HU000033342T2

(19) **HU**

(11) Lajstromszám: **E 033 342**

(13) **T2**

MAGYARORSZÁG
Szellemi Tulajdon Nemzeti Hivatala

EURÓPAI SZABADALOM
SZÖVEGÉNEK FORDÍTÁSA

(21) Magyar ügyszám: **E 11 769799**
(22) A bejelentés napja: **2011. 09. 05.**

(51) Int. Cl.: ***A61K 39/08*** *(2006.01)*
C07K 14/33 *(2006.01)*

(96) Az európai bejelentés bejelentési száma:
EP 20110769799

(86) A nemzetközi (PCT) bejelentési szám:
PCT/EP 11/065304

(97) Az európai bejelentés közzétételi adatai:
EP 2753352 A1 **2012. 03. 08.**

(87) A nemzetközi közzétételi szám:
WO 12028741

(97) Az európai szabadalom megadásának meghirdetési adatai:
EP 2753352 B1 **2017. 01. 25.**

(72) Feltaláló(k):
ELLINGSWORTH, Larry, Rockville Maryland 20850 (US)
FLYER, David, Olney Maryland 20832 (US)
TIAN, Jing-Hui, Germantown Maryland 20874 (US)
FUHRMANN, Steven, Germantown Maryland 20874 (US)
KLUEPFEL-STAHL, Stefanie, Bethesda Maryland 20817 (US)
GLENN, Gregory, Poolesville Maryland 20837 (US)
WESTRITSCHNIG, Kerstin, A-1190 Vienna (AT)

(73) Jogosult(ak):
Valneva Austria GmbH, 1030 Vienna (AT)
Intercell USA, Inc., Gaithersburg, MD 20878 (US)

(74) Képviselő:
SBGK Szabadalmi Ügyvivői Iroda, Budapest

(54) **C. difficile toxin A és toxin B proteinek izolált polipeptidei és alkalmazásaik**

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 753 352 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
25.01.2017 Bulletin 2017/04

(21) Application number: **11769799.5**

(22) Date of filing: **05.09.2011**

(51) Int Cl.:
A61K 39/08 *(2006.01)* ***C07K 14/33*** *(2006.01)*

(86) International application number:
PCT/EP2011/065304

(87) International publication number:
WO 2012/028741 (08.03.2012 Gazette 2012/10)

(54) **ISOLATED POLYPEPTIDE OF THE TOXIN A AND TOXIN B PROTEINS OF C. DIFFICILE AND USES THEREOF**

ISOLIERTES POLYPEPTID DER TOXIN A UND TOXIN B AUS C. DIFFICILE UND VERWENDUNG DAVON

POLYPEPTIDE ISOLÉ DE LA TOXINE A ET LA TOXINE B DE C. DIFFICILE ET SES UTILISATIONS

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(43) Date of publication of application:
16.07.2014 Bulletin 2014/29

(60) Divisional application:
16204515.7

(73) Proprietors:
- **Valneva Austria GmbH**
  1030 Vienna (AT)
- **Intercell USA, Inc.**
  Gaithersburg, MD 20878 (US)

(72) Inventors:
- **ELLINGSWORTH, Larry**
  Rockville
  Maryland 20850 (US)
- **FLYER, David**
  Olney
  Maryland 20832 (US)
- **TIAN, Jing-Hui**
  Germantown
  Maryland 20874 (US)
- **FUHRMANN, Steven**
  Germantown
  Maryland 20874 (US)
- **KLUEPFEL-STAHL, Stefanie**
  Bethesda
  Maryland 20817 (US)
- **GLENN, Gregory**
  Poolesville
  Maryland 20837 (US)
- **WESTRITSCHNIG, Kerstin**
  A-1190 Vienna (AT)

(74) Representative: **Dempster, Robert Charles et al**
Maschio & Soames IP Limited
30 Carlton Crescent
Southampton SO15 2EW (GB)

(56) References cited:
WO-A1-2010/017383 **US-B1- 6 733 760**

- **BELYI IOURI F ET AL: "Construction of a fusion protein carrying antigenic determinants of enteric clostridial toxins.", FEMS MICROBIOLOGY LETTERS, vol. 225, no. 2, 29 August 2003 (2003-08-29), pages 325-329, XP002666608, ISSN: 0378-1097 cited in the application**

Remarks:
The file contains technical information submitted after the application was filed and not included in this specification

EP 2 753 352 B1

Description

FIELD OF THE INVENTION

[0001] Disclosed herein is an isolated polypeptide containing the receptor binding domains of the *Clostridium difficile* toxin A and toxin B and its use as a vaccine. This isolated polypeptide provides anti-toxin immunity to both toxins.

BACKGROUND OF THE INVENTION

[0002] *Clostridium difficile* is the leading cause of nosocomial antibiotic associated diarrhea and has become a major health problem in hospitals, nursing home and other care facilities. The cost to hospitals has been estimated to be 2 billion dollars in Europe and 3.2 billion dollars in the United States.

[0003] The causative agent is a gram positive, spore forming anaerobic bacterium, commonly found through out the environment but also present in the intestinal tract of 2-3% of the healthy adult population. *C. difficile* associated disease (CDAD) is induced by the disruption of the normal colonic flora, usually the result of the administration of antibiotics. Following exposure to C. *difficile* spores in the environment, the organism may colonize the intestinal mucosa where the production of disease causing toxins can result in CDAD. Disease may range from mild uncomplicated diarrhea to severe pseudomembranous colitis and toxic megacolon.

[0004] CDAD has become increasingly more problematic in health care settings. A recent study reported that 31% of hospital patients who receive antibiotics become colonized with C. *difficile* and 56% of those patients who become colonized go on to develop CDAD. Overall, C. *difficile* is responsible for 10-25% of all antibiotic associated diarrheas, 50-75% of antibiotic related colitis and 90-100% of antibiotic related pseudomembranous colitis. Treatment of CDAD involves discontinuation of the causal antibiotic followed by treatment with either metronidazole or vancomycin. Relapsing after antibiotic treatment is discontinued occurs in approximately 20% of patients, often the result of recolonization by C. *difficile.*

[0005] In 2003, a *C. difficile* outbreak in Quebec, Canada indicated the emergence of a more virulent strain of *C. difficile* known as North American Phenotype 1/027 (NAP1). NAP1 has been associated with greater virulence, poor outcomes and greater morbidity and mortality rates compared to previous strains. The emergence of this strain adds to the problems already encountered in trying to contain the incidence of CDAD

[0006] Fidaxomicin (*Dificid©*) *for prevention of recurrent disease* is the first in a new class of narrow spectrum macrocyclic antibiotic drugs (Revill, P.; Serradell, N.; Bolos, J. (2006). "Tiacumicin B: macrolide antibiotic treatment of C. difficile-associated diarrhea". Drugs of the Future 31 (6): 494-497). It is a fermentation product obtained from the actinomycete *Dactylosporangium aurantiacum subspecies hamdenesis.* Fidaxomicin is non-systemic, meaning it is minimally absorbed into the bloodstream, it is bactericidal, and it has demonstrated selective eradication of pathogenic *Clostridium difficile* with minimal disruption to the multiple species of bacteria that make up the normal, healthy intestinal flora. The maintenance of normal physiological conditions in the colon can reduce the probability of *Clostridium difficile* infection recurrence (Johnson, Stuart (2009-06). "Recurrent Clostridium difficile infection: a review of risk factors, treatments, and outcomes". Journal of Infection 58 (6): 403-410). Although it is thought, that the introduction of this new class of antibiotic drug will improve the treatment of CDAD, there is still a medical need for a preventative drug, in particular for high risk patients such as the elderly and the immunocompromised patients.

[0007] CDAD is the result of the actions of two exotoxins produced by C. *difficile*, toxin A and toxin B (also referred to as CTA and CTB, respectively). Both toxins are high molecular weight (~300 kDa) secreted proteins that possess multiple functional domains (Voth DE and Ballard JD, Clinical Microbiology Reviews 18:247-263 (2005)). The N-terminal domain of both toxins contains ADP-glucosyltransferase activity that modifies Rho-like GTPases. This modification causes a loss of actin polymerization and cytoskeletal changes resulting in the disruption of the colonic epithelial tight junctions. This leads to excessive fluid exudation into the colon and a resulting diarrhea. The central domain contains a hydrophobic domain and is predicted to be involved in membrane transport. The C-terminal domain of both toxins contain multiple homologous regions called repeating units (RUs) that are involved in toxin binding to target cells (Ho et al, PNAS 102:18373-18378 (2005)). The repeating units are classified as either short (21-30 amino acids) or long (~50 amino acids). Repeating units combine to form clusters, each usually containing one long and 3 - 5 short repeating units. The full length toxin A possesses 39 repeating units (ARUs) organized into 8 clusters (Dove et al. Infect. Immun. 58:480-488 (1990), while the full length toxin B contains 24 repeating units (BRUs) organized into 5 clusters (Barroso et al., Nucleic Acids Res. 18:4004 (1990); Eichel-Streiber et al., Gene 96:107-113 (1992)).

[0008] A number of studies, from both animal models and from the clinic, have indicated a role for anti-toxin antibody in the protection from *C. difficile* associated disease. Hamsters immunized with formalin inactivated toxin A and toxin B generated high levels of anti-toxin antibody and were protected from a lethal challenge of *C. difficile* bacteria (Giannasca PJ and Warny M, Vaccine 22:848-856 (2004)). In addition, passive transfer of mouse anti-toxin antibody protected hamsters in a dose dependent manner. Kyne L et al. (The Lancet 357:189-193 (2001)) reported that the development

of an anti-toxin A antibody response during an initial episode of CDAD correlated with protection against disease recurrence.

[0009] The determinants recognized by protective anti-toxin antibodies have been localized to the C-terminal domain containing the repeating units which function as the receptor binding domain. Initially, Lyerly et al. (Current Microbiology 21:29-32 (1990)) revealed that the toxin A C-terminal domain containing 33 repeating units is capable of inducing the production of neutralizing anti-toxin antibody and may protect from *C. difficile* infection. In this study hamsters were injected subcutaneously with the purified recombinant polypeptide multiple times prior to challenge with the bacteria, however only partial protection was achieved. Another study (Ryan et al., Infect. Immun. 65:2941-49 (1997)) showed that the isolated polypeptide containing 720 amino acid residues from the C-terminus of CTA and the secretion signal of *E. coli* hemolysin A (expressed in *Vibrio cholerae*) induced protective systemic and mucosal immunity against a small dose of CTA in the rabbit CDAD model.

[0010] It was also reported that antibody response against the C-terminal domain of both toxin A and B was necessary to achieve full protection (Kink and Williams, Infect. Immun. 66:2018-25 (1998), U.S. Pat. No. 5,736,139 (1998)). This study revealed that the C-terminal domain of each toxin was most effective in generating toxin-neutralizing antibodies. It demonstrated the effectiveness of orally delivered avian antibodies (antitoxin) raised against C-terminal domain of CTA and CTB in the hamster lethal model. The results also indicate that the antitoxin may be effective in the treatment and management of CD AD in humans. In another study, human anti-toxin A and B monoclonal antibodies were reported confer protection against C. *difficile* induced mortality in hamsters (Babcock et al., Infect. Immun. 74:6339-6347 (2006)). Protection was only observed by antibodies directed against the receptor binding domain of either toxin and enhanced protection was observed following treatment with both anti-toxin A and B antibodies.

[0011] On the other hand, Ward et al. (Infect. Immun. 67: 5124-32 (1999)) considered 14 repeating units from C. *difficile* toxin A (14 CTA) for the study of adjuvant activity. The repeating units were cloned and expressed either with the N-terminal polyhistidine tag (14 CTA-HIS) or fused to the nontoxic binding domain from tetanus toxin (14 CTA-TETC). Both fusion proteins administered intranasally generated anti-toxin A serum antibodies but no response at the mucosal surface in mice. Enhanced systemic and mucosal anti-toxin A responses were seen following co-administration with *E. coli* heat-labile toxin (LT) or its mutated form LTR72. Based on the data, Ward *et al.* suggested using non-toxic 14 CTA-TETC fusion as a mucosal adjuvant in human vaccine directed against clostridial pathogens.

[0012] Recent biochemical studies on the repeating unit domains of C. *difficile* toxins has looked at the minimal sequence requirements for forming stable tertiary structure (Demarest SJ et al., J. Mol. Bio. 346:1197-1206 (2005)). An 11 repeating unit peptide derived from toxin A was found with a correct tertiary structure but 6 and 7 repeating units from toxins A and B did not. The correctly folded 11 repeating unit segment was found to maintain the receptor binding property. A second study examined the functional properties of toxin A fragments containing 6, 11 or 15 repeating units (Dingle T, Glycobiology 18:698-706 (2008)). Only the 11 and 15 repeat units were capable of competitively inhibiting the toxin neutralizing ability of anti-toxin A antibody. While all 3 fragments were found to have hemagglutinating activity, the longer fragments displayed higher hemagglutinating activity than the shorter ones. The data indicates that toxin receptor binding domain structure and immunogenicity are retained in domain fragments that contain greater than 11-14 repeats.

[0013] Thomas et al. (WO97/02836, U.S. Pat. No. 5,919,463 (1999)) also disclosed *C. difficile* toxin A, toxin B and certain fragments thereof (*e.g.*, C-terminal domain containing some or all of the repeating units) as mucosal adjuvants. They showed that intranasal administration of CTA or CTB significantly enhanced mucosal immune response to a heterologous antigen such as *Helicobacter pylori* urease, ovalbumin, or keyhole limpet hemocyanin (KLH) in multiple mouse compartments and was associated with protection against the challenge with *Helicobacter.* Additionally, the adjuvant activity of a toxin A fusion protein was evaluated: 794 C-terminal amino acid residues of CTA comprising ARUs (toxin A repeating units) were fused to glutatione-S-transferase (GST) and resulted polypeptide GST-ARU was expressed in *E. coli.* This study demonstrated significant enhancement of immune response by GST-ARU to co-administered antigens in serum and mucosal secretions.

