purified *C. difficile* antigens at a concentration of between 0.5 – 2 mg/ml (nominally 1 mg/ml) were dialysed against a suitable buffer (e.g. 10mM Hepes buffer pH 7.4 containing 150mM NaCl) and then formaldehyde added to a final concentration of 0.2% and incubated for up to 7 days at 35°C. After incubation, the formaldehyde may optionally be removed by dialysis against a suitable buffer, e.g. phosphate buffered saline.

For sheep, 2 ml of buffer solution containing between 10 and 500 µg of the above *C. difficile* antigen is mixed with 2.6 ml of Freund's adjuvant to form an emulsion. Mixing with the adjuvant is carried out for several minutes to ensure a stable emulsion. The complete form of the adjuvant is used for the primary immunisation and incomplete Freund's adjuvant for all subsequent boosts.

Example 6 – Generation of antibodies to antigens of the invention

A number of conventional factors are taken into consideration during the preparation of antiserum in order to achieve the optimal humoral antibody response. These include: breed of animal; choice of adjuvant; number and location of immunisation sites; quantity of immunogen; and number of and interval between doses. With conventional optimisation of these parameters is routine to obtain specific antibody levels in excess of 6 g/litre of serum.

For sheep, an emulsion of the antigen with Freund's adjuvant was prepared as described as in Example 5. The complete form of the adjuvant is used for the primary immunisation and incomplete Freund's adjuvant for all subsequent boosts. About 4.2 ml of the antigen/adjuvant mixture was used to immunise each sheep by i.m. injection and spread across 6 sites including the neck and all the upper limbs. This was repeated every 28 days. Blood samples were taken 14 days after each immunisation.

For comparison of the toxin-neutralising immune response to the different antigens, 3 sheep were used per antigen. They were immunised as above using an identical protocol and the same protein dose per immunisation.

Example 7 – Assessment of the neutralising efficacy of antisera to toxins using the *in vitro* cell assay

The toxin neutralizing activity of the antisera against *C. difficile* Toxins was measured by cytotoxicity assays using Vero cells. A fixed amount of either purified *C. difficile* Toxin A or Toxin B was mixed with various dilutions of the antibodies, incubated for 30min at 37°C and then applied to Vero cells growing on 96-well tissue culture plates. Both Toxin A and B possess cytotoxic activity which results in a characteristic rounding of the Vero cells over a period of 24 - 72 h. In the presence of neutralising antibodies this activity is inhibited and the neutralising strength of an antibody preparation may be assessed by the dilution required to neutralise the effect of a designated quantity of either Toxin A or B.

Data demonstrating the neutralising activity of ovine antibody to various recombinant *C. difficile* Toxin B antigens are shown in Table 1 and 2. In these experiments, various dilutions of ovine antibody were mixed with Toxin B at a final concentration of 0.5 ng/ml and incubated for 30min at 37°C and then applied to Vero cells as above and incubated at 37° and monitored over a period of 24 -72 h. The antibody dilutions which completely protect the cells against the cytotoxic effects of the Toxin B were calculated.

Table 1 shows the neutralising titres of an antigen of the invention and Table 2 shows the titres obtained using an antigen which consist of just the repeat regions.

Collectively, the data in Tables 1 and 2 show the superior capacity of antigens of the invention to elicit a toxin-neutralising immune response compared to fragments containing just the repeat domains.

Antibody toxin neutralisation titres were also estimated by colorimetric assays based on cell staining with crystal violet (Rothman (1986) J. Clin. Pathol. 39: 672-676). Vero cells were grown to confluence in 96-well cell culture plates. These assays were performed as described above using final concentrations of Toxin A and Toxin B in antibody mixtures of 50 ng/ml and 2 ng/ml, respectively. After overnight incubation, cells were washed gently with 200µl of Dulbecco's-PBS (Sigma) which was carefully removed before the cells were fixed with 70 µl ice cold ethanol for 2 min. The ethanol was then removed and 70 µl crystal violet (1% w/v in ethanol; Pro-Lab) added to the fixed cells and incubated for 30 minutes at 22°C. Plates were then washed carefully by immersion in deionized water to remove excess dye, dried at 37°C and then 200 µl of 50% (v/v) ethanol added. Plates were then incubated at 37°C in shaker incubator (300 rpm) for 2 h before being read at 492 nm. ED_{50} values were derived from the resulting toxin neutralisation curves using 4-or 5-pl nonlinear regression models (Figure 4). Thus, the ED_{50} titre is the dilution of the serum or antibody required to achieve the 50% toxin-neutralising endpoint in the assay. If antibody solutions of known IgG concentrations are used, the titres may also be expressed as the concentration of IgG required to achieve the 50% toxin-neutralising endpoint.

Table 3 shows the toxin-neutralising ED_{50} titres obtained using the crystal violet method for the serum generated using the central domains of both Toxin A and Toxin B. For both fragments, toxin-neutralising ED_{50} titres in excess of 1000 unit/ml were obtained for their respective sera (see also Figure 6).

Toxin-neutralizing ED_{50} titre values obtained for a sheep anti Toxin B (residues 767-1852) IgG solution are shown in Table 4. The neutralising titres against various toxinotypes of Toxin B were obtained for this fragment antiserum in order to assess its cross neutralising efficacy. Each purified toxinotype of Toxin B was normalised for toxicity in the assay and held at a fixed concentration of 16x the minimum toxin concentration which causes cell death in a 24 hr incubation period. Neutralising

potencies are expressed in µg/ml IgG required for 50% neutralisation of the above Toxin B concentration. Less than 4-fold difference in neutralising titres was observed which is indicative of good cross-neutralising efficacy.

Example 8 - Assessment of the *in vivo* efficacy of antiserum generated using recombinant antigens of the invention for treating CDI

To demonstrate the efficacy of the antisera generated, using recombinant antigens, to treat CDI in vivo, Syrian hamsters are passively immunised with antibodies which have neutralising activity against one or more of the toxins of *C. difficile*. For assessing the efficacy of a treatment formulation, hamsters will be given antibody either intravenously or by the intraperitoneal route at various times from 6 hours post-challenge to 240 hours post challenge with *C. difficile.*

Prior to passively immunisation hamsters are administered a broad spectrum antibiotic (e.g. clindamycin) and 12-72 h later challenged with *C. difficile* spores by mouth. Animals are then monitored for up to 15 days for symptoms of *C. difficile*-associated disease. Control, non-immunised animals develop signs of the disease (e.g. diarrhoea, swollen abdomen, lethargy, ruffled fur) while those treated with ovine antibody appear normal or show statistically significant reduced incidence of disease.

Example 9 - Vaccination by peptide/ peptide fragments of the invention

A vaccine, represented by a peptide/ peptide fragment of the invention is prepared by current Good Manufacturing Practice. Using such practices, peptides/ peptide fragments of the invention may be bound to an adjuvant of aluminium hydroxide which is commercially available (*e.g.* Alhydrogel). The vaccine would normally contain a combination of antigens of the invention derived from Toxin A and Toxin B but could also contain either Toxin A or B antigens. The vaccine may also contain Toxin A and B antigens in combination with other antigens of bacterial or viral origin.

Purified *C. difficile* Toxin A and/or Toxin B antigen of the invention may be treated with formaldehyde at a final concentration of 0.2% and incubated for up to 24 hours at 35^{o}C (as described in Example 5).

In addition to the antigens of the invention, a typical vaccine composition comprises:
A) A buffer (*e.g.*, Hepes buffer between 5 and 20 mM and pH between 7.0 and 7.5;
B) A salt component to make the vaccine physiologically isotonic (*e.g.* between 100 and 150 mM NaCl);
C) An adjuvant (*e.g.*, aluminium hydroxide at a final aluminium concentration of between 100 and 700µg per vaccine dose); and
D) A preservative (*e.g.*, Thiomersal at 0.01% or formaldehyde at 0.01%).

Such vaccine compositions are administered to humans by a variety of different immunisation regimens, such as:
1. A single dose (*e.g.*, 20 µg adsorbed fragment of the invention) in 0.5 ml administered sub-cutaneously.
2. Two doses (*e.g.*, of 10 µg adsorbed fragment of the invention) in 0.5 mls administered at 0 and 4 weeks.
3. Three doses (*e.g.*, of 10 µg adsorbed fragment of the invention) in 0.5 mls administered at 0, 2 and 12 weeks.