[0014] All of these studies suggest potential use of a non-toxic, recombinant protein comprising either C. *difficile* toxin A, or toxin B, or fragments thereof, or their combinations for producing an active vaccine against CDAD. Currently, no vaccine against *C. difficile* is commercially available, although a candidate vaccine consisting of formalin-detoxified entire toxins A and B has been evaluated in human phase I and IIa studies. It is reported that parenteral immunization with this vaccine induces anti-toxin IgG and toxin-neutralizing antibody responses (Kotloff KL et al., Infect. Immun. 69:988-995 (2001); Aboudola S et al., Infect. Immun. 71:1608-1610(2003)).

[0015] The literature further indicates that the construction of a recombinant fusion protein containing both toxin A and B receptor binding domains of *C. difficile*, either in their entirety or fragments thereof, would be an efficient and commercially viable approach for vaccine development. Such an approach has been attempted as a two part fusion protein of a 700 base pair fragment of toxin A and a 1300 base pair fragment of toxin B by Varfolomeeva et al. (Mol. Genetics, Microb. and Virol. 3:6-10 (2003)). This approach has also been described by Belyi and Varfolomeeva (FEMS Letters 225:325-9 (2003)) demonstrating construction of the recombinant fusion protein consisting of three parts: two C-terminal

domains composed of repeating units of *C. difficile* toxin A and toxin B followed by the fragment of *Clostridium perfringens* enterotoxin Cpe. The fusion protein was expressed in *E. coli* but the product was accumulated in inclusion bodies and was not stable. Moreover, the yield of pure product achieved in this study (50 μg per 100 ml culture) was considerably low.

[0016] Wilkins et al. (WO 00/61762, U.S. Pat. No. 6,733, 760 (2004)) also described the use of recombinant *C.difficile* toxin A and B repeating units (recombinant ARU and recombinant BRU) and their polysaccharide conjugates for the preparation of a vaccine against CDAD. The resulting recombinant ARU protein comprised 867 amino acid residues while the recombinant BRU protein contains 622 amino acids in length. Unlike the previously mentioned studies, this work demonstrated high-level expression of recombinant ARU and BRU soluble proteins in *E. coli.* Mice vaccinated with recombinant ARU and with polysaccharide-conjugated recombinant ARU both mounted a high level of neutralizing anti-toxin A antibodies and were highly protected against lethal challenge with *C.difficile* toxin A. In addition, Wilkins *et al.* suggested using a recombinant fusion protein consisting of both ARU and BRU for the preparation of a vaccine.

[0017] There is an interest in developing a vaccine against CDAD. A recombinant fusion protein consisting of ARU and BRU may be potentially useful as a vaccine.

[0018] Telfer at al. (WO 2010/017383) proposed using attenuated microorganisms, such as *Salmonella;* expressing on their surface an immunogenic polypeptide comprising *C. difficile* toxin A C-terminal repeat units and/or toxin B C-terminal repeat units for vaccination against CDAD by oral administration to a human subject. The recombinant polypeptide included at least 20 repeats of toxin A and/or 15 repeats of toxin B, and further included a secretion signal, e.g., ClyA tag of S. *typhimurium*.

SUMMARY OF THE INVENTION

[0019] The present invention provides an immunogenic or vaccine composition as defined in the appended claims. Disclosed herein are new tools and methods for the design, production and use of the toxin A and toxin B from *C.difficile.* Disclosed herein is an isolated polypeptide C-TAB comprising SEQ ID NO:2 (C-TAB.G5) or a derivative thereof, SEQ ID NO: 4 (C-TAB.G5.1). The C-TAB.G5 or C-TAB.G5.1 comprises 19 repeating units of the C-terminal domain of toxin A fused to 23 repeating units of the C-terminal domain of toxin B. Disclosed herein are compositions and formulations comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. The compositions or formulations may contain the isolated polypeptide, an additional antigen, an adjuvant, and/or an excipient. Alternatively, the compositions or formulations may consist essentially of the isolated polypeptide without an adjuvant or other active ingredients (but optionally comprising an excipient such as a carrier, buffer and/or stabilizer). Moreover, the compositions or formulations of the invention may be administered concomitantly with other drugs such as an antibiotic in particular e.g. in subjects with recurrent CDAD or in subjects requiring frequent and/or prolonged antibiotic use.

[0020] Disclosed herein is a vaccine comprising the isolated polypeptide described herein. The vaccine may further comprise an adjuvant, such as such as alum, an adjuvant derived from an ADP-ribosylating exotoxin or others. The vaccine may be administered in a one dose regimen, two dose regimen (administered e.g. within 3 to 20 days, e.g. after 10 to 15 days of the first dose), three dose regimen (administered e.g. after about 7 days and about 21 days of the first dose), or more than three dose regimen, preferably a two or three dose regimen, wherein the dose comprises a 20 μg to 200μg amount of the polypeptide of the invention.

[0021] Disclosed herein is a method of preventing, treating, or alleviating one or more symptoms of a disease, such as CDAD by administering the isolated polypeptide described herein to a subject in need thereof. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be administered to the subject intramuscularly or by other routes of delivery.

[0022] Disclosed herein is a method of preventing a disease, such as CDAD by administering the isolated polypeptide described herein or a composition comprising said polypeptide to a subject at risk of CDAD, such as e.g. a subject with the following profile: i) a subject with a weaker immune system such as e.g. an elderly subject (e.g. a subject above 65 years of age) or a subject below 2 years of age; ii) an immunocompromised subject such as e.g. a subject with AIDS; iii) a subject taking or planning to take immunosuppressing drugs; iv) a subject with planned hospitalization or a subject that is in hospital; v) a subject in or expected to go to an intensive care unit (ICU); vi) a subject that is undergoing or is planning to undergo gastrointestinal surgery; vii) a subject that is in or planning to go to a long-term care such as a nursing home; viii) a subject with co-morbidities requiring frequent and/or prolonged antibiotic use; ix) a subject that is a subject with two or more of the above mentioned profiles, such as e.g. an elderly subject that is planning to undergo a gastrointestinal surgery; x) a subject with inflammatory bowel disease; and/or xi) a subject with recurrent CDAD such as e.g. a subject having experienced one or more episodes of CDAD

[0023] Disclosed herein are methods of producing the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be produced from a nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide using a bacterial expression system, such as an *E. coli* expression system.

[0024] Disclosed herein is the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide wherein the 19 repeating units of toxin A are connected to the 23 repeating units of toxin B via a linker consisting of at least 4, 5, 6, 7, 8, 9, or 10 amino acid residues. By way of example, the linker may comprise the sequence RSMH (Arg-Ser-Met-His) (amino acids 439-442 of

SEQ ID NO: 2 or SEQ ID NO: 4).

[0025] Disclosed herein is a variant of the isolated polypeptide that comprises at least one mutation (e.g., insertion, substitution or deletion), for example in the ARU and/or BRU. The sequence of the variant may have 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2.

[0026] Disclosed herein are methods for producing the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide or variants thereof through recombinant DNA engineering, bacterial fermentation and protein purification. Disclosed herein are methods for constructing the nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. Disclosed herein are methods of producing the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide using a bacterial expression system, such as an *E. coli* expression system.

[0027] Disclosed herein are methods for preventing and treating CDAD in subjects in need thereof, such as humans. In this method the C-TAB.G5 or C-TAB.G5.1 is administered to a subject either alone or co-administered with one or more adjuvants such as alum or others. Subjects may be healthy individuals who are at risk for exposure to C. *difficile*, human subjects who have been treated and recovered from C. *difficile* infection and who are at risk for re-infection by *C. difficile*, or human subjects who are currently infected with C. *difficile* and whose condition may be improved by induction of C. *difficile* toxin-neutralizing antibody.

[0028] Disclosed herein is an immunogenic composition comprising C-TAB.G5 or C-TAB.G5.1. The immunogenic composition may further include an adjuvant to enhance an antigen-specific immune response and/or a pharmaceutically acceptable carrier and/or other components in a formulation suitable for application to a subject in need thereof. The immunogenic composition may be delivered by intramuscular (IM) delivery, intradermal (ID) delivery, subcutaneous (SC) delivery, intraperitoneal (IP) delivery, oral delivery, nasal delivery, buccal delivery, or rectal delivery.

[0029] Disclosed herein is an immunogenic composition that elicits antibodies that bind native C. *difficile* toxins and neutralize their cytotoxic activity thus providing long-term, active protection, and/or treatment against C. *difficile* associated disease (CDAD).

[0030] Disclosed herein are immunogenic compositions useful for the prevention or treatment of C. *difficile* associated disease in subjects in need thereof.

[0031] Disclosed herein are nucleic acids and fragments or variants thereof that encode C-TAB.G5 or C-TAB.G5.1. Disclosed herein are expression vectors comprising the nucleic acid encoding C-TAB.G5 or C-TAB.G5.1.

[0032] Disclosed herein are antibodies and fragments thereof, such as neutralizing, humanized, monoclonal, chimeric and polyclonal antibodies, specific for C-TAB.G5 or C-TAB.G5.1. The antibodies or fragments thereof may recognize toxin A and/or toxin B.

[0033] Disclosed herein is a vaccine comprising a polypeptide having the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4.

[0034] Disclosed herein are diagnostic kits comprising the nucleic acids, polypeptides and/or antibodies described herein.

[0035] Other embodiments and advantages of the invention are set forth in part in the description, which follows.

BRIEF DESCRIPTION OF THE DRAWINGS

[0036]

Figure 1A shows the nucleic acid encoding the C-TAB.G5 isolated polypeptide (SEQ ID NO: 1). Figure 1B shows the amino acid sequence of the C-TAB.G5 isolated polypeptide (SEQ ID NO: 2). The amino acid linker between the toxin A domain and the toxin B domain is underlined.

Figure 2A shows the nucleic acid encoding the C-TAB. G5.1 isolated polypeptide (SEQ ID NO: 3). Figure 2B shows the amino acid sequence of the C-TAB.G5.1 isolated polypeptide (SEQ ID NO: 4). The amino acid linker between the toxin A domain and the toxin B domain is underlined.

Figure 3 shows the enhancement of antibody production in C-TAB.G5 vaccinated mice by increasing doses of C-TAB.G5 and co-delivery with alum adjuvant. Mice received two vaccinations by IM injection. IgG titers for anti-C-TAB, anti-toxin A and anti-toxin B antibodies were evaluated by ELISA two weeks after the first and second injection.

Figure 4 shows a graphical representation of anti-C-TAB, anti-toxin A, and anti-toxin B IgG induction in mice receiving increasing doses of C-TAB.G5 with and without alum by two IM injection.

Figure 5 shows antibody titers over one log dose range in mice immunized with C-TAB.G5 in the presence or absence of alum. IgG titers were evaluated by ELISA two weeks after the second immunization. The data demonstrate that alum significantly augments antibody production in vaccinated mice.

Figure 6 shows protective effect in mice vaccinated with C-TAB.G5 (with and without alum) and then exposed to a lethal dose of toxin A or toxin B. Mice receiving two vaccinations (IM) in two week interval were challenged (IP) three weeks later. Toxin A and toxin B neutralizing antibodies (TNA) were assessed two weeks after the second injection, and the percent of animals survived the lethal challenge was determined. Increased doses of C-TAB.G5 conferred greater TNA production, as well as increased protection to the lethal challenge. The presence of alum further increased TNA production, as well as conferring higher survival at lower doses.

Figure 7 shows a comparison of antibody response and protection efficacy of C-TAB.G5 in vaccinated young (6-7 weeks) and old (18 months) mice. Mice receiving two vaccinations (IM) in two week interval were challenged (IP) three weeks later. ELISA IgG titers for anti-C-TAB, anti-toxin A and anti-toxin B antibodies, TNA production as well as overall survival were assessed. Young mice demonstrated higher antibody response even without alum, and both groups showed improved survival when vaccinated in the presence of alum.

Figure 8 shows a comparison of the kinetics of anti-C-TAB IgG antibody development in vaccinated young and old mice. Young mice demonstrated greater rates and earlier IgG production, and both groups demonstrated improved responses when vaccinated in the presence of alum.

Figure 9 shows a comparison in anti-C-TAB, anti-toxin A and anti-toxin B antibody production in mice immunized with either C-TAB.G5.1 or toxoid A and B mixture (1:1). Mice received two vaccinations IM injection. IgG titers for anti-C-TAB, anti-toxin A and anti-toxin B antibodies were evaluated by ELISA two weeks after the second injection. Immunization with toxoid induces antibody to the N-terminal portion of the toxin molecule while immunization with C-TAB induces antibody to the C-terminal portion of the toxin molecule.

Figure 10 shows a comparison in TNA production and protection against challenge with toxin A or B in mice immunized with either C-TAB.G5.1 or toxoid A and B mixture. Mice receiving two vaccinations (IM) in two week interval were challenged (IP) three weeks later with a lethal dose of toxin A or toxin B.

Figure 11 shows anti-C-TAB (A), anti-toxin A (B), and anti-toxin B (C) IgG production in hamsters immunized with C-TAB.G5.1 with and without alum. Hamsters received three vaccinations by IM injection on day 0 and day 14. IgG titers for anti-C-TAB, anti-toxin A and anti-toxin B antibodies were evaluated by ELISA on days 14, 28 and 35.

Figure 12 shows a graphical representation of anti-C-TAB IgG antibody development in hamsters immunized with C-TAB.G5.1 with or without alum.

Figure 13 shows a comparison in TNA and protection in hamsters immunized with C-TAB. G5.1 with or without alum. Two weeks after the third vaccination hamsters received a lethal dose of toxin A or toxin B by IP injection.

Fig 14 shows survival of hamsters vaccinated with C-TAB.G5.1 following the intragastric administration of a lethal dose of C. *difficile* spores. Survival data was plotted as Kaplan-Meier survival fit curves and statistical analysis was done using a log rank analysis. At all spore doses (10^2, 10^3 and 10^4), 100 % survival of hamsters in the vaccinated group was observed and survival was significantly enhanced when compared to the placebo group.

Figure 15 shows anti-C-TAB, anti-toxin A, and anti-toxin B antibody production in cyanomologous monkeys immunized with C-TAB.G5.1 in the presence or absence of alum. Two groups of monkeys (three per group, 4-6 years) received 200 μg of C-TAB.G5.1 with or without 250 μg alum. Blood samples were taken on study days 0, 14, 28 and 42. ELISA method was used to assess anti-C-TAB, anti-toxin A and anti-toxin B IgG titers.

Figure 16 shows a comparison of immunogenicity of C-TAB.G5 and C-TAB.G5.1 delivered over a 1 μg - 30 μg dose range either in PBS or histidine buffer. Mice received two vaccinations (IM) in two week interval. IgG titers for anti-C-TAB, anti-toxin A and anti-toxin B antibodies were evaluated by ELISA two weeks after the second injection. All three antibody titers were not significantly different (T-test analysis) between C-TAB.G5 delivered in PBS or histidine buffer and C-TAB.G5.1 delivered in histidine buffer.

Figure 17 shows a comparison of immunogenicity of C-TAB.G5, C-TABNCTB and C-TADCTB in mice. Mice received two vaccinations of each recombinant protein in two week interval by IM injection. All immunizations were done in the absence of alum adjuvant. IgG titers for anti-C-TAB, anti-toxin A and anti-toxin B antibodies were evaluated by ELISA two weeks after the second injection. All three fusion proteins demonstrate high immunogenicity.

Figure 18 shows protection against challenge with native toxin B in mice. Mice were immunized as indicated for Figure 17 and three weeks later they were challenged by IP injection with a lethal dose of native toxin B.

Figure 19 shows a comparison in TNA and protection in hamsters vaccinated with either C-TAB.G5.1 or C-TADCTB in the absence or presence of alum. Two weeks after the third vaccination hamsters received a lethal dose of toxin A or toxin B by IP injection.

Figure 20 shows TNA production and protection against challenge with toxin A or toxin B in mice immunized with C-TAB.G5.1 in different regimens. Comparison in TNA production and protection between groups of mice vaccinated by IM injection three times on day 0, 3 and 14, or on day 0, 7 and 21, or on day 0, 14 and 28. Three weeks after the last injection mice were challenged with a lethal dose of toxin A or toxin B (panel A is in table form and panel B is in graph form).

Figure 21 shows protection (survival) against challenge with C. *difficile* toxin A (55 ng/mouse)in mice immunized with a single shot of 10μg C-TAB.G5.1 and 12.5 μg alum (in 100 μl). Said challenge was done 21 days, 35 days or 49 days after immunization.

DETAILED DESCRIPTION

General Description

[0037] Disclosed herein is an immunogenic composition for inducing protective and/or therapeutic immune responses to C. *difficile* toxins A and B comprising use of a isolated polypeptide C-TAB.G5 (SEQ ID NO: 2) or a derivative thereof, C-TAB.G5.1 (SEQ ID NO: 4). that comprises 19 repeating units (RU) of toxin A and 23 repeating units (RU) of toxin B or peptide fragments, or variants thereof.

[0038] Disclosed herein are methods of producing the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide and the method of preparing the composition (e.g. a vaccine) useful for prevention and/or treatment of CDAD in mammals. The following description provides more details and examples for the construction, expression, and purification of the recombinant isolated polypeptides, their use as antigens for inducing a specific-immune response as well as evaluating protection in subjects. The subjects may be animals or humans.

[0039] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides for use in the methods and compositions described herein may be prepared using any of several standard methods. For example, the C-TAB.G5 or C-TAB.G5.1 may be produced using standard recombinant DNA techniques, wherein a suitable host cell is transformed with an appropriate expression vector containing a part of a toxin-encoding nucleic acid fragment (see *e.g.* Dove et al., Infect. Immun. 58:480-8 (1990), and Barroso et al., Nucleic Acids Research 18:4004 (1990). Any of a wide variety of expression systems may be used to produce the recombinant polypeptides. C-TAB.G5 or C-TAB.G5.1 may be produced in a prokaryotic host *(e.g.* a bacterium, such as *E. coli* or *Bacillus)* or in an eukaryotic host *(e.g.* yeast cells, mammalian cells (*e.g.* COS1, NIH3T3, or JEG3 cells), or insect cells (*e.g. Spodoptera frugiperda* (SF9) cells)). Such cells are available, for example, from the American Type Culture Collection (ATCC). The method of transformation and transfection and the choice of expression vector will depend on the host system selected. Transformation and transfection methods are described by, *e.g.,* Ausubel *et al.*, ISBN: 047132938X C-TAB.G5 or C-TAB.G5.1, particularly short fragments, may also be produced by chemical synthesis, e.g., by the methods described in Solid Phase Peptide Synthesis, 1984, 2nd ed., Stewart and Young, Eds., Pierce Chemical Co., Rockford, Ill., or by standard in vitro translation methods.

[0040] In addition to the C-TAB.G5 or C-TAB.G5.1 sequences, disclosed herein are variants thereof that are functionally active and immunogenic. The variants may have the same level of immunogenicity as C-TAB.G5 or C-TAB.G5.1. The variant may have amino acid substitutions, deletions, or insertions as compared to SEQ ID NO: 2 or SEQ ID NO: 4. Genes encoding C-TAB.G5 or C-TAB.G5.1 or variants thereof may be made using standard methods (see below; also see, *e.g.* Ausubel *et al.*, supra).

[0041] In addition to the C-TAB.G5 or C-TAB.G5.1 sequences, disclosed herein are further derivatives of C-TAB.G5 that comprise additional repeats. By way of example, a fusion protein, C-TABNCTB (SEQ ID NO: 18, encoded by SEQ ID NO: 17), comprises, like C-TAB.G5, 19 repeating units of CTA (amino acids 2272-2710), 23 repeating units of CTB (amino acids 1850-2366), and a further additional 10 repeats of CTB (amino acids 1834-2057) fused to the C-terminus of CTB. A further variant, C-TADCTB fusion protein (SEQ ID NO: 20, encoded by SEQ ID NO: 19) comprises C-TAB.G5 (19 repeats of CTA and 23 repeats of CTB) plus an additional 24 repeating units of CTB (amino acids 1834-2366) fused to the C-terminus of C-TAB.G5. A variant may also comprise additional copies of C-TAB.G5 or portions thereof. For example, C-TADCTB comprises a double portion of the repeating units of CTB present in C-TAB.G5.

[0042] Disclosed herein are methods for high level expression C-TAB.G5 or C-TAB. G5.1 in bacterial system such as *E. coli* comprising introducing a nucleic acid encoding C-TAB.G5 or C-TAB.G5.1 into a bacterial host cell and expressing

C-TAB.G5 or C-TAB.G5.1.

[0043] In addition, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein may be covalently coupled or cross-linked to adjuvants (see, *e.g.,* Cryz et al., Vaccine 13:67-71(1994); Liang et al., J. Immunology 141:1495-501 (1988) and Czerkinsky et al., Infect. Immun. 57:1072-77 (1989)).

[0044] Disclosed herein is a vaccine comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide that can protect and provide therapy against CDAD. The vaccine described herein comprises a novel antigen which can be delivered intramuscularly (IM), intradermally (ID), subcutaneously (SC), orally, nasally, buccally, or rectally routes. The vaccine may provide immune protection or induce antibodies for passive immunization.

[0045] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein provides a vaccine to immunize against CDAD. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein or variants thereof, is a combined vaccine candidate targeted to broaden the protective coverage against C. *difficile* associated diseases, such as CDAD, to a level not known or published hitherto. This concept of a single vaccine offering protection or a diminished severity of C. *difficile* associated diseases represents a unique step forward in managing public health at a global level and especially reducing the severity of epidemics (e.g. nursing homes, cruise ships).

[0046] As used herein, "toxin A protein" or "toxin B protein" refers to toxic proteins of C. *difficile* that are primarily responsible for CDAD. Toxin A and toxin B comprise multiple repeating units responsible for immunogenicity in the C-terminal binding domains.

[0047] As used herein "wild-type" or "native" refers to a full length protein comprised of a nucleic acid or amino acid sequence as would be found endogenously in a host cell.

[0048] As used herein, the terms *"Clostridium difficile* associated disease", *"Clostridium difficile* related disease", *"Clostridium difficile*-associated disease", *"Clostridium difficile* toxin-mediated disease", *"Clostridium difficile* infection", and "CDAD" refer to diseases caused, directly or indirectly, by infection with *Clostridium difficile.*

[0049] "Antigen" refers to a substance that induces a specific immune response when presented to immune cells of an organism. For example, an antigen may be a nucleic acid, a protein, a polypeptide, a peptide, a glycoprotein, a carbohydrate, a lipid, a glycolipid, a lipoprotein, a fusion protein, a phospholipid, or a conjugate of a combination thereof. An antigen may comprise a single immunogenic epitope, or a multiplicity of immunogenic epitopes recognized by a B-cell receptor (*i.e.,* antibody on the membrane of the B cell) or a T-cell receptor. Antigen may be provided as a virus-like-particle (VLP) or a whole microbe or microorganism such as, for example, a bacterium or virion. The antigen may be an inactivated or attenuated live virus. The antigen may be obtained from an extract or lysate, either from whole cells or membrane alone; or antigen may be chemically synthesized or produced by recombinant means. An antigen may be administered by itself or with an adjuvant. A single antigen molecule may have both antigen and adjuvant properties.

[0050] By "adjuvant" is meant any substance that is used to specifically or non-specifically potentiate an antigen-specific immune response, perhaps through activation of antigen presenting cells. Examples of adjuvants include an oil emulsion (e.g., complete or incomplete Freund's adjuvant), Montanide incomplete Seppic adjuvant such as ISA, oil in water emulsion adjuvants such as the Ribi adjuvant system, syntax adjuvant formulation containing muramyl dipeptide, aluminum salt adjuvant (ALUM), polycationic polymer, especially polycationic peptide, especially polyarginine or a peptide containing at least two LysLeuLys motifs, especially KLKLLLLLKLK, immunostimulatory oligodeoxynucleotide (ODN) containing non-methylated cytosine-guanine dinucleotides (CpG) in a defined base context (e.g., as described in WO 96/02555) or ODNs based on inosine and cytidine (*e.g.*, as described in WO 01/93903), or deoxynucleic acid containing deoxy-inosine and/or deoxyuridine residues (as described in WO 01/93905 and WO 02/095027), especially Oligo(dIdC)$_{13}$ (as described in WO 01/93903 and WO 01/93905), neuroactive compound, especially human growth hormone (described in WO 01/24822), or combinations thereof, a chemokine (*e.g.*, defensins 1 or 2, RANTES, MIP1-α, MIP-2, interleukin-8, or a cytokine (*e.g.*, interleukin-1β, -2, -6, -10 or -12; interferon-γ; tumor necrosis factor-α; or granulocyte-monocyte-colony stimulating factor) (reviewed in Nohria and Rubin, 1994), a muramyl dipeptide variant (*e.g.,* murabutide, threonyl-MDP or muramyl tripeptide), synthetic variants of MDP, a heat shock protein or a variant, a variant of Leishmania major LeIF (Skeiky et al., 1995, J. Exp. Med. 181: 1527-1537), non-toxic variants of bacterial ADP-ribosylating exotoxins (bAREs) including variants at the trypsin cleavage site (Dickenson and Clements, (1995) Infection and Immunity 63 (5): 1617-1623) and/or affecting ADP-ribosylation (Douce *et al.*, 1997) or chemically detoxified bAREs (toxoids), QS21, Quill A, N-acetylmuramyl-L-alanyl-D-isoglutamyl-L-alanine-2-[1,2-dipalmitoyl-s-glycero-3-(hydroxyphosphoryloxy)]ethylamide (MTP-PE) and compositions containing a metabolizable oil and an emulsifying agent. An adjuvant may be administered with an antigen or may be administered by itself, either by the same route as that of the antigen or by a different route than that of the antigen. A single adjuvant molecule may have both adjuvant and antigen properties.

[0051] By "effective amount" is meant an amount of a therapeutic agent sufficient to induce or enhance an antigen-specific immune response, for an antigen, or treat or diagnose a condition, for a drug. Such induction of an immune response may provide a treatment such as, for example, immunoprotection, desensitization, immunosuppression, modulation of autoimmune disease, potentiation of cancer immunosurveillance, or therapeutic vaccination against an established infectious disease. Treatment includes curing, amelioration, or prevention.

[0052] By "nucleic acid" is meant either a single deoxyribonucleic acid base or a ribonucleic acid or a sequence thereof

joined by phosphodiester bonds.

[0053] By "therapeutic agent" is meant any molecule capable of use in treating a disease, alleviating the symptoms of a disease, preventing a disease, or diagnosing a disease. For example, a therapeutic agent may be an antigen or a drug.

[0054] By "subject" is meant an animal. The subject may be any animal, including any vertebrate. The subject may be a domestic livestock, laboratory animal (including but not limited to, rodents such as a rat, hamster, gerbil, or mouse) or pet animal. In one embodiment, the animal may be a mammal. Examples of mammals include humans, primates, marsupials, canines, monkeys, rodents, felines, apes, whales, dolphins, cows, pigs, and horses. The subject may be in need of treatment of a disease or may be in need of a prophylactic treatment.

[0055] As used herein, the term "antibody" means an immunoglobulin molecule or a fragment of an immunoglobulin molecule having the ability to specifically bind to a particular antigen. Antibodies are well known to those of ordinary skill in the science of immunology. As used herein, the term "antibody" means not only full-length antibody molecules but also fragments of antibody molecules retaining antigen binding ability. Such fragments are also well known in the art and are regularly employed both *in vitro* and *in vivo.* In particular, as used herein, the term "antibody" means not only full-length immunoglobulin molecules but also antigen binding active fragments such as the well-known active fragments F(ab')2, Fab, Fv, and Fd.

[0056] As used herein, the term "variants" may include proteins and/or polypeptides and/or peptides that are different from a wild-type polypeptide, wherein one or more residues have been conservatively substituted with a functionally similar residue, and further which displays substantially identical functional properties of the wild-type polypeptide. Examples of conservative substitutions include substitution of one non-polar (hydrophobic) residue for another (*e.g.* isoleucine, valine, leucine or methionine) for another, substitution of one polar (hydrophilic) residue for another (*e.g.* between arginine and lysine, between glutamine and asparagine, between glycine and serine), substitution of one basic residue for another (*e.g.* lysine, arginine or histidine), or substitution of one acidic residue for another (*e.g.* aspartic acid or glutamic acid). A variant may include any polypeptide having a tertiary structure substantially identical to a polypeptide of the invention which also displays the functional properties of the polypeptides as described herein. A variant may be a mutant of a wild-type polypeptide.