These vaccination regimens confer levels of protection against exposure to the homologous serotypes of *C. difficile* toxins

Example 10 - Clinical use of antibodies produced using antigens of the invention

Three examples serve to illustrate the therapeutic value of the systemic ovine antibody products, produced using antigens of the invention, in patients with differing degrees of seventy in their CDI.

Mild CDI

A 67 year old male is admitted to a coronary care unit following a severe myocardial infarction. Whilst making an uneventful recovery he develops a mild diarrhoea without any other signs or symptoms. Because there have been recent episodes of CDI in the hospital, a faecal sample is sent immediately for testing and found to contain both Toxin A and Toxin B. After isolation to a single room with its own toilet he receives 250mg of the ovine F(ab')2 intravenously followed by a second injection

two days later. His diarrhoea stops quickly and he makes a full recovery without the need of either metranidazole or vancomycin.

Severe CDI with risk of relapse

A female aged 81 falls in her home and sustained a fractured left hip. She is immediately admitted to hospital and the hip is pinned successfully. Her frail condition prevented early discharge and, a few days later, she develops a productive cough for which she was given a wide spectrum antibiotic. After a further eight days she develops profuse diarrhoea with abdominal pain and tenderness and CDI is diagnosed by the appropriate faecal tests. At the time there is also evidence of systemic manifestations of the infection including a markedly raised white blood cell count, and of significant fluid loss with dehydration. The patient is started immediately on oral vancomycin and, at the same time, receives the first of five daily injections of 250mg of the ovine F(ab')2 - based product intravenously. There is a rapid resolution of the signs and symptoms and of the laboratory manifestations of CDI. However, in order to avoid the risk of relapse of her CDI following stopping vancomycin, she continues to be treated for a further two weeks on an oral form of the antibody therapy. She experiences no relapse.

Severe CDI with complications

An 87 year old female develops bronchopneumonia while resident in long-stay care facilities. The local general practitioner starts her on a course of antibiotic therapy with immediate benefit. However, eight days after stopping the antibiotic she experiences severe diarrhoea. Her condition starts to deteriorate necessitating admission to hospital where Toxin A is detected in her faeces by an ELISA test. By this time she is extremely ill with evidence of circulatory failure and her diarrhoea has stopped. The latter is found to be due a combination of paralytic ileus and toxic megacolon and an emergency total colectomy is considered essential. Since such surgery is associated with mortality in excess of 60% she receives intravenous replacement therapy together with the contents of two ampoules (500mg) of antibody product. By the time she is taken to the operating theatre four hours later, her general condition had improved significantly and she survives surgery.

Table 1 - Neutralisation titres obtained by immunisation of sheep with a recombinant Toxin B-derived antigen (TxBcentral; residues 767-1852) of the invention

Antigen	No of Doses	Immunisation period (weeks)	Neutralisation titre against Toxin B (0.5ng/ml)
Recombinant Toxin B (residues 767-1852) at 100µg/dose	2	6	480
	3	10	5,120
	4	14	5,120
	5	18	5,120

Table 2 - Neutralisation titres obtained by immunisation of sheep with a recombinant Toxin B-derived antigen (TxB2, residues 1756-2366) representing the repeat regions

Antigen	No of Doses	Immunisation period (weeks)	Neutralisation titre against Toxin B (0.5ng/ml)
Recombinant Toxin B (residues 1756-2366) at 100µg/dose	2	6	<10
	3	10	10
	4	14	10
	5	18	80

Table 3. Neutralisation ED_{50} titres obtained by immunisation of sheep with a recombinant antigensToxin B (residues 767-1852) and Toxin A (residues 543-1851)

Antigen	ELISA Titre	Neutralising titre ED_{50}		
		1	2	Mean
Toxin B (767-1852)	1×10^5	7946 ±1027	8225 ±418	8086
Toxin A (542-1850)	2×10^5	2472 ±199	2096 ±222	2284

For each antigen, 5 doses of 100µg were given monthly to each of 3 sheep and the serum analysed at 18 weeks. ELISA titres, derived from 14 week samples, represent serum dilutions (pool from 3 animals) which gave a signal of 0.5 A_{450} above background and are the mean of duplicate determinations. For the crystal violet ED_{50} assay, Toxin B was used at a fixed concentration of 2ng/ml and Toxin A at 50ng/ml.

Table 4. Neutralisation ED_{50} titres against various Toxin B toxinotypes using serum antiserum generated to recombinant Toxin B (residues 767-1852)

Immunising Antigen	Assay Toxinotype	Neutralising Potency ED_{50} (µg/ml IgG)		
		1	2	Mean
Toxin B (residues 767-1852)	Toxin B (0)	2.5 ± 0.22	2.1 ± 0.20	2.3
	Toxin B (3)	7.9 ± 0.64	7.5 ± 1.20	7.7
	Toxin B (5)	7.7 ± 0.73	8.1 ± 0.71	7.9
	Toxin B (10)	7.2 ± 0.67	9.0 ± 0.75	8.1

Antibodies to Toxin B (residues 767-1852) (toxinotype 0 sequence) were assessed for their capacity to neutralise other Toxin B toxinotypes. Purified Toxin B toxinotypes (0, 3, 5 and 10) were each titrated in the cell assay and used at a fixed concentration of 16x the minimum toxin concentration which causes cell death in a 24 hr incubation period. Neutralising potencies are expressed in µg/ml IgG required for 50% neutralisation of the above Toxin B concentration.

SED ID NOs

SEQ ID NO: 1 - *Clostridium difficile* Toxin A (Toxinotype 0)

MSLISKEELIKLAYSIRPRENEYKTILTNLDEYNKLTTNNNENKYLQLKKLNESIDVFMNKYKTSSRNRALSNL
KKDILKEVILIKNSNTSPVEKNLHFVWIGGEVSDIALEYIKQWADINAEYNIKLWYDSEAFLVNTLKKAIVESST
TEALQLLEEEIQNPQFDNMKFYKKRMEFIYDRQKRFINYYKSQINKPTVPTIDDIIKSHLVSEYNRDETVLESY
RTNSLRKINSNHGIDIRANSLFTEQELLNIYSQELLNRGNLAAASDIVRLLALKNFGGVYLDVDMLPGIHSDLF
KTISRPSSIGLDRWEMIKLEAIMKYKKYINNYTSENFDKLDQQLKDNFKLIIESKSEKSEIFSKLENLNVSDLEI
KIAFALGSVINQALISKQGSYLTNLVIEQVKNRYQFLNQHLNPAIESDNNFTDTTKIFHDSLFNSATAENSMFL
TKIAPYLQVGFMPEARSTISLSGPGAYASAYYDFINLQENTIEKTLKASDLIEFKFPENNLSQLTEQEINSLWS
FDQASAKYQFEKYVRDYTGGSLSEDNGVDFNKNTALDKNYLLNNKIPSNNVEEAGSKNYVHYIIQLQGDDI
SYEATCNLFSKNPKNSIIIQRNMNESAKSYFLSDDGESILELNKYRIPERLKNKEKVKVTFIGHGKDEFNTSE
FARLSVDSLSNEISSFLDTIKLDISPKNVEVNLLGCNMFSYDFNVEETYPGKLLLSIMDKITSTLPDVNKNSITI
GANQYEVRINSEGRKELLAHSGKWINKEEAIMSDLSSKEYIFFDSIDNKLKAKSKNIPGLASISEDIKTLLLDA
SVSPDTKFILNNLKLNIESSIGDYIYYEKLEPVKNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFED
ISKNNSTYSVRFINKSNGESVYVETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNTLNAAFF
IQSLIDYSSNKDVLNDLSTSVKVQLYAQLFSTGLNTIYDSIQLVNLISNAVNDTINVLPTITEGIPIVSTILDGINL
GAAIKELLDEHDPLLKKELEAKVGVLAINMSLSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKAT
SVVNYFNHLSESKKYGPLKTEDDKILVPIDDLVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPS
ISSHIPSLSIYSAIGIETENLDFSKKIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLYPGKFYWR
FYAFFDYAITTLKPVYEDTNIKIKLDKDTRNFIMPTITTNEIRNKLSYSFDGAGGTYSLLLSSYPISTNINLSKDD
LWIFNIDNEVREISIENGTIKKGKLIKDVLSKIDINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVA
KSYSLLLSGDKNYLISNLSNTIEKINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFYND
STLEFNSKDFIAEDINVFMKDDINTITGKYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFVKNS
DGHHNTSNFMNLFLDNISFWKLFGFENINFVIDKYFTLVGKTNLGYVEFICDNNKNIDIYFGEWKTSSSKSTIF
SGNGRNVVVEPIYNPDTGEDISTSLDFSYEPLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNT
FHKKVNINLDSSSFEYKWSTEGSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSN
FKSFNSENELDRDHLGFKIIDNKTYYYDEDSKLVKGLININNSLFYFDPIEFNLVTGWQTINGKKYYFDINTGA
ALTSYKIINGKHFYFNNDGVMQLGVFKGPDGFEYFAPANTQNNNIEGQAIVYQSKFLTLNGKKYYFDNNSK
AVTGWRIINNEKYYFNPNNAIAAVGLQVIDNNKYYFNPDTAIISKGWQTVNGSRYYFDTDTAIAFNGYKTIDG
KHFYFDSDCVVKIGVFSTSNGFEYFAPANTYNNNIEGQAIVYQSKFLTLNGKKYYFDNNSKAVTGLQTIDSK
KYYFNTNTAEAATGWQTIDGKKYYFNTNTAEAATGWQTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGI
MQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNEFLTLNGKKYYFGSDSKAVTGWRIINNKKYYFNPNNAI
AAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANNESKMVTGVFKGPNGFEYFAPANTHNNNIEGQAIV
YQNKFLTLNGKKYYFDNDSKAVTGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTI
DGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLN
GKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGI
MQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYYFEPNTA
MGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSK
AVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGIYG

SEQ ID NO: 2 - *C. difficile* Toxin B (Toxinotype 0)

MSLVNRKQLEKMANVRFRTQEDEYVAILDALEEYHNMSENTVVEKYLKLKDINSLTDIYIDTYKKSGRNKAL
KKFKEYLVTEVLELKNNNLTPVEKNLHFVWIGGQINDTAINYINQWKDVNSDYNVNVFYDSNAFLINTLKKTV
VESAINDTLESFRENLNDPRFDYNKFFRKRMEIIYDKQKNFINYYKAQREENPELIIDDIVKTYLSNEYSKEID
ELNTYIEESLNKITQNSGNDVRNFEEFKNGESFNLYEQELVERWNLAAASDILRISALKEIGGMYLDVDMLP
GIQPDLFESIEKPSSVTVDFWEMTKLEAIMKYKEYIPEYTSEHFDMLDEEVQSSFESVLASKSDKSEIFSSLG
DMEASPLEVKIAFNSKGIINQGLISVKDSYCSNLIVKQIENRYKILNNSLNPAISEDNDFNTTTNTFIDSIMAEA
NADNGRFMMELGKYLRVGFFPDVKTTINLSGPEAYAAAYQDLLMFKEGSMNIHLIEADLRNFEISKTNISQS
TEQEMASLWSFDDARAKAQFEEYKRNYFEGSLGEDDNLDFSQNIVVDKEYLLEKISSLARSSERGYIHYIV
QLQGDKISYEAACNLFAKTPYDSVLFQKNIEDSEIAYYYNPGDGEIQEIDKYKIPSIISDRPKIKLTFIGHGKDE
FNTDIFAGFDVDSLSTEIEAAIDLAKEDISPKSIEINLLGCNMFSYSINVEETYPGKLLLKVKDKISELMPSISQD
SIIVSANQYEVRINSEGRRELLDHSGEWINKEESIIKDISSKEYISFNPKENKITVKSKNLPELSTLLQEIRNNS
NSSDIELEEKVMLTECEINVISNIDTQIVEERIEEAKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSHFIS
FEDISETDEGFSIRFINKETGESIFVETEKTIFSEYANHITEEISKIKGTIFDTVNGKLVKKVNLDTTHEVNTLNA
AFFIQSLIEYNSSKESLSNLSVAMKVQVYAQLFSTGLNTITDAAKVVELVSTALDETIDLLPTLSEGLPIIATIID
GVSLGAAIKELSETSDPLLRQEIEAKIGIMAVNLTTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLR
DKATKVVDYFKHVSLVETEGVFTLLDDKIMMPQDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDH
FFSAPSITYREPHLSIYDVLEVQKEELDLSKDLMVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNY
EGEFYWRYFAFIADALITTLKPRYEDTNIRINLDSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYNMG
INIELSESDVWIIDVDNVVRDVTIESDKIKKGDLIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILE
GINAIIEVDLLSKSYKLLISGELKILMLNSNHIQQKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVS

ELPDVVLISKVYMDDSKPSFGYYSNNLKDVKVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDES
GVAEILKFMNRKGNTNTSDSLMSFLESMNIKSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFI
KFNTLETNYTLYVGNRQNMIVEPNYDLDDSGDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYET
NNTYPEVIVLDANYINEKINVNINDLSIRYVWSNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNF
SDKQDVPVSEIILSFTPSYYEDGLIGYDLGLVSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNLITGFVT
VGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLII
DENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNKHYFD
DSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDD
SFTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQN
IDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDM
ENESDKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGED
GVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTA
QLVISE

SEQ ID NO:3 - *C. difficile* Toxin A (Toxinotype 3)

MSLISKEELIKLAYSIRPRENEYKTILTNLDEYNKLTTNNNENKYLQLKKLNESIDVFMNKYKNSSRNRALSNL
KKDILKEVILIKNSNTSPVEKNLHFVWIGGEVSDIALEYIKQWADINAEYNIKLWYDSEAFLVNTLKKAIVESST
TEALQLLEEEIQNPQFDNMKFYKKRMEFIYDRQKRFINYYKSQINKPTVPTIDDIIKSHLVSEYNRDETLLESY
RTNSLRKINSNHGIDIRANSLFTEQELLNIYSQELLNRGNLAAASDIVRLLALKNFGGVYLDVDMLPGIHSDLF
KTIPRPSSIGLDRWEMIKLEAIMKYKKYINNYTSENFDKLDQQLKDNFKLIIESKSEKSEIFSKLENLNVSDLEI
KIAFALGSVINQALISKQGSYLTNLVIEQVKNRYQFLNQHLNPAIESDNNFTDTTKIFHDSLFNSATAENSMFL
TKIAPYLQVGFMPEARSTISLSGPGAYASAYYDFINLQENTIEKTLKASDLIEFKFPENNLSQLTEQEINSLWS
FDQASAKYQFEKYVRDYTGGSLSEDNGVDFNKNTALDKNYLLNNKIPSNNVEEAGSKNYVHYIIQLQGDDI
SYEATCNLFSKNPKNSIIIQRNMNESAKSYFLSDDGESILELNKYRIPERLKNKEKVKVTFIGHGKDEFNTSE
FARLSVDSLSNEISSFLDTIKLDISPKNVEVNLLGCNMFSYDFNVEETYPGKLLLSIMDKITSTLPDVNKDSITI
GANQYEVRINSEGRKELLAHSGKWINKEEAIMSDLSSKEYIFFDSIDNKLKAKSKNIPGLASISEDIKTLLLDA
SVSPDTKFILNNLKLNIESSIGDYIYYEKLEPVKNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFED
ISKNNSTYSVRFINKSNGESVYVETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNTLNAAFF
IQSLIDYSSNKDVLNDLSTSVKVQLYAQLFSTGLNTIYDSIQLVNLISNAVNDTINVLPTITEGIPIVSTILDGINL
GAAIKELLDEHDPLLKKELEAKVGVLAINMSLSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKAT
SVVNYFNHLSESKEYGPLKTEDDKILVPIDDLVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPY
ISSHIPSLSVYSAIGIKTENLDFSKKIMMLPNAPSRVFWWETGAVPGLRSLENNGTKLLDSIRDLYPGKFYW
RFYAFFDYAITTLKPVYEDTNTKIKLDKDTRNFIMPTITTDEIRNKLSYSFDGAGGTYSLLLSSYPISMNINLSK
DDLWIFNIDNEVREISIENGTIKKGNLIEDVLSKIDINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEIN
LVAKSYSLLLSGDKNYLISNLSNTIEKINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFY
NGSTLEFNSKDFIAEDINVFMKDDINTITGKYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFV
KNSDGHHNTSNFMNLFLNNISFWKLFGFENINFVIDKYFTLVGKTNLGYVEFICDNNKNIDIYFGEWKTSSSK
STIFSGNGRNVVVEPIYNPDTGEDISTSLDFSYEPLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLN
PNTFHKKVNINLDSSSFEYKWSTEGSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYI
MSNFKSFNSENELDRDHLGFKIIDNKTYYYDEDSKLVKGLININNSLFYFDPIESNLVTGWQTINGKKYYFDI
NTGAASTSYKIINGKHFYFNNNGVMQLGVFKGPDGFEYFAPANTQNNNIEGQAIVYQSKFLTLNGKKYYFD
NDSKAVTGWRIINNEKYYFNPNNAIAAVGLQVIDNNKYYFNPDTAIISKGWQTVNGSRYYFDTDTAIAFNGY
KTIDGKHFYFDSDCVVKIGVFSGSNGFEYFAPANTYNNNIEGQAIVYQSKFLTLNGKKYYFDNNSKAVTGW
QTIDSKKYYFNTNTAEAATGWQTIDGKKYYFNTNTAEAATGWQTIDGKKYYFNTNTSIASTGYTIINGKYFYF
NTDGIMQIGVFKVPNGFEYFAPANTHNNNIEGQAILYQNKFLTLNGKKYYFGSDSKAITGWQTIDGKKYYFN
PNNAIAATHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANNESKMVTGVFKGPNGFEYFAPANTHNNNI
EGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDSKKYYFNLNTAVAVTGWQTIDGEKYYFNLNTAEAAT
GWQTIDGKRYYFNTNTYIASTGYTIINGKHFYFNTDGIMQIGVFKGPDGFEYFAPANTHNNNIEGQAILYQNK
FLTLNGKKYYFGSDSKAVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTYIASTGYTIISGKHFYF
NTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNDSKAATGWATIDGNRYYF
EPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGPNGFEYFAPANTDANNIDGQAIRYQNRFLHLLGKIYYF
GNNSKAVTGWQTINSKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGIYG

SEQ ID NO: 4 - *C. difficile* Toxin B (Toxinotype 3)

MSLVNRKQLEKMANVRFRVQEDEYVAILDALEEYHNMSENTVVEKYLKLKDINSLTDIYIDTYKKSGRNKAL
KKFKEYLVTEVLELKNNNLTPVEKNLHFVWIGGQINDTAINYINQWKDVNSDYNVNVFYDSNAFLINTLKKTI
VESATNDTLESFRENLNDPRFDYNKFYRKRMEIIYDKQKNFINYYKTQREENPDLIIDDIVKIYLSNEYSKDID
ELNSYIEESLNKVTENSGNDVRNFEEFKGGESFKLYEQELVERWNLAAASDILRISALKEVGGVYLDVDMLP
GIQPDLFESIEKPSSVTVDFWEMVKLEAIMKYKEYIPGYTSEHFDMLDEEVQSSFESVLASKSDKSEIFSSLG
DMEASPLEVKIAFNSKGIINQGLISVKDSYCSNLIVKQIENRYKILNNSLNPAISEDNDFNTTTNAFIDSIMAEA
NADNGRFMMELGKYLRVGFFPDVKTTINLSGPEAYAAAYQDLLMFKEGSMNIHLIEADLRNFEISKTNISQS
TEQEMASLWSFDDARAKAQFEEYKKNYFEGSLGEDDNLDFSQNTVVDKEYLLEKISSLARSSERGYIHYIV
QLQGDKISYEAACNLFAKTPYDSVLFQKNIEDSEIAYYYNPGDGEIQEIDKYKIPSIISDRPKIKLTFIGHGKDE

FNTDIFAGLDVDSLSTEIETAIDLAKEDISPKSIEINLLGCNMFSYSVNVEETYPGKLLLRVKDKVSELMPSISQ
DSIIVSANQYEVRINSEGRRELLDHSGEWINKEESIIKDISSKEYISFNPKENKIIVKSKNLPELSTLLQEIRNNS
NSSDIELEEKVMLAECEINVISNIDTQVVEGRIEEAKSLTSDSINYIKNEFKLIESISDALYDLKQQNELEESHFI
SFEDILETDEGFSIRFIDKETGESIFVETEKAIFSEYANHITEEISKIKGTIFDTVNGKLVKKVNLDATHEVNTLN
AAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQLFSTGLNTITDAAKVVELVSTALDETIDLLPTLSEGLPVIATII
DGVSLGAAIKELSETSDPLLRQEIEAKIGIMAVNLTAATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELIL
RDKATKVVDYFSHISLAESEGAFTSLDDKIMMPQDDLVISEIDFNNNSITLGKCEIWRMEGGSGHTVTDDID
HFFSAPSITYREPHLSIYDVLEVQKEELDLSKDLMVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDN
YEGEFYWRYFAFIADALITTLKPRYEDTNIRINLDSNTRSFIVPVITTEYIREKLSYSFYGSGGTYALSLSQYN
MNINIELNENDTWVIDVDNVVRDVTIESDKIKKGDLIENILSKLSIEDNKIILDNHEINFSGTLNGGNGFVSLTFS
ILEGINAVIEVDLLSKSYKVLISGELKTLMANSNSVQQKIDYIGLNSELQKNIPYSFMDDKGKENGFINCSTKE
GLFVSELSDVVLISKVYMDNSKPLFGYCSNDLKDVKVITKDDVIILTGYYLKDDIKISLSFTIQDENTIKLNGVY
LDENGVAEILKFMNKKGSTNTSDSLMSFLESMNIKSIFINSLQSNTKLILDTNFIISGTTSIGQFEFICDKDNNIQ
PYFIKFNTLETKYTLYVGNRQNMIVEPNYDLDDSGDISSTVINFSQKYLYGIDSCVNKVIISPNIYTDEINITPIY
EANNTYPEVIVLDTNYISEKININDLSIRYVWSNDGSDFILMSTDEENKVSQVKIRFTNVFKGNTISDKISFN
FSDKQDVSINKVISTFTPSYYVEGLLNYDLGLISLYNEKFYINNFGMMVSGLVYINDSLYYFKPPIKNLITGFTT
IGDDKYYFNPDNGGAASVGETIIDGKNYYFSQNGVLQTGVFSTEDGFKYFAPADTLDENLEGEAIDFTGKLT
IDENVYYFGDNYRAAIEWQTLDDEVYYFSTDTGRAFKGLNQIGDDKFYFNSDGIMQKGFVNINDKTFYFDD
SGVMKSGYTEIDGKYFYFAENGEMQIGVFNTADGFKYFAHHDEDLGNEEGEALSYSGILNFNNKIYYFDDS
FTAVVGWKDLEDGSKYYFDEDTAEAYIGISIINDGKYYFNDSGIMQIGFVTINNEVFYFSDSGIVESGMQNID
DNYFYIDENGLVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDME
NESDKYYFDPETKKAYKGINVIDDIKYYFDENGIMRTGLITFEDNHYYFNEDGIMQYGYLNIEDKTFYFSEDG
IMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQL
VISE

SEQ ID NO: 5 - *C. difficile* Toxin A 542-1850 (toxinotype 0)

LSEDNGVDFNKNTALDKNYLLNNKIPSNNVEEAGSKNYVHYIIQLQGDDISYEATCNLFSKNPKNSIIIQRNM
NESAKSYFLSDDGESILELNKYRIPERLKNKEKVKVTFIGHGKDEFNTSEFARLSVDSLSNEISSFLDTIKLDI
SPKNVEVNLLGCNMFSYDFNVEETYPGKLLLSIMDKITSTLPDVNKNSITIGANQYEVRINSEGRKELLAHSG
KWINKEEAIMSDLSSKEYIFFDSIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIESSIGDYI
YYEKLEPVKNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSVRFINKSNGESVYV
ETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNTLNAAFFIQSLIDYSSNKDVLNDLSTSVKV
QLYAQLFSTGLNTIYDSIQLVNLISNAVNDTINVLPTITEGIPIVSTILDGINLGAAIKELLDEHDPLLKKELEAKV
GVLAINMSLSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKATSVVNYFNHLSESKKYGPLKTED
DKILVPIDDLVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIPSLSIYSAIGIETENLDFSK
KIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLYPGKFYWRFYAFFDYAITTLKPVYEDTNIKIK
LDKDTRNFIMPTITTNEIRNKLSYSFDGAGGTYSLLLSSYPISTNINLSKDDLWIFNIDNEVREISIENGTIKKGK
LIKDVLSKIDINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVAKSYSLLLSGDKNYLISNLSNTIE
KINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFYNDSTLEFNSKDFIAEDINVFMKDDI
NTITGKYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFVKNSDGHHNTSNFMNLFLDNISFWK
LFGFENINFVIDKYFTLVGKTNLGYVEFICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTGEDI
STSLDFSYEPLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINLDSSSFEYKWSTE
GSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSFNSENELDRDHLGFKIIDN
KTYYYDEDSKLVKGLININNSLFYFDPIEFNL