[0057] As used herein "treatment" may include any type of intervention used in an attempt to alter the natural course of the individual or cell. Treatment may include, but is not limited to, administration of *e.g.,* a pharmaceutical composition, alone or in combination with other treatment modalities generally known in the art. The "treatment" may be performed prophylactically, or subsequent to the initiation of a pathologic event.

[0058] As used herein, "pharmaceutically acceptable carrier" may include any material which, when combined with an active ingredient, allows the ingredient to retain biological activity and is non-reactive with the subject's immune system. The pharmaceutically acceptable carriers and/or excipients may include buffers, stabilizers, diluents, preservatives, and solubilizers. In general, the nature of the carrier or excipients will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (e. g. powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. In addition to biologically neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

[0059] As used herein, "fusion" may refer to nucleic acids and polypeptides that comprise sequences that are not found naturally associated with each other in the order or context in which they are placed according to the present invention. A fusion nucleic acid or polypeptide does not necessarily comprise the natural sequence of the nucleic acid or polypeptide in its entirety. Fusion proteins have the two or more segments joined together through normal peptide bonds. Fusion nucleic acids have the two or more segments joined together through normal phosphodiester bonds.

Isolated polypeptides

[0060] Disclosed herein are the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides as set forth in SEQ ID NO: 2 and SEQ ID NO: 4, respectively, that comprises 19 repeating units of *C. difficile* toxin A and 23 repeating units of C. *difficile* toxin B. A homolog of C-TAB.G5, such as C-TAB.G5.1, may differ from C-TAB.G5 by 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids. The C-TAB. G5.1 polypeptide is a fusion protein containing the same C-terminal domain of toxin B as C-TAB.G5, but the C-terminal domain of toxin A derived from C. *difficile* VPI-10463 strain which is a homolog of the according C-TAB.G5 polypeptide derived from C. *difficile* 630 strain and differs by two amino acids at positions 155-156. The C-TAB.G5.1 coding sequence, as set forth in SEQ ID NO: 3, was codon optimized for improved expression within an *E. coli* host cell. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein may be effective in neutralizing the toxic effects of *C. difficile* toxin A and toxin B.

[0061] Toxin A and toxin B are encoded by the *trdA* (SEQ ID NO: 5) and *trdB* (SEQ ID NO: 7) genes, of the *C. difficile*

strain 630, respectively. Structurally, the *C. difficile* toxins comprise an ADP-glucosyl transferase domain, a cysteine protease domain, a hydrophobic region, and a receptor binding region. The C-terminal domain contains highly repetitive units (RUs) (also known as combined repetitive oligope ptides (CROPS)). The RUs may be long or short oligopeptides and may comprise 20 to 50 amino acids with a consensus YYF motif that is repeated. The RUs are grouped in clusters. As an example, toxin A, strain 630 (SEQ ID NO: 6) encoded by the wild-type *trdA* gene (SEQ ID NO: 5) contains 39 RUs. The 39 RUs are grouped into 8 clusters. Toxin B, strain 630 (SEQ ID NO: 8) encoded by the wild-type *trdB* gene (SEQ ID NO: 7) contains 24 RUs which are grouped into 5 clusters. Tables 1 and 2 below show the amino acid positions of each of the RUs in C. *difficile* toxin A and toxin B encoded by the *trdA* gene and *trdB* gene.

Table 1: Toxin A Repeating Units (ARU)

CLUSTER	REPEAT	AA START (SEQ ID NO: 6)	SEQ
1	S1	1832	GLININNSLFYFDPIEFNLVT
	S1	1853	GWQTINGKKYYFDINTGAAL
	S3	1874	SYKIINGKHFYFNNDGVMQL
	L	1894	GVFKGPDGFEYFAPANTQNNNIEGQAIVYQS
2	S1	1925	KFLTLNGKKYYFDNNSKAVT
	S2	1945	GWRIINNEKYYFNPNNAIAAV
	S3	1966	GLQVIDNNKYYFNPDTAIISK
	S4	1987	GWQTVNGSRYYFDTDTAIAFN
	S5	2008	GYKTIDGKHFYFDSDCVVKI
	L	2028	GVFSTSNGFEYFAPANTYNNNIEGQAIVYQS
3	S1	2059	KFLTLNGKKYYFDNNSKAVT
	S2	2079	GLQTIDSKKYYFNTNTAEAAT
	S3	2100	GWQTIDGKKYYFNTNTAEAAT
	S4	2121	GWQTIDGKKYYFNTNTAIAST
	S5	2142	GYTIINGKHFYFNTDGIMQI
	L	2162	GVFKGPNGFEYFAPANTDANNIEGQAILYQN
4	S1	2193	EFLTLNGKKYYFGSDSKAVT
	S2	2213	GWRIINNKKYYFNPNNAIAAI
	S3	2234	HLCTINNDKYYFSYDGILQN
	S4	2254	GYITIERNNFYFDANNESKMVT
	L	2276	GVFKGPNGFEYFAPANTHNNNIEGQAIVYQN
5	S1	2307	KFLTLNGKKYYFDNDSKAVT
	S2	2328	GWQTIDGKKYYFNLNTAEAAT
	S3	2349	GWQTIDGKKYYFNLNTAEAAT
	S4	2370	GWQTIDGKKYYFNTNTFIAST
	S5	2391	GYTSINGKHFYFNTDGIMQI
	L	2411	GVFKGPNGFEYFAPANTDANNIEGQAILYQN
6	S1	2441	KFLTLNGKKYYFGSDSKAVT
	S2	2460	GLRTIDGKKYYFNTNTAVAVT
	S3	2482	GWQTINGKKYYFNTNTSIAST
	S4	2503	GYTIISGKHFYFNTDGIMQI
	L	2523	GVFKGPDGFEYFAPANTDANNIEGQAIRYQN

(continued)

CLUSTER	REPEAT	AA START (SEQ ID NO: 6)	SEQ
7	S1	2554	RFLYLHDNIYYFGNNSKAAT
	S1	2574	GWVTIDGNRYYFEPNTAMGAN
	S3	2595	GYKTIDNKNFYFRNGLPQI
	L	2614	GVFKGSNGFEYFAPANTDANNIEGQAIRYQN
8	S1	2645	RFLHLLGKIYYFGNNSKAVT
	S2	2665	GWQTINGKVYYFMPDTAMAAAG
	S3	2687	GLFEIDGVIYFFGVDGVKAPGIYG
S: indicates a Short repeating unit L: indicates a Long repeating unit			

Table 2: Toxin B Repeating Units (BRU)

CLUSTER	REPEAT	AA START (SEQ ID NO: 8)	SEQ
1	S1	1834	GLIYINDSLYYFKPPVNNLIT
	S2	1854	GFVTVGDDKYYFNPINGGAASI
	S3	1877	GETIIDDKNYYFNQSGVLQT
	L	1897	GVFSTEDGFKYFAPANTLDENLEGEAIDFT
2	S1	1927	GKLIIDENIYYFDDNYRGAV
	S2	1947	EWKELDGEMHYFSPETGKAFK
	S3	1968	GLNQIGDYKYYSNSDGVMQK
	S4	1988	GFVNINDKTFYFDDSGVMKS
	S5	2008	GYTEIDGKHFYFAENGEMQI
	L	2028	GVFNTEDGFKYFAHHNEDLGNEEGEEISYS
3	S1	2058	GILNFNNKIYYFDD SFTAVG
	S2	2078	WKDLEDGSKYYFDEDTAEAI
	S3	2098	GLSLINDGQYYFNDDGIMQV
	S4	2119	GFVTINDKVFYFSD SGIIES
	S5	2139	GVQNIDDNYFYIDDNGIVQI
	L	2159	GVFDTSDGYKYFAPANTVNDNIYGQAVEYS
4	S1	2189	GLVRVGEDVYYFGETYTIETGWI
	S2	2213	YDMENESDKYYFNPETKKACK
	S3	2234	GINLIDDIKYYFDEKGIMRT
	S4	2254	GLISFENNNYYFNENGEMQF
	S5	2274	GYINIEDKMFYFGEDGVMQI
	L	2294	GVFNTPDGFKYFAHQNTLDENFEGESINYT
5	S1	2324	GWLDLDEKRYYFTDEYIAAT
	S2	2344	GSVIIDGEEYYFDPDTAQLVISE
S: indicates a Short repeating unit L: indicates a Long repeating unit			

[0062] Accordingly, the C-TAB.G5 and C-TAB.G5.1 isolated polypeptides comprises 19 RUs from the C-terminal domain of C. *difficile* toxin A and 23 RUs from the C-terminal domain of C. *difficile* toxin B, respectively. The C-TAB.G5 or C-TAB.G5.1 comprises toxin A amino acids 2272-2710 of SEQ ID NO: 6 fused to toxin B amino acids 1850-2366 of SEQ ID NO: 8. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide comprises the amino acid sequence as set forth in SEQ ID NO: 2 and SEQ ID NO: 4, respectively.

[0063] The respective RUs in the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may also be from variants of C. *difficile* toxin A or toxin B. These RUs in the C-TAB isolated polypeptide may also be a combination of naturally occurring or variants of *C. difficile* toxin A or toxin B.

[0064] The RUs in the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides comprise long RUs and short RUs, and the long RUs and the short RUs are arranged into a cluster. The C-TAB.G5 or C-TAB. G5.1 isolated polypeptides of the present invention comprises 4 clusters of 3 to 5 short RUs followed by one long RU of C. *difficile* toxin A and 5 clusters of 3 to 5 short RUs followed by one long RU of C. *difficile* toxin B.

[0065] The short and long RUs contain conserved motifs. The short repeating unit may comprise 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or 26 amino acids. Each short repeating unit may comprise conserved tyrosine motifs, such as YYF, FYF, YFF, FYI, or HYF. A short repeat unit may further comprise an aspartate/histidine residue prior to the tyrosine motif if the following repeating unit is a long repeating unit. The long repeating unit may comprise 27, 28, 29, 30, 31, 32, 33, 34, or 35 amino acids. Each long repeating unit may comprise a tyrosine repeat motif such as FEYF, FKYF, or YKYF.

[0066] In the present invention, the toxin A and toxin B portions of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides may be fused directly together. The toxin A and toxin B portions may be spaced apart by a linker region. A linker region may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 to 15, 20 to 30, 40, 45, or 50 amino acids. Those skilled in the art will recognize that the linker region may be adapted to alter the positioning of the toxin A and toxin B portions so that in their expressed and folded shape each toxin repeating unit in the C-TAB.G5 or C-TAB. G5.1 isolated polypeptides is positioned to optimally expose potential epitopes and to retain its immunogenicity. The RUs and the clusters in the C-TAB isolated polypeptides may also be separated by linkers. In one embodiment, the linker comprises the peptide RSMH (439-442 of SEQ ID NO: 2 or SEQ ID NO: 4).

[0067] The C-TAB isolated polypeptides described herein may have at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity or sequence similarity with SEQ ID NO: 2 or SEQ ID NO: 4. As known in the art "similarity" between two polypeptides or polynucleotides is determined by comparing the amino acid or nucleotide sequence and its conserved nucleotide or amino acid substitutes of one polynucleotide or polypeptide to the sequence of a second polynucleotide or polypeptide. Also known in the art is "identity" which means the degree of sequence relatedness between two polypeptide or two polynucleotide sequences as determined by the identity of the match between two strings of such sequences. Both identity and similarity can be readily calculated (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity and similarity between two polynucleotide or polypeptide sequences, the terms "identity" and "similarity" are well known to skilled artisans (Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipman, D., SIAM J. Applied Math. 48:1073 (1988).

[0068] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein are immunogenic. For example, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein may have at least 50%, 60%, 70%, 80%, or 90% of the immunological activity of the corresponding bacterial toxin A, and the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides may have at least 50%, 60%, 70%, 80%, or 90% of the immunological activity of the corresponding bacterial toxin B. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein may be used as vaccines for treating, preventing, or alleviating the symptoms of CDAD

[0069] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein also include variants of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, respectively. The variants may have amino acid insertions, substitutions and/or deletions that have minimal to no effect on the activity, function or shape of the isolated polypeptide. Examples of such substitutions include the substitution of one non-polar residue for another, the substitution of one polar residue for another, the substitution of one basic residue for another, or the substitution of one acidic residue for another. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide variants may further include insertions, substitutions and/or deletions of amino acids in a comparison to the amino acid sequence of the extracellular domain of native toxin A or toxin B that yield minimal effect on the activity, function and/or structure of the polypeptide. Those skilled in the art will recognize non-natural amino acids may also be used. Non-natural amino acids include, for example,

beta-alanine (beta-Ala), or other omega-amino acids, such as 3-amino propionic, 2,3-diamino propionic (2,3-diaP), 4-amino butyric and so forth, alpha-aminisobutyric acid (Aib), sarcosine (Sat), ornithine (Orn), citrulline (Cit), t-butylalanine (t-BuA), t-butylglycine (t-BuG), N-methylisoleucine (N-MeIle), phenylglycine (Phg), and cyclohexylalanine (Cha), norleucine (Nle), cysteic acid (Cya) 2-naphthylalanine (2-Nal); 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (Tic); beta-2-thienylalanine (Thi); and methionine sulfoxide (MSO).

[0070] The nucleotide sequences encoding C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein may be codon optimized to enhance expression in varying host cells. Codon optimization refers to modifying the nucleotide sequence in order to enhance protein expression in a host cell of interest by replacing one or more codons of the native sequence with codons that are more frequently used in the genes of that host cell or in the genes of the host the cell was derived from. Various species exhibit particular bias for certain codons of a particular amino acid. Disclosed herein is codon-optimized nucleotide sequence encoding the C-TAB.G5.1 isolated polypeptide for enhanced expression in *E.coli.*

[0071] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein may be prepared by any known techniques. For example, the isolated polypeptides may be expressed through genetic engineering. By way of example, the translation of recombinant DNA. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides may also be prepared synthetically. By way of example, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides may be synthesized using the solid-phase synthetic technique initially described by Merrifield (J. Am Chem. Soc. 85:2149-2154). Other polypeptide synthesis techniques may be found, for example, Kent et al. (1985) in Synthetic Peptides in Biology and Medicine, eds. Alitalo et al., Elsevier Science Publishers, 295-358.

[0072] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein may be isolated or obtained in substantially pure form. Substantially pure means that the proteins and/or polypeptides and/or peptides are essentially free of other substances with which they may be found in nature or *in vivo* systems to an extent practical and appropriate for their intended use. In particular, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides are sufficiently pure and are sufficiently free from other biological constituents of their host cells so as to be useful in, for example, generating antibodies, sequencing, or producing pharmaceutical preparations. By techniques well known in the art, substantially pure polypeptides may be produced in light of the nucleic acid and amino acid sequences disclosed herein. Because a substantially purified isolated polypeptide of the invention may be admixed with a pharmaceutically acceptable carrier in a pharmaceutical preparation, the isolated polypeptide may comprise only a certain percentage by weight of the preparation. The isolated polypeptide is nonetheless substantially pure in that it has been substantially separated from the substances with which it may be associated in living systems.

[0073] Disclosed herein are isolated C-TAB.G5 or C-TAB.G5.1 isolated polypeptides comprising additional polypeptides. The additional polypeptides may be fragments of a larger polypeptide. In one embodiment, there are one, two, three, four, or more additional polypeptides fused to the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In some embodiments, the additional polypeptides are fused toward the amino terminus of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In other embodiments, the additional polypeptides are fused toward the carboxyl terminus of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In further embodiments, the additional polypeptides flank the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In yet further embodiments, the additional polypeptides are dispersed between the toxin A portion and the toxin B portion of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides.

[0074] In some embodiments, the additional polypeptides aid in directing the secretion or subcellular localization of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. Such polypeptides are referred to as a "signal sequence." A secretory signal is described, for example U.S. Pat. 6,291,212 and U.S. Pat 5,547,871.

[0075] Secretory signal sequence encodes secretory peptides. A secretory peptide is an amino acid sequence that acts to direct the secretion of C-TAB.G5 or C-TAB.G5.1 from a cell. Secretory peptides are generally characterized by a core of hydrophobic amino acids and are typically (but not exclusively) found at the amino termini of newly synthesized proteins. The secretory peptide may be cleaved from C-TAB.G5 or C-TAB.G5.1 isolated polypeptide during secretion. Secretory peptides may contain processing sites that allow cleavage of the signal peptide from the mature protein as it passes through the secretory pathway. Processing sites may be encoded within the signal peptide or may be added to the signal peptide by, for *example, in vitro* mutagenesis. Secretory signal sequences may be required for a complex series of post-translational processing steps to allow for secretion of C-TAB.G5 or C-TAB.G5.1. The signal sequence may immediately follow the initiation codon and encodes a signal peptide at the amino-terminal end of C-TAB.G5 or C-TAB.G5.1. The signal sequence may precede the stop codon and encodes a signal peptide at the carboxy-terminal end of C-TAB.G5 or C-TAB.G5.1. In most cases, the signal sequence is cleaved off by a specific protease, called a signal peptidase. Examples of a secretory signal sequences include, but are not limited to ompA, pelB, and ST pre-pro.

[0076] In some embodiments, the additional polypeptides aid the stabilization, structure and/or the purification of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In some embodiments the additional polypeptides may comprise an epitope. In other embodiments, the additional polypeptides may comprise an affinity tag. By way of example, fusion of a polypeptide comprising an epitope and/or an affinity tag to the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may aid purification and/or identification of the polypeptide. By way of example, the polypeptide segment may be a His-tag, a

myc-tag, an S-peptide tag, a MBP tag (maltose binding protein), a GST tag (glutathione S-transferase), a FLAG tag, a thioredoxin tag, a GFP tag (green fluorescent protein), a BCCP (biotin carboxyl carrier protein), a calmodulin tag, a Strep tag, an HSV-epitope tag, a V5-epitope tag, and a CBP tag. The use of such epitopes and affinity tags is known to those skilled in the art.

[0077] In further embodiments, the additional polypeptides may provide a C-TAB.G5 or C-TAB. G5.1 isolated polypeptide comprising sites for cleavage of the polypeptide. As an example, a polypeptide may be cleaved by hydrolysis of the peptide bond. In some embodiments, the cleavage is performed by an enzyme. In some embodiments, cleavage occurs in the cell. In other embodiments, cleavage occurs through artificial manipulation and/or artificial introduction of a cleaving enzyme. By way of example, cleavage enzymes may include pepsin, trypsin, chymotrypsin, thrombin, and/or Factor Xa. Cleavage allows ease of isolating the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides from the polypeptides. Cleavage may further allow for the separation of the toxin A portion from the toxin B portion. Cleavage may also allow isolation of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide fused to polypeptides from other polypeptides, such as through cleavage of an epitope utilized to purify the expressed protein.

[0078] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides may further possess additional structural modifications not shared with the same organically synthesized peptide, such as adenylation, carboxylation, glycosylation, hydroxylation, methylation, phosphorylation or myristylation. These added structural modifications may be further be selected or preferred by the appropriate choice of recombinant expression system. On the other hand, fusion polypeptides may have its sequence extended by the principles and practice of organic synthesis.

[0079] Disclosed herein are nucleic acids encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides comprising a polypeptide portion obtained from C. *difficile* toxin A and a polypeptide portion obtained from C. *difficile* toxin B. Nucleic acids may include single or double stranded forms of deoxyribonucleotides or ribonucleotides or polymers thereof. Disclosed herein are ribonucleic acids encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. Disclosed herein are for nucleic acids that hybridize under stringent conditions to a nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide and the complement thereof. Stringent conditions refer to the degree of homology between a probe and a filter-bound nucleic acid; the higher the stringency, the higher percent homology between the probe and filter bound nucleic acid. The temperature for a stringent wash may be determined based on the Tm of the nucleic acid (based on G/C content). Stringent conditions may further be affected by the concentration of salt in a buffer, such as standard sodium citrate (SSC). The present invention provides for nucleic acids having about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence similarity or sequence identity with SEQ ID NO: 1.

[0080] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may further comprise a linker region, for instance a linker less than about 50, 40, 30, 20, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid residues. The linker can be covalently linked to and between the polypeptide portion derived from toxin A or portion thereof and the polypeptide portion derived from toxin B.

[0081] Disclosed herein are nucleic acids encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides that are degenerate to SEQ ID NO: 1 or SEQ ID NO: 3, respectively. The degeneracy of the genetic code permits variations of the nucleotide sequence of a toxin A protein, a toxin B protein and/or isolated polypeptide of interest, while still producing a polypeptide having the identical amino acid sequence as the polypeptide encoded by the native DNA sequence. The procedure, known as "codon optimization" (described in U.S. Patent 5,547,871) provides one with a means of designing such an altered DNA sequence. The design of codon optimized genes should take into account a variety of factors, including the frequency of codon usage in an organism, nearest neighbor frequencies, RNA stability, the potential for secondary structure formation, the route of synthesis and the intended future DNA manipulations of that gene. In particular, available methods may be used to alter the codons encoding a given isolated polypeptide with those most readily recognized by yeast when yeast expression systems are used, or by insect cells when the insect cell expression system is used. The degeneracy of the genetic code also permits the same amino acid sequence to be encoded and translated in many different ways. For example, leucine, serine and arginine are each encoded by six different codons, while valine, proline, threonine, alanine and glycine are each encoded by four different codons. However, the frequency of use of such synonymous codons varies from genome to genome among eukaryotes and prokaryotes. For example, synonymous codon-choice patterns among mammals are very similar, while evolutionarily distant organisms such as yeast (such as *S. cerevisiae*), bacteria (such as *E. coli)* and insects (such as *D. melanogaster*) reveal a clearly different pattern of genomic codon use frequencies (Grantham, R., et al., Nucl. Acid Res., 8, 49-62 (1980); Grantham, R., et al., Nucl. Acid Res., 9, 43-74 (1981); Maroyama, T., et al., Nucl. Acid Res., 14, 151-197 (1986); Aota, S., et al., Nucl. Acid Res., 16, 315-402 (1988); Wada, K., et al., Nucl. Acid Res., 19 Supp., 1981-1985 (1991); Kurland, C. G., FEBS Lett., 285, 165-169 (1991)). These differences in codon-choice patterns appear to contribute to the overall expression levels of individual genes by modulating peptide elongation rates. (Kurland, C. G., FEBS Lett., 285, 165-169 (1991); Pedersen, S., EMBO J., 3, 2895-2898 (1984); Sorensen, M. A., J. Mol. Biol., 207, 365-377 (1989); Randall, L. L., et al., Eur. J. Biochem., 107, 375-379 (1980); Curran, J. F., and Yarus, M., J. Mol. Biol., 209, 65-77 (1989); Varenne, S., et al., J. Mol. Biol., 180, 549-576 (1984), Varenne, S., et al., J. Mol, Biol., 180, 549-576 (1984); Garel, J.-P., J. Theor. Biol., 43, 211-225 (1974); Ikemura, T., J. Mol. Biol., 146, 1-21 (1981); Ikemura, T., J. Mol. Biol., 151, 389-409 (1981)).

[0082] The preferred codon usage frequencies for a synthetic gene should reflect the codon usages of nuclear genes derived from the exact (or as closely related as possible) genome of the cell/organism that is intended to be used for recombinant protein expression.

[0083] Preferred methods to determine identity are designed to give the largest match between the two sequences tested. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, et al., Nucl. Acid Res. 12(1):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, et al., J. Mol. Biol. 215:403 (1990)). The degree of similarity or identity referred to above is determined as the degree of identity between the two sequences, often indicating a derivation of the first sequence from the second. The degree of identity between two nucleic acids may be determined by means of computer programs known in the art such as GAP provided in the GCG program package (Needleman and Wunsch J. Mol. Biol. 48:443-453 (1970)). For purposes of determining the degree of identity between two nucleic acids for the present invention, GAP is used with the following settings: GAP creation penalty of 5.0 and GAP extension penalty of 0.3.

[0084] Disclosed herein is a vector comprising a nucleic acid encoding for the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. A vector may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA, although RNA vectors are also available. Vectors include, but are not limited to, plasmids and phagemids. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase.

[0085] Vectors may further contain a promoter sequence. A promoter may include an untranslated nucleic acid usually located upstream of the coding region that contains the site for initiating transcription of the nucleic acid. The promoter region may also include other elements that act as regulators of gene expression. In further embodiments of the invention, the expression vector contains an additional region to aid in selection of cells that have the expression vector incorporated. The promoter sequence is often bounded (inclusively) at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes.

[0086] Vectors may further contain one or more marker sequences suitable for use in the identification and selection of cells which have been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β-galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques. Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

[0087] An expression vector is one into which a desired nucleic acid may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Expression refers to the transcription and/or translation of an endogenous gene, transgene or coding region in a cell.

[0088] A coding sequence and regulatory sequences are operably joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

[0089] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein may be produced by expressing the encoding nucleic acid in host cells. The nucleic acid may be transformed or transfected into host cells. Accordingly, some aspects described herein include the transformation and/or transfection of nucleic acid encoding the C-TAB.G5 or C-TAB. G5.1 isolated polypeptides. Transformation is the introduction of exogenous or heterologous nucleic acid to the interior of a prokaryotic cell. Transfection is the introduction of exogenous or heterologous nucleic acid to the interior of a eukaryotic cell. The transforming or transfecting nucleic acid may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes, for example, the transforming nucleic acid may

be maintained on an episomal element such as a plasmid or viral vector. With respect to eukaryotic cells, a stably transfected cell is one in which the transfecting nucleic acid has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transfected nucleic acid.

[0090] Higher eukaryotic cell cultures may be used to express the proteins of the present invention, whether from vertebrate or invertebrate cells, including insects, and the procedures of propagation thereof are known (see, for example, Kruse et al. (1973) Tissue Culture, Academic Press).

[0091] Host cells and vectors for replicating the nucleic acids and for expressing the encoded C-TAB.G5 or C-TAB.G5.1 isolated polypeptides are also provided. Any vectors or host cells may be used, whether prokaryotic or eukaryotic. Many vectors and host cells are known in the art for such purposes. It is well within the skill of the art to select an appropriate set for the desired application.

[0092] DNA sequences encoding toxin A and toxin B, or portions thereof may be cloned from a variety of genomic or cDNA libraries derived from C. *difficile* and other known toxin A and toxin B expressing prokaryotes known in the art. The techniques for isolating such DNA sequences using probe-based methods are conventional techniques and are well known to those skilled in the art. Probes for isolating such DNA sequences may be based on published DNA or protein sequences. Alternatively, the polymerase chain reaction (PCR) method disclosed by Mullis et al. (U.S. Pat. No. 4,683,195) and Mullis (U.S. Pat. No. 4,683,202) may be used. The choice of library and selection of probes for the isolation of such DNA sequences is within the level of ordinary skill in the art.

[0093] Suitable host cells maybe derived from prokaryotes or eukaryotes. Suitable prokaryote hosts include: *Pseudomonas* such as *P. aeruginosa, Escherichia coli*, *Staphylococcus* such as *S. aureus* and *S. epidermidis, Serratia marcescens*, *Bacillus* such as *B. subtillis* and *B. megaterium, Clostridium sporogenes*, *Enterococcus faecalis*, *Micrococcus* such as *M. luteus* and *M. roseus*, and *Proteus vulgaris.* Suitable host cells for expressing the polypeptides of the present invention in higher eukaryotes include: yeasts such as *Saccharomyces* (*e.g. S. cerevisiae*); 293 (human embryonic kidney) (ATCC CRL-1573); 293F (Invitrogen, Carlsbad CA); 293T and variant 293T/17 (293tsA1609neo and variant ATCC CRL-11268) (human embryonic kidney transformed by SV40 T antigen); COS-7 (monkey kidney CVI line transformed by SV40)(ATCC CRL1651); BHK (baby hamster kidney cells) (ATCC CRL10); CHO (Chinese hamster ovary cells); mouse Sertoli cells; CVI (monkey kidney cells) (ATCC CCL70); VERO76 (African green monkey kidney cells) (ATCC CRL1587); HeLa (human cervical carcinoma cells) (ATCC CCL2); MDCK (canine kidney cells) (ATCC CCL34); BRL3A (buffalo rat liver cells) (ATCC CRL1442); W138 (human lung cells) (ATCC CCL75); HepG2 (human liver cells) (HB8065); and MMT 060652 (mouse mammary tumor) (ATCC CCL51).