SEQ ID NO: 6 - *C. difficile* Toxin A 542-1850 (toxinotype 3)

LSEDNGVDFNKNTALDKNYLLNNKIPSNNVEEAGSKNYVHYIIQLQGDDISYEATCNLFSKNPKNSIIIQRNM
NESAKSYFLSDDGESILELNKYRIPERLKNKEKVKVTFIGHGKDEFNTSEFARLSVDSLSNEISSFLDTIKLDI
SPKNVEVNLLGCNMFSYDFNVEETYPGKLLLSIMDKITSTLPDVNKDSITIGANQYEVRINSEGRKELLAHSG
KWINKEEAIMSDLSSKEYIFFDSIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIESSIGDYI
YYEKLEPVKNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSVRFINKSNGESVYV
ETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNTLNAAFFIQSLIDYSSNKDVLNDLSTSVKV
QLYAQLFSTGLNTIYDSIQLVNLISNAVNDTINVLPTITEGIPIVSTILDGINLGAAIKELLDEHDPLLKKELEAKV
GVLAINMSLSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKATSVVNYFNHLSESKEYGPLKTED
DKILVPIDDLVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPYISSHIPSLSVYSAIGIKTENLDFS
KKIMMLPNAPSRVFWWETGAVPGLRSLENNGTKLLDSIRDLYPGKFYWRFYAFFDYAITTLKPVYEDTNTKI
KLDKDTRNFIMPTITTDEIRNKLSYSFDGAGGTYSLLLSSYPISMNINLSKDDLWIFNIDNEVREISIENGTIKK
GNLIEDVLSKIDINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVAKSYSLLLSGDKNYLISNLSN
TIEKINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFYNGSTLEFNSKDFIAEDINVFMK
DDINTITGKYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFVKNSDGHHNTSNFMNLFLNNISF
WKLFGFENINFVIDKYFTLVGKTNLGYVEFICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTG
EDISTSLDFSYEPLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINLDSSSFEYKWS

TEGSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSFNSENELDRDHLGFKII
DNKTYYYDEDSKLVKGLININNSLFYFDPIESNL

SEQ ID NO: 7 - *C. difficile* Toxin A 542-1850 (toxinotype 0) Cysteine protease negative
LSEDNGVDFNKNTALDKNYLLNNKIPSNNVEEAGSKNYVHYIIQLQGDDISYEATCNLFSKNPKNSIIIQRNM
NESAKSYFLSDDGESILELNKYRIPERLKNKEKVKVTFIGHGKDEFNTSEFARLSVDSLSNEISSFLDTIKLDI
SPKNVEVNLLGANMFSYDFNVEETYPGKLLLSIMDKITSTLPDVNKNSITIGANQYEVRINSEGRKELLAHSG
KWINKEEAIMSDLSSKEYIFFDSIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIESSIGDYI
YYEKLEPVKNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSVRFINKSNGESVYV
ETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNTLNAAFFIQSLIDYSSNKDVLNDLSTSVKV
QLYAQLFSTGLNTIYDSIQLVNLISNAVNDTINVLPTITEGIPIVSTILDGINLGAAIKELLDEHDPLLKKELEAKV
GVLAINMSLSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKATSVVNYFNHLSESKKYGPLKTED
DKILVPIDDLVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIPSLSIYSAIGIETENLDFSK
KIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLYPGKFYWRFYAFFDYAITTLKPVYEDTNIKIK
LDKDTRNFIMPTITTNEIRNKLSYSFDGAGGTYSLLLSSYPISTNINLSKDDLWIFNIDNEVREISIENGTIKKGK
LIKDVLSKIDINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVAKSYSLLLSGDKNYLISNLSNTIE
KINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFYNDSTLEFNSKDFIAEDINVFMKDDI
NTITGKYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFVKNSDGHHNTSNFMNLFLDNISFWK
LFGFENINFVIDKYFTLVGKTNLGYVEFICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTGEDI
STSLDFSYEPLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINLDSSSFEYKWSTE
GSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSFNSENELDRDHLGFKIIDN
KTYYYDEDSKLVKGLININNSLFYFDPIEFNL

SEQ ID NO: 8 - *C. difficile* Toxin A 770-1850
MSDLSSKEYIFFDSIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIESSIGDYIYYEKLEPV
KNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSVRFINKSNGESVYVETEKEIFSK
YSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNTLNAAFFIQSLIDYSSNKDVLNDLSTSVKVQLYAQLFS
TGLNTIYDSIQLVNLISNAVNDTINVLPTITEGIPIVSTILDGINLGAAIKELLDEHDPLLKKELEAKVGVLAINMS
LSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKATSVVNYFNHLSESKKYGPLKTEDDKILVPIDD
LVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIPSLSIYSAIGIETENLDFSKKIMMLPNA
PSRVFWWETGAVPGLRSLENDGTRLLDSIRDLYPGKFYWRFYAFFDYAITTLKPVYEDTNIKIKLDKDTRNFI
MPTITTNEIRNKLSYSFDGAGGTYSLLLSSYPISTNINLSKDDLWIFNIDNEVREISIENGTIKKGKLIKDVLSKI
DINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVAKSYSLLLSGDKNYLISNLSNTIEKINTLGLD
SKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFYNDSTLEFNSKDFIAEDINVFMKDDINTITGKYY
VDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFVKNSDGHHNTSNFMNLFLDNISFWKLFGFENIN
FVIDKYFTLVGKTNLGYVEFICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTGEDISTSLDFSY
EPLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINLDSSSFEYKWSTEGSDFILVRY
LEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSFNSENELDRDHLGFKIIDNKTYYYDED
SKLVKGLININNSLFYFDPIEFNL

SEQ ID NO: 9 - *C. difficile* Toxin A 1130-1850
SESKKYGPLKTEDDKILVPIDDLVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIPSLSIY
SAIGIETENLDFSKKIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLYPGKFYWRFYAFFDYAIT
TLKPVYEDTNIKIKLDKDTRNFIMPTITTNEIRNKLSYSFDGAGGTYSLLLSSYPISTNINLSKDDLWIFNIDNEV
REISIENGTIKKGKLIKDVLSKIDINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVAKSYSLLLSG
DKNYLISNLSNTIEKINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFYNDSTLEFNSKD
FIAEDINVFMKDDINTITGKYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFVKNSDGHHNTSN
FMNLFLDNISFWKLFGFENINFVIDKYFTLVGKTNLGYVEFICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNV
VVEPIYNPDTGEDISTSLDFSYEPLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINL
DSSSFEYKWSTEGSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSFNSEN
ELDRDHLGFKIIDNKTYYYDEDSKLVKGLININNSLFYFDPIEFNL

SEQ ID NO: 10 - *C. difficile* Toxin B (toxinotype 0) 543-1852
LGEDDNLDFSQNIVVDKEYLLEKISSLARSSERGYIHYIVQLQGDKISYEAACNLFAKTPYDSVLFQKNIEDS
EIAYYYNPGDGEIQEIDKYKIPSIISDRPKIKLTFIGHGKDEFNTDIFAGFDVDSLSTEIEAAIDLAKEDISPKSIEI
NLLGCNMFSYSINVEETYPGKLLLKVKDKISELMPSISQDSIIVSANQYEVRINSEGRRELLDHSGEWINKEE
SIIKDISSKEYISFNPKENKITVKSKNLPELSTLLQEIRNNSNSSDIELEEKVMLTECEINVISNIDTQIVEERIEE
AKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSHFISFEDISETDEGFSIRFINKETGESIFVETEKTIFSE
YANHITEEISKIKGTIFDTVNGKLVKKVNLDTTHEVNTLNAAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQLF
STGLNTITDAAKVVELVSTALDETIDLLPTLSEGLPIIATIIDGVSLGAAIKELSETSDPLLRQEIEAKIGIMAVNL
TTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLRDKATKVVDYFKHVSLVETEGVFTLLDDKIMMP
QDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSKDL
MVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTNIRINL

DSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKKGD
LIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVDLLSKSYKLLISGELKILMLNSNHIQ
QKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVYMDDSKPSFGYYSNNLKDV
KVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDESGVAEILKFMNRKGNTNTSDSLMSFLESMNI
KSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFIKFNTLETNYTLYVGNRQNMIVEPNYDLDDS
GDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYINEKINVNINDLSIRYVW
SNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEIILSFTPSYYEDGLIGYDLGL
VSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNL

SEQ ID NO: 11 - *C. difficile* Toxin B (toxinotype 3) 543-1852
LGEDDNLDFSQNTVVDKEYLLEKISSLARSSERGYIHYIVQLQGDKISYEAACNLFAKTPYDSVLFQKNIEDS
EIAYYYNPGDGEIQEIDKYKIPSIISDRPKIKLTFIGHGKDEFNTDIFAGLDVDSLSTEIETAIDLAKEDISPKSIEI
NLLGCNMFSYSVNVEETYPGKLLLRVKDKVSELMPSISQDSIIVSANQYEVRINSEGRRELLDHSGEWINKE
ESIIKDISSKEYISFNPKENKIIVKSKNLPELSTLLQEIRNNSNSSDIELEEKVMLAECEINVISNIDTQVVEGRIE
EAKSLTSDSINYIKNEFKLIESISDALYDLKQQNELEESHFISFEDILETDEGFSIRFIDKETGESIFVETEKAIFS
EYANHITEEISKIKGTIFDTVNGKLVKKVNLDATHEVNTLNAAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQL
FSTGLNTITDAAKVVELVSTALDETIDLLPTLSEGLPVIATIIDGVSLGAAIKELSETSDPLLRQEIEAKIGIMAVN
LTAATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELILRDKATKVVDYFSHISLAESEGAFTSLDDKIMMP
QDDLVISEIDFNNNSITLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSKDL
MVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTNIRINL
DSNTRSFIVPVITTEYIREKLSYSFYGSGGTYALSLSQYNMNINIELNENDTWVIDVDNVVRDVTIESDKIKKG
DLIENILSKLSIEDNKIILDNHEINFSGTLNGGNGFVSLTFSILEGINAVIEVDLLSKSYKVLISGELKTLMANSNS
VQQKIDYIGLNSELQKNIPYSFMDDKGKENGFINCSTKEGLFVSELSDVVLISKVYMDNSKPLFGYCSNDLK
DVKVITKDDVIILTGYYLKDDIKISLSFTIQDENTIKLNGVYLDENGVAEILKFMNKKGSTNTSDSLMSFLESMN
IKSIFINSLQSNTKLILDTNFIISGTTSIGQFEFICDKDNNIQPYFIKFNTLETKYTLYVGNRQNMIVEPNYDLDD
SGDISSTVINFSQKYLYGIDSCVNKVIISPNIYTDEINITPIYEANNTYPEVIVLDTNYISEKININDLSIRYVWS
NDGSDFILMSTDEENKVSQVKIRFTNVFKGNTISDKISFNFSDKQDVSINKVISTFTPSYYVEGLLNYDLGLIS
LYNEKFYINNFGMMVSGLVYINDSLYYFKPPIKNL

SEQ ID NO: 12 - *C. difficile* Toxin B 543-1852 Cysteine protease negative
LGEDDNLDFSQNIVVDKEYLLEKISSLARSSERGYIHYIVQLQGDKISYEAACNLFAKTPYDSVLFQKNIEDS
EIAYYYNPGDGEIQEIDKYKIPSIISDRPKIKLTFIGHGKDEFNTDIFAGFDVDSLSTEIEAAIDLAKEDISPKSIEI
NLLGANMFSYSINVEETYPGKLLLKVKDKISELMPSISQDSIIVSANQYEVRINSEGRRELLDHSGEWINKEE
SIIKDISSKEYISFNPKENKITVKSKNLPELSTLLQEIRNNSNSSDIELEEKVMLTECEINVISNIDTQIVEERIEE
AKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSHFISFEDISETDEGFSIRFINKETGESIFVETEKTIFSE
YANHITEEISKIKGTIFDTVNGKLVKKVNLDTTHEVNTLNAAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQLF
STGLNTITDAAKVVELVSTALDETIDLLPTLSEGLPIIATIIDGVSLGAAIKELSETSDPLLRQEIEAKIGIMAVNL
TTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLRDKATKVVDYFKHVSLVETEGVFTLLDDKIMMP
QDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSKDL
MVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTNIRINL
DSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKKGD
LIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVDLLSKSYKLLISGELKILMLNSNHIQ
QKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVYMDDSKPSFGYYSNNLKDV
KVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDESGVAEILKFMNRKGNTNTSDSLMSFLESMNI
KSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFIKFNTLETNYTLYVGNRQNMIVEPNYDLDDS
GDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYINEKINVNINDLSIRYVW
SNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEIILSFTPSYYEDGLIGYDLGL
VSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNL

SEQ ID NO: 13 - *C. difficile* Toxin B 767-1852
SIIKDISSKEYISFNPKENKITVKSKNLPELSTLLQEIRNNSNSSDIELEEKVMLTECEINVISNIDTQIVEERIEE
AKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSHFISFEDISETDEGFSIRFINKETGESIFVETEKTIFSE
YANHITEEISKIKGTIFDTVNGKLVKKVNLDTTHEVNTLNAAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQLF
STGLNTITDAAKVVELVSTALDETIDLLPTLSEGLPIIATIIDGVSLGAAIKELSETSDPLLRQEIEAKIGIMAVNL
TTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLRDKATKVVDYFKHVSLVETEGVFTLLDDKIMMP
QDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSKDL
MVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTNIRINL
DSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKKGD
LIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVDLLSKSYKLLISGELKILMLNSNHIQ
QKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVYMDDSKPSFGYYSNNLKDV
KVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDESGVAEILKFMNRKGNTNTSDSLMSFLESMNI
KSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFIKFNTLETNYTLYVGNRQNMIVEPNYDLDDS
GDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYINEKINVNINDLSIRYVW

SNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEIILSFTPSYYEDGLIGYDLGL
VSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNL

SEQ ID NO: 14 - *C. difficile* Toxin B 1145-1852
MPQDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSK
DLMVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTNIRI
NLDSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKK
GDLIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVDLLSKSYKLLISGELKILMLNSNH
IQQKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVYMDDSKPSFGYYSNNLKD
VKVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDESGVAEILKFMNRKGNTNTSDSLMSFLESMN
IKSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFIKFNTLETNYTLYVGNRQNMIVEPNYDLDD
SGDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYINEKINVNINDLSIRYV
WSNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEIILSFTPSYYEDGLIGYDLG
LVSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNL

SEQ ID NO: 15 - *C. difficile* Toxin B 1350-1852
NVVRDVTIESDKIKKGDLIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVDLLSKSYKL
LISGELKILMLNSNHIQQKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVYMDD
SKPSFGYYSNNLKDVKVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDESGVAEILKFMNRKGNT
NTSDSLMSFLESMNIKSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFIKFNTLETNYTLYVGN
RQNMIVEPNYDLDDSGDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYIN
EKINVNINDLSIRYVWSNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEIILSFT
PSYYEDGLIGYDLGLVSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNL

SEQ ID NO: 16 – Protein domain from amino acid residues 27-401 of *C. difficile* CD2767
SNDKEMRAAWISTVYNLDWPKTKNNEAKQKKEYTDLLDKLKSVGINTAVVQVRPKSDALYKSNINPWSEYL
TGTQGKDPGYDPLPFLIEEAHKRGMEFHAWFNPYRITMADESIDKLPANHPAKKNPSWVVKHGNKYYYDP
GLPEVRKYIVDSIAEVVQNYDIDGVHFDDYFYPGVSFNDTATYQKYGKGQNKDDWRRENVNTLLRDVKASI
KSIKPNVVFGVSPAGIWRNKSSDPTGSDTSGNESYVGTYADTRAWIKQGLIDYVVPQLYWPIGLKAADYSK
LVAWWANEVKGTNVDLYIGQGIYKQGQSSYGGQNIAKEIVQQVTLNRKYSEIKGSMYFSAKDIANSTSIQK
DLKSLYSSSEEPVTPPSNVKV

SEQ ID NO: 17 – CD2767 (27-401) Toxin A (542-1850)-fusion protein in a expression contruct
MGSSHHHHHHSSGLVPRGSHMSNDKEMRAAWISTVYNLDWPKTKNNEAKQKKEYTDLLDKLKSVGINTA
VVQVRPKSDALYKSNINPWSEYLTGTQGKDPGYDPLPFLIEEAHKRGMEFHAWFNPYRITMADESIDKLPA
NHPAKKNPSWVVKHGNKYYYDPGLPEVRKYIVDSIAEVVQNYDIDGVHFDDYFYPGVSFNDTATYQKYGK
GQNKDNWRRENVNTLLRDVKASIKSIKPNVVFGVSPAGIWRNKSSDPTGSDTSGNESYVGTYADTRAWIK
QGLIDYVVPQLYWPIGLKAADYSKLVAWWANEVKGTNVDLYIGQGIYKQGQSSYGGQNIAKEIVQQVTLNR
KYSEIKGSMYFSAKDIANSTSIQKDLKSLYSSSEEPVTPPSNVKVAAAPFTLSEDNGVDFNKNTALDKNYLL
NNKIPSNNVEEAGSKNYVHYIIQLQGDDISYEATCNLFSKNPKNSIIIQRNMNESAKSYFLSDDGESILELNKY
RIPERLKNKEKVKVTFIGHGKDEFNTSEFARLSVDSLSNEISSFLDTIKLDISPKNVEVNLLGSNMFSYDFNV
EETYPGKLLLSIMDKITSTLPDVNKNSITIGANQYEVRINSEGRKELLAHSGKWINKEEAIMSDLSSKEYIFFD
SIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIESSIGDYIYYEKLEPVKNIIHNSIDDLIDEF
NLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSVRFINKSNGESVYVETEKEIFSKYSEHITKEISTIKN
SIITDVNGNLLDNIQLDHTSQVNTLNAAFFIQSLIDYSSNKDVLNDLSTSVKVQLYAQLFSTGLNTIYDSIQLVN
LISNAVNDTINVLPTITEGIPIVSTILDGINLGAAIKELLDEHDPLLKKELEAKVGVLAINMSLSIAATVASIVGIGA
EVTIFLLPIAGISAGIPSLVNNELILHDKATSVVNYFNHLSESKKYGPLKTEDDKILVPIDDLVISEIDFNNNSIKL
GTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIPSLSIYSAIGIETENLDFSKKIMMLPNAPSRVFWWETGAV
PGLRSLENDGTRLLDSIRDLYPGKFYWRFYAFFDYAITTLKPVYEDTNIKIKLDKDTRNFIMPTITTNEIRNKLS
YSFDGAGGTYSLLLSSYPISTNINLSKDDLWIFNIDNEVREISIENGTIKKGKLIKDVLSKIDINKNKLIIGNQTID
FSGDIDNKDRYIFLTCELDDKISLIIEINLVAKSYSLLLSGDKNYLISNLSNIIEKINTLGLDSKNIAYNYTDESNN
KYFGAISKTSQKSIIHYKKDSKNILEFYNDSTLEFNSKDFIAEDINVFMKDDINTITGKYYVDNNTDKSIDFSISL
VSKNQVKVNGLYLNESVYSSYLDFVKNSDGHHNTSNFMNLFLDNISFWKLFGFENINFVIDKYFTLVGKTNL
GYVEFICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTGEDISTSLDFSYEPLYGIDRYINKVLI
APDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINLDSSSFEYKWSTEGSDFILVRYLEESNKKILQKIRIK
GILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSFNSENELDRDHLGFKIIDNKTYYYDEDSKLVKGLININNSLF
YFDPIEFNL

SEQ ID 18 –CD2767 (27-401) Toxin 6 (767-1852)-fusion protein in a expression contruct
SNDKEMRAAWISTVYNLDWPKTKNNEAKQKKEYTDLLDKLKSVGINTAVVQVRPKSDALYKSNINPWSEYL
TGTQGKDPGYDPLPFLIEEAHKRGMEFHAWFNPYRITMADESIDKLPANHPAKKNPSWVVKHGNKYYYDP
GLPEVRKYIVDSIAEVVQNYDIDGVHFDDYFYPGVSFNDTATYQKYGKGQNKDDWRRENVNTLLRDVKASI
KSIKPNVVFGVSPAGIWRNKSSDPTGSDTSGNESYVGTYADTRAWIKQGLIDYVVPQLYWPIGLKAADYSK

LVAWWANEVKGTNVDLYIGQGIYKQGQSSYGGQNIAKEIVQQVTLNRKYSEIKGSMYFSAKDIANSTSIQK
DLKSLYSSSEEPVTPPSNVKVSIIKDISSKEYISFNPKENKITVKSKNLPELSTLLQEIRNNSNSSDIELEEKVM
LTECEINVISNIDTQIVEERIEEAKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSHFISFEDISETDEGFS
IRFINKETGESIFVETEKTIFSEYANHITEEISKIKGTIFDTVNGKLVKKVNLDTTHEVNTLNAAFFIQSLIEYNSS
KESLSNLSVAMKVQVYAQLFSTGLNTITDAAKVVELVSTALDETIDLLPTLSEGLPIIATIIDGVSLGAAIKELSE
TSDPLLRQEIEAKIGIMAVNLTTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLRDKATKVVDYFKH
VSLVETEGVFTLLDDKIMMPQDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPH
LSIYDVLEVQKEELDLSKDLMVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEGEFYWRYFAFI
ADALITTLKPRYEDTNIRINLDSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYNMGINIELSESDVWII
DVDNVVRDVTIESDKIKKGDLIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVDLLSK
SYKLLISGELKILMLNSNHIQQKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVY
MDDSKPSFGYYSNNLKDVKVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDESGVAEILKFMNRK
GNTNTSDSLMSFLESMNIKSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFIKFNTLETNYTLY
VGNRQNMIVEPNYDLDDSGDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDA
NYINEKINVNINDLSIRYVWSNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEII
LSFTPSYYEDGLIGYDLGLVSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNL

SEQ ID 19 – Toxin A (residues 542 -1850; TxACPD) within the construct: 6His-Thioredoxin-TxACPD
MGSSHHHHHHSHMASDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKL
NIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLARALVPRGSVTSLYKKAGSAAAP
FTLSEDNGVDFNKNTALDKNYLLNNKIPSNNVEEAGSKNYVHYIIQLQGDDISYEATCNLFSKNPKNSIIIQRN
MNESAKSYFLSDDGESILELNKYRIPERLKNKEKVKVTFIGHGKDEFNTSEFARLSVDSLSNEISSFLDTIKLD
ISPKNVEVNLLGSNMFSYDFNVEETYPGKLLLSIMDKITSTLPDVNKNSITIGANQYEVRINSEGRKELLAHS
GKWINKEEAIMSDLSSKEYIFFDSIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIESSIGD
YIYYEKLEPVKNIIHNSIDDLIDEFNLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSVRFINKSNGESVY
VETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNTLNAAFFIQSLIDYSSNKDVLNDLSTSVK
VQLYAQLFSTGLNTIYDSIQLVNLISNAVNDTINVLPTITEGIPIVSTILDGINLGAAIKELLDEHDPLLKKELEAK
VGVLAINMSLSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKATSVVNYFNHLSESKKYGPLKTE
DDKILVPIDDLVISEIDFNNNSIKLGTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIPSLSIYSAIGIETENLDFS
KKIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLYPGKFYWRFYAFFDYAITTLKPVYEDTNIKI
KLDKDTRNFIMPTITTNEIRNKLSYSFDGAGGTYSLLLSSYPISTNINLSKDDLWIFNIDNEVREISIENGTIKKG
KLIKDVLSKIDINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVAKSYSLLLSGDKNYLISNLSNIIE
KINTLGLDSKNIAYNYTDESNNKYFGAISKTSQKSIIHYKKDSKNILEFYNDSTLEFNSKDFIAEDINVFMKDDI
NTITGKYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFVKNSDGHHNTSNFMNLFLDNISFWK
LFGFENINFVIDKYFTLVGKTNLGYVEFICDNNKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTGEDI
STSLDFSYEPLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNINLDSSSFEYKWSTE
GSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSIDFKDIKKLSLGYIMSNFKSFNSENELDRDHLGFKIIDN
KTYYYDEDSKLVKGLININNSLFYFDPIEFNL

SEQ ID 20 - Toxin B (residues 767 -1852; TxBc) within the construct: 6His-Thioredoxin-TxBc
MGSSHHHHHHSHMASDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKL
NIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLARALVPRGSVTSLYKKAGSAAAP
FTSIIKDISSKEYISFNPKENKITVKSKNLPELSTLLQEIRNNSNSSDIELEEKVMLTECEINVISNIDTQIVEERIE
EAKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSHFISFEDISETDEGFSIRFINKETGESIFVETEKTIFS
EYANHITEEISKIKGTIFDTVNGKLVKKVNLDTTHEVNTLNAAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQL
FSTGLNTITDAAKVVELVSTALDETIDLLPTLSEGLPIIATIIDGVSLGAAIKELSETSDPLLRQEIEAKIGIMAVN
LTTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLRDKATKVVDYFKHVSLVETEGVFTLLDDKIMMP
QDDLVISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSKDL
MVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTNIRINL
DSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKKGD
LIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVDLLSKSYKLLISGELKILMLNSNHIQ
QKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVYMDDSKPSFGYYSNNLKDV
KVITKDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDESGVAEILKFMNRKGNTNTSDSLMSFLESMNI
KSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFIKFNTLETNYTLYVGNRQNMIVEPNYDLDDS
GDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYINEKINVNINDLSIRYVW
SNDGNDFILMSTSEENKVSQVKIRFVNVFKDKTLANKLSFNFSDKQDVPVSEIILSFTPSYYEDGLIGYDLGL
VSLYNEKFYINNFGMMVSGLIYINDSLYYFKPPVNNL KGGRADPAFLYKVVSAWSHPQFEK

CLAIMS

1. A polypeptide, consisting of or comprising an amino acid sequence having at least 80% sequence identity with an amino acid sequence consisting of residues 1500-1700 of a *C. difficile* Toxin A sequence or a *C. difficile* Toxin B sequence;

with the proviso that the polypeptide does not include one or more Repeat Unit (RU) located between amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or residues 1853-2366 of a *C. difficile* Toxin B.