[0094] Disclosed herein are nucleic acids encoding an isolated polypeptide comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides and additional polypeptides. Vectors useful for constructing eukaryotic expression systems for the production of fusion polypeptides comprise nucleic acid encoding the isolated polypeptide operatively linked to an appropriate transcriptional activation sequence, such as a promoter and/or operator. Other typical features may include appropriate ribosome binding sites, termination codons, enhancers, terminators, or replicon elements. These additional features can be inserted into the vector at the appropriate site or sites by conventional splicing techniques such as restriction endonuclease digestion and ligation.

[0095] In some embodiments, additional nucleic acids may be fused to the nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. The fused nucleic acid may encode polypeptides that may aid in purification and/or immunogenicity and/or stability without shifting the codon reading frame of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. The fused nucleic acids may encode a secretory sequence, that may or may not be cleaved from the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. The fused nucleic acids may not elongate the expressed polypeptide significantly. The fused nucleic acids may encode for less than sixty extra amino acids to the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In some embodiments, the fused nucleic acids follow after the nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In other embodiments, the fused nucleic acids precede the nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In other embodiments, the fused nucleic acids flank the nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides.

[0096] In some embodiments, the fused nucleic acids may encode for a polypeptide to aid purification of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. In some embodiments the fused nucleic acid will encode for an epitope and/or an affinity tag. Examples of polypeptides that aid purification include, but are not limited to, a His-tag, a myc-tag, an S-peptide tag, a MBP tag, a GST tag, a FLAG tag, a thioredoxin tag, a GFP tag, a BCCP, a calmodulin tag, a Strep tag, an HSV-epitope tag, a V5-epitope tag, and a CBP tag. In other embodiments, the fused nucleic acid may encode for a C-TAB.G5 or C-TAB.G5.1 isolated polypeptide that has a site directed for, or prone to, cleavage. In one embodiment, the fused nucleic acid may encode for polypeptides comprising sites of enzymatic cleavage. In further embodiments, the enzymatic cleavage may aid in isolating the C-TAB.G5 or C-TAB.G5.1isolated polypeptides, as well as other fused polypeptide segments, from yet other polypeptides. By way of example, an intermediary nucleic acid that encodes for an enzymatic cleavage site placed between nucleic acids that encode for C-TAB.G5 or C-TAB.G5.1 isolated polypeptide and an epitope may allow for later separation of the expressed C-TAB.G5 or C-TAB.G5.1 isolated polypeptides and the

epitope. Such sites may also be present between the toxin A portion and the toxin B portion.

[0097] Disclosed herein are expression systems designed to assist in expressing and providing the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. The expression system may comprise a host cell transformed or transfected with a nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. The host cell may be a prokaryote. The prokaryote may be *E. coli.* The host cell may be an eukaryotic cell.

[0098] The expression system may further comprise agents to aid in selection of host cells successfully transformed or transfected with a nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides. For example, the nucleic acid encoding the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide may further express a gene to assist the host cell in resistance to antibiotics, such as genes to resist kanamycin or gentamycin or ampicillin or penicillin. Such resistant genes will allow for selection of host cells that have properly incorporated the nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide, as is known to those skilled in the art.

[0099] Another aspect described herein is directed to the generation of antibodies. Examples of antibodies described herein, include, but are not limited to, antibodies produced by immunizing a subject with the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. Antibodies generated by immunizing with the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may bind specifically to toxin A or toxin B, or they may cross react with the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. The antibodies produced by the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein may be characterized using methods well known in the art.

[0100] The antibodies produced by using the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide of the present invention can encompass monoclonal antibodies, polyclonal antibodies, antibody fragments (e.g., Fab, Fab', F(ab')2, Fv, Fc, etc.), chimeric antibodies, bispecific antibodies, heavy chain only antibodies, heteroconjugate antibodies, single chain (ScFv), single domain antibodies, variants thereof, isolated polypeptides comprising an antibody portion, humanized antibodies, and any other modified configuration of the immunoglobulin molecule that comprises an antigen recognition site of the required specificity, including glycosylation variants of antibodies, amino acid sequence variants of antibodies, and covalently modified antibodies. Preferred antibodies are derived from murine, rat, human, rabbit, canine, porcine, dromedary, camel, llama, feline, primate, or any other origin (including chimeric, fragment and/or humanized antibodies).

[0101] In other embodiments, the antibodies produced by immunizing with the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide are then humanized by methods known in the art. A humanized antibody is an immunoglobulin molecule that contains minimal sequence derived from non-human immunoglobulin. In yet other embodiments, fully human antibodies are obtained by using commercially available mice that have been engineered to express specific human immunoglobulin proteins. In other embodiments, the antibodies are chimeric. A chimeric antibody is an antibody that combines characteristics from two different antibodies. Methods of preparing chimeric antibodies are known in the art.

[0102] In other embodiments, the nucleotide sequence that encodes the antibodies is obtained and then cloned into a vector for expression or propagation. In another embodiment, antibodies are made recombinantly and expressed using methods known in the art. By way of example, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be used as an antigen for the purposes of isolating recombinant antibodies by these techniques. Antibodies can be made recombinantly by using the gene sequence to express the antibody recombinantly in host cells. Methods for making variants of antibodies and recombinant antibodies are known in the art.

[0103] In other embodiments, the antibodies are bound to a carrier by conventional methods in the art, for use in, for example, isolating or purifying native toxin A or toxin B or detecting native toxin A or toxin B or C. *difficile* in a biological sample or specimen.

Compositions and formulations

[0104] Disclosed herein are compositions comprising C-TAB.G5 or C-TAB. G5.1 isolated polypeptides. The compositions may be pharmaceutical compositions comprising the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide and a pharmaceutically acceptable carrier. The compositions used in the methods described herein generally comprise, by way of example and not limitation, and effective amount of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide (e.g., an amount sufficient to induce an immune response) described herein or an antibody against the C-TAB.G5 or C-TAB. G5.1 isolated polypeptides (e.g., an amount of a neutralizing antibody sufficient to mitigate infection, alleviate a symptom of infection and/or prevent infection). The pharmaceutical composition may further comprise pharmaceutically acceptable carriers, excipients, or stabilizers known in the art (see generally Remington, (2005) The Science and Practice of Pharmacy, Lippincott, Williams and Wilkins).

[0105] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein may be used for methods for immunizing or treating humans and/or animals with the CDAD. Therefore, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptides may be used within a pharmaceutical composition. The pharmaceutical composition described herein may further encompass pharmaceutically acceptable carriers and/or excipients. The pharmaceutically acceptable carriers and/or excipients useful in this invention are conventional and may include buffers, stabilizers, diluents, preservatives, and solubilizers. Remington's Pharmaceutical Sciences, by E. W. Martin, Mack Publishing Co., Easton, PA, 15th Edition (1975), describes

compositions and formulations suitable for pharmaceutical delivery of the polypeptides herein disclosed. In general, the nature of the carrier or excipients will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (e. g. powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. In addition to biologically neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

[0106] In one embodiment the pharmaceutical composition may further comprise an immunostimulatory substance, such as an adjuvant. The adjuvant can be selected based on the method of administration and may include mineral oil-based adjuvants such as Freund's complete and incomplete adjuvant, Montanide incomplete Seppic adjuvant such as ISA, oil in water emulsion adjuvants such as the Ribi adjuvant system, syntax adjuvant formulation containing muramyl dipeptide, aluminum hydroxide or aluminum salt adjuvant (alum), polycationic polymer, especially polycationic peptide, especially polyarginine or a peptide containing at least two LysLeuLys motifs, especially KLKLLLLLKLK, immunostimulatory oligodeoxynucleotide (ODN) containing non-methylated cytosine-guanine dinucleotides (CpG) in a defined base context *(e.g.* as described in WO 96/02555) or ODNs based on inosine and cytidine *(e.g.* as described in WO 01/93903), or deoxynucleic acid containing deoxy-inosine and/or deoxyuridine residues (as described in WO 01/93905 and WO 02/095027), especially Oligo(dIdC)$_{13}$ (as described in WO 01/93903 and WO 01/93905), neuroactive compound, especially human growth hormone (described in WO 01/24822), or combinations thereof. Such combinations are according to the ones e.g. described in WO 01/93905, WO 02/32451, WO 01/54720, WO 01/93903, WO 02/13857, WO 02/095027 and WO 03/047602. Preferably, the adjuvant is aluminum hydroxide adjuvant.

[0107] Acceptable carriers, excipients, or stabilizers are nontoxic to recipients at the dosages and concentrations that are administered. Carriers, excipients or stabilizers may further comprise buffers. Examples of excipients include, but are not limited to, carbohydrates (such as monosaccharide and disaccharide), sugars (such as sucrose, mannitol, and sorbitol), phosphate, citrate, antioxidants (such as ascorbic acid and methionine), preservatives (such as phenol, butanol, benzanol; alkyl parabens, catechol, octadecyldimethylbenzyl ammonium chloride, hexamethoniuni chloride, resorcinol, cyclohexanol, 3-pentanol, benzalkonium chloride, benzethonium chloride, and m-cresol), low molecular weight polypeptides, proteins (such as serum albumin or immunoglobulins), hydrophilic polymers amino acids, chelating agents (such as EDTA), salt-forming counter-ions, metal complexes (such as Zn-protein complexes), and nonionic surfactants (such as TWEEN™ and polyethylene glycol).

[0108] The pharmaceutical composition described herein may further comprise additional agents that serve to enhance and/or complement the desired effect. By way of example, to enhance the immunogenicity the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide of the invention being administered as a subunit vaccine, the pharmaceutical composition may further comprise an adjuvant.

[0109] An example of a pharmaceutical composition may be an immunogenic composition. Disclosed herein are immunogenic compositions comprising the C-TAB.G5 or C-TAB. G5.1 isolated polypeptides. The immunogenic composition may further include a pharmaceutically acceptable carrier or other carriers and/or excipients in a formulation suitable for injection in a mammal. An immunogenic composition is any composition of material that elicits an immune response in a mammalian host when the immunogenic composition is injected or otherwise introduced. The immune response may be humoral, cellular, or both. A booster effect refers to an increased immune response to an immunogenic composition upon subsequent exposure of the mammalian host to the same immunogenic composition. A humoral response results in the production of antibodies by the mammalian host upon exposure to the immunogenic composition.

[0110] The immunogenic compositions described herein elicit an immune response in a mammalian host, including humans and other animals. The immune response may be either a cellular dependent response or an antibody dependent response or both; and further the response may provide immunological memory or a booster effect or both in the mammalian host. These immunogenic compositions are useful as vaccines and may provide a protective response by the mammalian subject or host to infection by strains of C. *difficile.*

[0111] Disclosed herein are methods for producing an immunogenic composition by constructing the nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide and expressing C-TAB.G5 or C-TAB.G5.1 isolated polypeptide component in a microbial host; recovering the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide from a culture of the host; conjugating the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide to a second protein component, and recovering the conjugated protein and polysaccharide component. The nucleic acid encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be maintained throughout the growth of the host by constant and stable selective pressure. Maintenance of the expression vector may be conferred by incorporation in the expression vector of a genetic sequence that encodes a selective genotype, the expression of which in the microbial host cell results in a selective phenotype. A selective genotype sequence may also include a gene complementing a conditional lethal mutation. Other genetic sequences may be incorporated in the expression vector, such as other drug resistance genes or genes that complement lethal

mutations. Microbial hosts may include: Gram positive bacteria; Gram negative bacteria, such as *E. coli*; yeasts; filamentous fungi; mammalian cells; insect cells; or plant cells.

[0112] The methods described herein also provide for a level of expression of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide in the host at a level greater than about 50 mg/liter of the culture, a level greater than about 100 mg/liter, a level greater than about 500 mg/liter, or a level greater than about 1 g/liter. The protein may be recovered by any number of methods known to those in the art for the isolation and recovery of proteins, such as by ammonium sulfate precipitation followed by ion exchange chromatography.

[0113] Disclosed herein are methods for preparing the immunogenic composition that provides that the protein component is conjugated to a second protein component by one of a number of means known to those in the art, such as an amidization reaction.

[0114] Disclosed herein are formulations comprising the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide for treating and preventing CDAD. In one embodiment, the formulation may include the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein, an adjuvant, and a pharmaceutically acceptable carrier. In another embodiment, the formulation includes the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein, or consists essentially of one or more C-TAB.G5 or C-TAB.G5.1 isolated polypeptides described herein. The formulation may comprise the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein and an adjuvant. The formulation may further include an additional antigen or a drug. Moreover, the formulation may include one or more drugs and may in addition to the isolated polypeptide and/or adjuvant include one or more drugs.

[0115] The formulation comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be in liquid or dry form. A dry formulation may be easily stored and transported. Dry formulations break the cold chain required from the vaccine's place of manufacture to the locale where vaccination occurs. Alternatively, the dry, active ingredient of the formulation per se may be an improvement by providing a solid particulate form that is taken up and processed by antigen presenting cells. These possible mechanisms are discussed not to limit the scope of the invention or its equivalents, but to provide insight into the operation of the invention and to guide the use of this formulation in immunization and vaccination.

[0116] Dry formulations of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be provided in various forms: for example, fine or granulated powders, lyophilized powder, uniform films, pellets, and tablets. It may be air dried, dried with elevated temperature, lyophilized, freeze or spray dried, coated or sprayed on a solid substrate and then dried, dusted on a solid substrate, quickly frozen and then slowly dried under vacuum, or combinations thereof. If different molecules are active ingredients of the formulation, they may be mixed in solution and then dried, or mixed in dry form only.

[0117] Formulations comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide in liquid or solid form, such as a dry form, may be applied with one or more adjuvants at the same or separate sites or simultaneously or in frequent, repeated applications. The formulation may include other antigens such that administration of the formulation induces an immune response to multiple antigens. In such a case, the other antigens may have different chemical structures so as to induce an immune response specific for different antigens. At least one antigen and/or adjuvant may be maintained in dry form prior to administration. Subsequent release of liquid from a reservoir or entry of liquid into a reservoir containing the dry ingredient of the formulation will at least partially dissolve that ingredient.

[0118] Solids (e.g., particles of nanometer or micrometer dimensions) may also be incorporated in the formulation. Solid forms (e.g., nanoparticles or microparticles) may aid in dispersion or solubilization of active ingredients; provide a point of attachment for adjuvant, C-TAB.G5 or C-TAB. G5.1 isolated polypeptide, or both to a substrate that can be opsonized by antigen presenting cells, or combinations thereof. Prolonged release of the formulation from a porous solid formed as a sheet, rod, or bead acts as a depot.

[0119] At least one ingredient or component of the formulation (*i.e.,* C-TAB.G5 or C-TAB. G5.1 isolated polypeptide, adjuvant, or drug) may be provided in dry form prior to administration of the formulation. This formulation may also be used in conjunction with conventional enteral, mucosal, or parenteral immunization techniques.

[0120] The formulation comprising the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide may be manufactured under aseptic conditions acceptable to appropriate regulatory agencies (*e.g.*, Food and Drug Administration, EMEA for biologicals and vaccines. Optionally, components such as desiccants, excipients, stabilizers, humectants, preservatives, or combinations thereof may be included in the formulation even though they are immunologically inactive. They may, however, have other desirable properties or characteristics.

[0121] Processes for manufacturing a pharmaceutical formulation are well known. The components of the formulation may be combined with a pharmaceutically-acceptable carrier or vehicle, as well as any combination of optional additives (*e.g.*, diluents, binders, excipients, stabilizers, desiccants, preservatives, colorings). The use of solid carriers, and the addition of excipients to assist in solubilization of dry components or stabilizers of immunogenic or adjuvant activity, are preferred embodiments. See, generally, Ullmann's Encyclopedia of Industrial Chemistry, 6th Ed. (electronic edition, 2003); Remington's Pharmaceutical Sciences, 22nd (Gennaro, 2005, Mack Publishing); Pharmaceutical Dosage Forms, 2nd Ed. (various editors, 1989-1998, Marcel Dekker); and Pharmaceutical Dosage Forms and Drug Delivery Systems (Ansel et al., 2005, Williams & Wilkins).

[0122] Good manufacturing practices are known in the pharmaceutical industry and regulated by government agencies

(*e.g.*, Food and Drug Administration, EMEA. Sterile liquid formulations may be prepared by dissolving an intended component of the formulation in a sufficient amount of an appropriate solvent, followed by sterilization by filtration to remove contaminating microbes. Generally, dispersions are prepared by incorporating the various sterilized components of the formulation into a sterile vehicle which contains the basic dispersion medium. For production of solid forms that are required to be sterile, vacuum drying or freeze drying can be used.

[0123] In general, solid dosage forms (*e.g.*, powders, granules, pellets, tablets) can be made from at least one active ingredient or component of the formulation.

[0124] Suitable tableting procedures are known. The formulation may also be produced by encapsulating solid forms of at least one active ingredient, or keeping them separate from liquids in compartments or chambers. The size of each dose and the interval of dosing to the subject may be used to determine a suitable size and shape of the tablet, capsule, compartment, or chamber.

[0125] Formulations will contain an effective amount of the active ingredients (*e.g.*, drug, antigen and adjuvant) together with carrier or suitable amounts of vehicle in order to provide pharmaceutically-acceptable compositions suitable for administration to a human or animal.

[0126] The relative amounts of active ingredients, such as amounts of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide, within a dose and the dosing schedule may be adjusted appropriately for efficacious administration to a subject (*e.g.*, animal or human). This adjustment may also depend on the subject's particular disease or condition, and whether treatment or prophylaxis is intended. To simplify administration of the formulation to the subject, each unit dose contains the active ingredients in predetermined amounts for a single round of immunization.

[0127] There are numerous causes of polypeptide instability or degradation, including hydrolysis and denaturation. In the case of denaturation, the conformation or three-dimensional structure of the protein is disturbed and the protein unfolds from its usual globular structure. Rather than refolding to its natural conformation, hydrophobic interaction may cause clumping of molecules together (*i.e.,* aggregation) or refolding to an unnatural conformation. Either of these results may entail diminution or loss of immunogenic or adjuvant activity. Stabilizers may be added to lessen or prevent such problems.

[0128] The formulation, or any intermediate in its production, may be pretreated with protective agents (*i.e.,* cryoprotectants and dry stabilizers) and then subjected to cooling rates and final temperatures that minimize ice crystal formation. By proper selection of cryoprotective agents and use of pre-selected drying parameters, almost any formulation might be cryoprepared for a suitable desired end use.

[0129] It should be understood in the following discussion of optional additives like excipients, stabilizers, desiccants, and preservatives are described by their function. Thus, a particular chemical may act as some combination of recipient, stabilizer, desiccant, and/or preservative. Such chemical would be immunologically-inactive because it does not directly induce an immune response, but it increases the response by enhancing immunological activity of the antigen or adjuvant: for example, by reducing modification of the antigen or adjuvant, or denaturation during drying and dissolving cycles.

[0130] Stabilizers include cyclodextrin and variants thereof (see U.S. Pat. No. 5,730,969). Suitable preservatives such as sucrose, mannitol, sorbitol, trehalose, dextran, and glycerin can also be added to stabilize the final formulation (Howell and Miller, 1983). A stabilizer selected from nonionic surfactants, D-glucose, D-galactose, D-xylose, D-glucuronic acid, salts of D-glucuronic acid, trehalose, dextrans, hydroxyethyl starches, and mixtures thereof may be added to the formulation. Addition of an alkali metal salt or magnesium chloride may stabilize the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide, optionally including serum albumin and freeze-drying to further enhance stability. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may also be stabilized by contacting it with a saccharide selected from the group consisting of dextran, chondroitin sulfuric acid, starch, glycogen, insulin, dextrin, and alginic acid salt. Other sugars that can be added include monosaccharides, disaccharides, sugar alcohols, and mixtures thereof *(e.g.,* glucose, mannose, galactose, fructose, sucrose, maltose, lactose, mannitol, xylitol). Polyols may stabilize a polypeptide, and are water-miscible or water-soluble. Suitable polyols may be polyhydroxy alcohols, monosaccharides and disaccharides including mannitol, glycerol, ethylene glycol, propylene glycol, trimethyl glycol, vinyl pyrrolidone, glucose, fructose, arabinose, mannose, maltose, sucrose, and polymers thereof. Various excipients may also stabilize polypeptides, including serum albumin, amino acids, heparin, fatty acids and phospholipids, surfactants, metals, polyols, reducing agents, metal chelating agents, polyvinyl pyrrolidone, hydrolyzed gelatin, and ammonium sulfate.

[0131] As an example, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide formulation can be stabilized in sucrose, trehalose, poly(lactic acid) (PLA) and poly(lactide-co-glycolide) (PLGA) microspheres by suitable choice of excipient or stabilizer (Sanchez *et al.,* 1999). Sucrose, or trehalose may be advantageously used as an additive because it is a non-reducing saccharide, and therefore does not cause aminocarbonyl reactions with substances bearing amino groups such as proteins. Sucrose or trehalose may be combined with other stabilizers such as saccharides.

[0132] Additionally, the formulation comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may include therapeutic agents, such as e.g. anesthetics, analgesics, anti-inflammatories, steroids, antibiotics, antiarthritics, anorectics, antihistamines, and antineoplastics. Examples of such therapeutic agents include lidocaine and nonsteroidal anti-inflammatory drugs (NSAID). In another embodiment, the therapeutic agents are antigens and adjuvants. In still another

embodiment, the formulation comprising antigen and/or adjuvant may be applied separately but along with other therapeutic agents, such e.g anesthetics, analgesics, anti-inflammatories, steroids, antibiotics, antiarthritics, anorectics, antihistamines, and antineoplastics. In a preferred embodiment, the antibiotics are fidaxomicin, metronidazole or vancomycin.

[0133] The formulation comprising the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide may be delivered via various routes of administration such as e.g. intramuscularly.

[0134] Polymers may be added to the formulation and may act as an excipient, stabilizer, and/or preservative of an active ingredient as well as reducing the concentration of the active ingredient that saturates a solution used to dissolve the dry form of the active ingredient. Such reduction occurs because the polymer reduces the effective volume of the solution by filling the "empty" space. Thus, quantities of antigen/adjuvant can be conserved without reducing the amount of saturated solution. An important thermodynamic consideration is that an active ingredient in the saturated solution will be "driven" into regions of lower concentration. In solution, polymers can also stabilize and/or preserve the antigen/adjuvant-activity of solubilized ingredients of the formulation. Such polymers include ethylene or propylene glycol, vinyl pyrrolidone, and 0-cyclodextrin polymers and copolymers.

[0135] A single or unit dose of the formulation comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide suitable for administration is provided. The amount of adjuvant and/or C-TAB.G5 or C-TAB. G5.1 isolated polypeptide in the unit dose may be anywhere in a broad range from about 0.001 μg to about 10 mg. This range may be from about 0.1 ug to about 1 mg; a narrower range is from about 5 μg to about 500 μg. Other suitable ranges are between about 20 μg to about 200 μg, such as e.g. about 20 μg, about 75 μg or about 200 μg.. A preferred dose for a C-TAB.G5 or C-TAB.G5.1 isolated polypeptide is from about 20 μg or 200 μg or less. The ratio between C-TAB.G5 or C-TAB.G5.1 isolated polypeptide and adjuvant may be about 1:1 or about 1:1.25, but higher ratios may also be used (*e.g.,* about 1:10 or less), or lower ratios of C-TAB isolated polypeptide to adjuvant may also be used (*e.g.,* about 10:1 or more).

[0136] The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be used as an antigen and may be presented to immune cells, and an antigen-specific immune response is induced. This may occur before, during, or after infection by a pathogen, such as *C. difficile.* Only C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be required, but no additional adjuvant, if the immunogenicity of the formulation is sufficient to not require adjuvant activity. The formulation may include an additional antigen such that application of the formulation induces an immune response against multiple antigens (*i.e.,* multivalent). Antigen-specific lymphocytes may participate in the immune response and, in the case of participation by B lymphocytes, antigen-specific antibodies may be part of the immune response. The formulations described above may include desiccants, excipients, humectants, stabilizers, and preservatives known in the art.

[0137] The formulation comprising the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide described herein may be used to treat a subject (*e.g.,* a human or animal in need of treatment such as prevention of disease, protection from effects of infection, reducing or alleviating the symptoms of a disease, such as CDAD, or combinations thereof). E.g. the formulation comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide described herein may be used to treat a subject at risk of CDAD, such as e.g. a subject with the following profile: i) a subject with a weaker immune system such as e.g. an elderly subject (e.g. a subject above 65 years of age) or a subject below 2 years of age; ii) an immunocompromised subject such as e.g. a subject with AIDS; iii) a subject taking or planning to take immunosuppressing drugs; iv) a subject with planned hospitalization or a subject that is in hospital; v) a subject in or expected to go to an intensive care unit (ICU); vi) a subject that is undergoing or is planning to undergo gastrointestinal surgery; vii) a subject that is in or planning to go to a long-term care such as a nursing home; viii) a subject with co-morbidities requiring frequent and/or prolonged antibiotic use; ix) a subject that is a subject with two or more of the above mentioned profiles, such as e.g. an elderly subject that is planning to undergo a gastrointestinal surgery; x) a subject with inflammatory bowel disease; and/or xi) a subject with recurrent CDAD such as e.g. a subject having experienced one or more episodes of CDAD.

[0138] The treatment may vaccinate the subject against infection by the pathogen or against its pathogenic effects such as those caused by toxin secretion. The formulation may be used therapeutically to treat existing disease, protectively to prevent disease, to reduce the severity and/or duration of disease, to ameliorate symptoms of disease, or combinations thereof.

[0139] The formulations comprising C-TAB.G5 or C-TAB.G5.1 isolated polypeptides may be delivered by various routes of administration including but not limited to oral, subcutaneous, intradermal, intravenous, intra-arterial, intramuscular, intracardial, intraspinal, intrathoracical, intraperitoneal, intraventricular, and/or sublingual routes.

[0140] The formulation may also comprise one or more adjuvants or combinations of adjuvants. Usually, the adjuvant and the formulation are mixed prior to presentation of the antigen but, alternatively, they may be separately presented within a short interval of time.