2. The polypeptide according to Claim 1, wherein said polypeptide consists of or comprises an amino acid sequence having at least 80% sequence identity with an amino acid sequence consisting of residues 1450-1750 of a *C. difficile* Toxin A sequence or a *C. difficile* Toxin B sequence.

3. The polypeptide according to Claim 1 or Claim 2, wherein said polypeptide consists of or comprises an amino acid sequence having at least 80% sequence identity with an amino acid sequence consisting of residues 1400-1800 of a *C. difficile* Toxin A sequence or a *C. difficile* Toxin B sequence.

4. The polypeptide according to any preceding claim, wherein said polypeptide does not include the amino acid sequence provided by amino acid residues 1851-2710 of *C. difficile* Toxin A and/ or amino acid residues 1853-2366 of a *C. difficile* Toxin B.

5. The polypeptide according to any preceding claim, wherein the polypeptide consists of or comprises an amino acid sequence having at least 80% sequence identity with an amino acid sequence consisting of residues 1330-1850 or 770-1850 or 542-1850 of a *C. difficile* Toxin A sequence.

6. The polypeptide according to any of Claims 1-4, wherein the polypeptide consists of or comprises an amino acid sequence having at least 80% sequence identity with an amino acid sequence consisting of residues 1145-1852 or 767-1852 or 543-1852 of a *C. difficile* Toxin B sequence.

7. The polypeptide according to any preceding claim, wherein said polypeptide further comprises a second amino sequence selected from one or more of a bacterial surface layer protein, NusA, thioredoxin, maltose-binding protein, a ubiquitin-like peptide, and a fragment of any of said second amino acid sequences.

8. The polypeptide according to Claim 7, wherein the second amino acid sequence comprises a *C. difficile* surface layer protein.

9. The polypeptide according to Claim 8, wherein the *C. difficile* surface layer protein consists of or comprises a CD 2767 peptide, for example a peptide consisting of or comprising an amino acid sequence that has at least 80% sequence identity with amino acid residues 27-401 of a CD 6767 peptide.

10. A polypeptide according to any preceding claim, for use as an antigen in a method of generating an antibody that binds to a *C. difficile* Toxin A and/ or to a *C. difficile* Toxin B;

preferably wherein said antibody comprises whole IgG, an Fab antibody or an $F(ab')_2$ antibody.

11. An *in vitro* method for isolating antibodies that bind to *C. difficile* Toxin A and/ or to *C. difficile* Toxin B, said method comprising

a) immobilising on a surface (for example on a matrix within a column) one or more *C. difficile* polypeptide(s) according to any of Claims 1-7;
b) contacting the immobilised polypeptide(s) with a solution containing antibodies that bind to *C. difficile* Toxin A and/ or Toxin B;
c) allowing said antibodies to bind to said polypeptide(s), thereby forming a bound complex of antibody and polypeptide(s);
d) washing away any unbound antibody or protein; and
e) eluting the bound antibodies from the surface, thereby providing affinity-purified antibodies.

12. An antibody that binds to a *C. difficile* Toxin A and/ or to a *C. difficile* Toxin B, wherein said antibody is obtainable by a method according to Claim 8 or Claim 9, and wherein said antibody does not substantially bind to the effector domain of a *C. difficile* Toxin A or Toxin B and/ or does not substantially bind to a RU of a *C. difficile* Toxin A or Toxin B;

preferably wherein said antibody comprises whole IgG, an Fab antibody or an $F(ab')_2$ antibody.

13. An antibody according to Claim 12, wherein said antibody binds to an epitope presented by residues 1350-1850 or 1130-1850 or 770-1850 or 542-1185 of a *C. difficile* Toxin A and/ or by residues 1350-1852 or 1145-1852 or 767-1852 or 543-1852 of a *C. difficile* Toxin B.

14. A polypeptide according to any of Claims 1-9 and/ or an antibody according to Claim 12 or Claim 13, for use in the prevention, treatment or suppression of CDI in a mammal (e.g. a human).

15. A method for preventing, treating or suppressing CDI in a mammal (e.g. a human), said method comprising administering to said human a therapeutically effective amount of a polypeptide according to any of Claims 1-9 and/ or an antibody according to Claim 12 or Claim 13.

16. A method according to Claim 15, wherein said polypeptide is administered as part of a combination therapy with one or more additional antigen selected from a *C. difficile* non-Toxin antigen, an inactivated or attenuated whole cell *C. difficile* bacterium, and/ or a *C. difficile* cell extract; and

a. optionally one or more antigen selected from 1) an antigen from a bacterium that causes nosocomial infection, and 2) an inactivated or attenuated whole cell bacterium that causes nosocomial infection; and
b. optionally one or more antibiotic, such as an antibiotic effective against a bacterium that causes nosocomial infection.

17. A method according to Claim 15, wherein said antibody is administered as part of a combination therapy with one or more additional antibodies selected

from an antibody that binds to a *C. difficile* non-Toxin antigen, and an antibody that binds to a whole cell *C. difficile* bacterium; and

a. optionally one or more an antibody selected from 1) an antibody that binds to antigen from a bacterium that causes nosocomial infection, and 2) an antibody that binds to a whole cell bacterium that causes nosocomial infection; and
b. optionally one or more antibiotic, such as an antibiotic effective against a bacterium that causes nosocomial infection.

18. A polypeptide for use according to Claim 14, wherein said use includes administration of said polypeptide as part of a combination therapy with one or more additional antigen selected from a *C. difficile* non-Toxin antigen, an inactivated or attenuated whole cell *C. difficile* bacterium, and/ or a *C. difficile* cell extract; and

a. optionally one or more antigen selected from 1) an antigen from a bacterium that causes nosocomial infection, and 2) an inactivated or attenuated whole cell bacterium that causes nosocomial infection; and
b. optionally one or more antibiotic, such as an antibiotic effective against a bacterium that causes nosocomial infection.

19. An antibody for use according to Claim 14, wherein said use includes administration of said antibody as part of a combination therapy with one or more additional antibodies selected from an antibody that binds to a *C. difficile* non-Toxin antigen, and an antibody that binds to a whole cell *C. difficile* bacterium; and

a. optionally one or more an antibody selected from 1) an antibody that binds to antigen from a bacterium that causes nosocomial infection, and 2) an antibody that binds to a whole cell bacterium that causes nosocomial infection; and
b. optionally one or more antibiotic, such as an antibiotic effective against a bacterium that causes nosocomial infection.

20. Use of an antibody according to Claim 12 or Claim 13 in an *in vitro* immunoassay method for confirming the presence or absence of a *C. difficile*

infection in a patient sample, wherein the presence of a *C. difficile* infection is confirmed by detecting the binding of said antibody to a *C. difficile* Toxin A and/ or Toxin B present in said sample, and wherein failure to detect the binding of said antibody to a *C. difficile* Toxin A and/ or Toxin B present in said sample confirms the absence of a *C. difficile* infection.

21. Use of a polypeptide according to any of Claims 1-9 in an *in vitro* immunoassay method for confirming the presence or absence of a *C. difficile* infection in a patient sample, wherein the presence of a *C. difficile* infection is confirmed by detecting the binding of said fusion protein to an antibody present in said sample, and wherein failure to detect the binding of said fusion protein to an antibody present in said sample confirms the absence of a *C. difficile* infection.