[0141] Adjuvants include, for example, an oil emulsion (*e.g.,* complete or incomplete Freund's adjuvant), Montanide incomplete Seppic adjuvant such as ISA, oil in water emulsion adjuvants such as the Ribi adjuvant system, syntax adjuvant formulation containing muramyl dipeptide, aluminum hydroxide or salt adjuvant (ALUM), polycationic polymer, especially polycationic peptide, especially polyarginine or a peptide containing at least two LysLeuLys motifs, especially KLKLLLLLKLK, immunostimulatory oligodeoxynucleotide (ODN) containing non-methylated cytosine-guanine dinucle-

otides (CpG) in a defined base context (*e.g.* as described in WO 96/02555) or ODNs based on inosine and cytidine (*e.g.* as described in WO 01/93903), or deoxynucleic acid containing deoxy-inosine and/or deoxyuridine residues (as described in WO 01/93905 and WO 02/095027), especially Oligo(dIdC)$_{13}$ (as described in WO 01/93903 and WO 01/93905), neuroactive compound, especially human growth hormone (described in WO 01/24822), or combinations thereof, a chemokine (*e.g.,* defensins 1 or 2, RANTES, MIP1-α, MIP-2, interleukin-8, or a cytokine (*e.g.,* interleukin-1β, -2, -6, -10 or -12; interferon-γ; tumor necrosis factor-α; or granulocyte-monocyte-colony stimulating factor) (reviewed in Nohria and Rubin, 1994), a muramyl dipeptide variant (*e.g.,* murabutide, threonyl-MDP or muramyl tripeptide), synthetic variants of MDP, a heat shock protein or a variant, a variant of Leishmania major LeIF (Skeiky *et al.,* 1995), non-toxic variants of bacterial ADP-ribosylating exotoxins (bAREs) including variants at the trypsin cleavage site (Dickenson and Clements, 1995) and/or affecting ADP-ribosylation (Douce *et al.,* 1997), or chemically detoxified bAREs (toxoids), QS21, Quill A, N-acetylmuramyl-L-alanyl-D-isoglutamyl-L-alanine-2-[1,2-dipalmitoyl-s-glycero-3-(hydroxyphosphoryloxy)]ethylamide (MTP-PE) and compositions containing a metabolizable oil and an emulsifying agent, wherein the oil and emulsifying agent are present in the form of an oil-in-water emulsion having oil droplets substantially all of which are less than one micron in diameter (see, for example, EP 0399843). Also, see Richards *et al.* (1995) for other adjuvants useful in immunization.

[0142] An adjuvant may be chosen to preferentially induce antibody or cellular effectors, specific antibody isotypes (*e.g.,* IgM, IgD, IgA1, IgA2, secretory IgA, IgE, IgG1, IgG2, IgG3, and/or IgG4), or specific T-cell subsets (*e.g.,* CTL, Th1, Th2 and/or T_{DTH}) (see, for example, Munoz *et al.,* 1990; Glenn *et al.,* 1995).

[0143] Unmethylated CpG dinucleotides or motifs are known to activate B cells and macrophages (Stacey *et al.,* 1996). Other forms of DNA can be used as adjuvants. Bacterial DNAs are among a class of structures which have patterns allowing the immune system to recognize their pathogenic origins to stimulate the innate immune response leading to adaptive immune responses (Medzhitov and Janeway, 1997, Curr. Opin. Immunol. 9(1): 4-9). These structures are called pathogen-associated molecular patterns (PAMPs) and include lipopolysaccharides, teichoic acids, unmethylated CpG motifs, double-stranded RNA, and mannins. PAMPs induce endogenous signals that can mediate the inflammatory response, act as co-stimulators of T-cell function and control the effector function. The ability of PAMPs to induce these responses play a role in their potential as adjuvants and their targets are APCs such as macrophages and dendritic cells. PAMPs could also be used in conjunction with other adjuvants to induce different co-stimulatory molecules and control different effector functions to guide the immune response, for example from a Th2 to a Th1 response.

[0144] Disclosed herein is use of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide as vaccinating agent. The vaccines or immunogenic compositions described herein may employ an effective amount of the antigen. There will be included an amount of antigen which will cause the subject to produce a specific and sufficient immunological response so as to impart protection to the subject from subsequent exposure to C. *difficile.* The antigen may be the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. In one embodiment, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide is administered by itself or in combination with an adjuvant.

[0145] Disclosed herein is use of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide as a subunit vaccine. A subunit vaccine refers to the use of a fragment of a pathogen as an inoculating agent. Those skilled in the art will know subunit vaccines offer a means to generate antibodies to a particular part or region of a pathogen.

[0146] Dosage schedule of administration and efficacy of the vaccine can be determined by methods known in the art. The amount of the vaccine and the immunization regimen may depend on the particular antigen and the adjuvant employed, the mode and frequency of administration, and the desired effect (*e.g.,* protection and/or treatment). In general, the vaccine described herein may be administered in amounts ranging between 1 μg and 100 mg, such as e.g. between 60 μg and 600 μg. A single dose of the vaccine comprising the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be in a range from about 1 μg to about 1 mg, preferably from about 5 μg to about 500 μg, more preferably from about 20 μg to about 200 μg. The ratio between C-TAB.G5 or C-TAB.G5.1 isolated polypeptide and adjuvant such as alum may be about 1:1 such as e.g. 1:1.25, but higher ratios may also be used (*e.g.,* about 1:10 or less), or lower ratios may also be used

[0147] (e.g., about 10:1 or more). In an embodiment, in the vaccine comprising the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide the adjuvant aluminum hydroxide will be used in a range from about 50 μg/mL to about 200 μg/mL, preferably in the amount about 125 μg/mL of the final formulation.

[0148] The vaccine comprising the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide can be administered orally, intravenously, subcutaneously, intra-arterially, intramuscularly, intracardially, intraspinally, intrathoracically, intraperitoneally, intraventricularly, and/or sublingually.

[0149] The immunization regimen can be determined by methods known in the art. Administration of the vaccine can be repeated as is determined to be necessary by one skilled in the art. For example, a priming dose may be followed by 1, 2, 3 or more booster doses at weekly, bi-weekly or monthly intervals. In an embodiment of the present invention, the priming dose is followed by one or two booster administration in intervals from about 7 to about 14 days such as e.g. after 7 days and 21 days after first prime. In a preferred embodiment, the therapeutically effective amount of the vaccine is administered two or three times in intervals of 14 days +/- 1, 2 or 3 days (bi-weekly) to a subject. In an

embodiment of the present invention, the therapeutically effective amount of the vaccine is administered once.

[0150] Still another aspect is directed to the population which can be treated according to the present invention. In one embodiment, the population includes healthy individuals who are at risk of exposure to *C. difficile,* especially, the individuals impending hospitalization or residence in a care facility, as well as personals in hospitals, nursing homes and other care facilities. In another embodiment, the population includes previously infected patients who relapsed after discontinuation of antibiotic treatment, or patients for whom antibiotic treatment is not efficient.

[0151] In one more embodiment of the invention, the population includes individuals who are at least 18 years or more of age. In one preferred embodiment, the human subject is from 18 to 65 years old. In another preferred embodiment, the human subject is elderly individuals over 65 years of age. The latter age group being the most vulnerable population suffering from *C. difficile* infections. In some more embodiment, the human subject is younger than 18 years of age.

Methods of using the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide

[0152] Disclosed herein are methods of using the isolated polypeptide. For example, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide may be used to prevent or treat diseases associated with *C. difficile.* By way of example, introducing the isolated polypeptides of the present invention into the immune system of a subject may induce an immune response that includes the subject producing antibodies directed against the isolated polypeptide. Such antibodies are useful for recognizing *C. difficile.*

[0153] Disclosed herein are methods of delivering isolated polypeptides to a subject comprising administering the isolated polypeptide to a subject. The isolated polypeptide may be administered as a liquid or as a solid. The isolated polypeptide may further include a pharmaceutically acceptable carrier.

[0154] Disclosed herein are methods for identifying and isolating variable domains of an antibody that recognize and bind to toxin A and or toxin B comprising use of the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide to produce an immune response, purifying and then characterizing the antibodies produced in response to the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. Identified epitopes may be of use for cloning further antibodies or fragments thereof.

[0155] One aspect of the present disclosure is directed in part to the treatment, the prevention, and the detection of *C. difficile.* In some embodiments, a subject, such as an animal, receives treatment and/or prevention and/or detection of *C. difficile.* In other embodiments, the animal is a human. For example, the polypeptides described herein may be used to raise antibodies to *C. difficile* in vivo. By way of further example, the polypeptides described herein may be used to determine if a subject produces antibodies to *C. difficile.* In some embodiments, the polypeptide is used to isolate antibodies. By way of example, polypeptides may be bound to an affinity matrix.

[0156] By way of further example, the nucleic acid described herein can be used to transform and/transfect cells to recombinantly produce the polypeptides and/or antibodies of the present invention. The nucleic acids described herein may also be used, for example, to determine if a subject is infected with *C. difficile.* By way of example, this can be achieved using methods of radiolabeled hybridization.

[0157] By way of further example, the antibodies described herein can be used to recognize an infection by *C. difficile.* By way of example, the antibodies can recognize native toxin A and/or toxin B as an antigen. The antibodies described herein can also be used to fight an infection by *C. difficile.* By way of example, humanized antibodies or antibody fragments or monoclonal antibodies can employ a subject's own immune response to a *C. difficile* infection. By way of further example, the antibodies described herein may be coupled to a cytokine or a toxin or an enzyme or a marker to assist in treating and detecting an infection.

[0158] Further aspects of the present disclosure relate to diagnostic assays. The present disclosure is of use with many assays known in the art. Those skilled in the art will recognize the wide array of research based uses for the polypeptides, nucleic acids and antibodies described herein. The polypeptides, antibodies and nucleic acids described herein may, for example, be labeled, such as with a radioactive, chemilu minescent, fluorescent and/or dye molecules. The antibodies, nucleic acids and polypeptides described herein lend themselves to use assays for example DNA assays (such as southern blotting), RNA assays (such as northern blotting), protein assays (such as western blotting), chromatographic assays (such as gas, liquid, HPLC, size-exclusion), immunoassays (such as ELISA) and structural assays (such as crystallography and NMR spectroscopy). The antibodies, polypeptides and nucleic acids described herein may further be used as probes. Assays which amplify the signals from a probe are also known to those skilled in the art.

Kits

[0159] Disclosed herein are kits comprising by way of example, and not limitation, nucleic acids encoding the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide, the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide, and/or antibodies against the C-TAB.G5 or C-TAB.G5.1 isolated polypeptide. The kits may include one or more containers and instructions for use in accordance with any of the methods described herein. The C-TAB.G5 or C-TAB.G5.1 isolated polypeptide and/or antibodies described herein may be used in a variety of assays including immunoassays for detecting *C. difficile.*

In one embodiment, the C-TAB.G5 or C-TAB. G5.1 isolated polypeptide serves to function as an antigen for the purposes of detecting antibody in biological samples. The containers may be unit doses, bulk packages (*e.g.,* multidose packages) or sub-unit doses. The kits of this invention are in suitable packaging. Also contemplated are packages for use in combination with a specific device, such as an inhaler, nasal administration device or an infusion device. A kit may have a sterile access port. The container may also have a sterile access port. Kits may optionally provide additional components such as buffers and interpretive information.

[0160] The kits may be used to detect the presence of *C. difficile* or to detect a disease associated with *C. difficile,* such as CDAD. The kits may be used to prevent or treat diseases associated with *C. difficile.* The kits described herein may also be used to alleviate the symptoms of a disease associated with *C. difficile.*

[0161] Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the claimed invention. The following working examples therefore, specifically point out preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

EXAMPLES

Example 1: Preparation of the C-TAB.G5 and C-TAB.G5.1 isolated polypeptides.

[0162] This Example describes the preparation of isolated polypeptides comprising portions of the *C. difficile* toxins A (CTA) and B (CTB) for expression in *E. coli* cells. The method described below can be used for making various isolated polypeptides comprising CTA and CTB. As an example, an isolated polypeptide comprising a portion of the C-terminal domain of CTA and a portion of the C-terminal domain of CTB is described.

Example 1.1: Cloning of the C-TAB.G5 and C-TAB.G5.1 gene constructs.

[0163] The portion of CTA gene (Accession No. YP-001087137) encoding amino acids 2026 to 2710 of the C-terminal domain was amplified by PCR from genomic DNA of *C. difficile* strain 630 (ATCC BAA-1382) using the following primers:

forward: 5'- caccACTAGTatgaacttagtaactggatggc -3'(SEQ ID NO: 9) and
reverse: 5'- CTCGAGttagccatatatcccaggggc -3' (SEQ ID NO: 10).

[0164] Amplification with the forward primer created a *SpeI* site, and amplification with the reverse primer created of a *XhoI* site.

[0165] The portion of CTB gene (Accession No: YP-00108735) encoding amino acids 1850 to 2366 of the C-terminal domain was amplified by PCR using the following primers:

forward: 5'- caccATGCATatgagtttagttaatagaaaacag -3' (SEQ ID NO: 11) and

reverse: 5'- ggcCTCGAGctattcactaatcactaattgagc -3' (SEQ ID NO: 12).

[0166] Amplification with the forward primer created a *NsiI* site, and amplification with the reverse primer created a *XhoI* site.

[0167] PCR reactions were performed using PCR Super-Mix (Invitrogen). The cycle conditions was 95°C for 2 minutes, 95°C for 45 seconds, 55°C for 50 seconds, 68°C for 8 minutes (30 cycles), and 72°C for 10 minutes. The PCR products were purified with Quick gene extraction kit (Invitrogen) and ligated into the PCR 2.1 TOPO vector (Invitrogen). The ligation mixtures were used to transform *E. coli* Mech-1 cells by heat shock. The transformants were plated on plates of ImMedia Amp Blue (Invitrogen). White colonies were picked and cultured in 15 ml tubes with 4 ml of LB medium containing 100 μg/ml ampicillin. Cultures were incubated overnight at 37°C and plasmids were extracted with Quick plasmid miniprep kit (Invitrogen).

[0168] The CTA gene fragment in the PCR 2.1-TOPO/TA vector was digested with *SpeI* and *XhoI,* and the fragment was cloned into an intermediate vector, also digested with *SpeI* and *XhoI,* using T4 DNA Ligase. A linker containing three restriction sites (*BglII-NsiI-SacI*) was then inserted at the 3' end of the CTA gene fragment by PCR using the following set of synthetic primers:

forward: 5'- AGATCTATGCATGAGCTCctcgagcccaaaacgaaaggctcagc -3' (SEQ ID NO: 13)
reverse: 5'- cggtccggggccatatatcccaggggcttttactcc -3' (SEQ ID NO: 14).

[0169] The CTB gene fragment in PCR 2.1 -TOPO/TB was digested with *NsiI* and *XhoI,* and the digested CTB gene

fragment was ligated to the intermediate vector containing the CTA gene and linker, which was also digested with *NsiI* and *XhoI.* The CTB gene was inserted 3' to the linker giving the construct sequence 5'-CTA-linker-CTB-3'. This fusion construct is referred to as C-TAB.V1 intermediate vector.

[0170] The C-TAB.G5 gene was amplified by PCR from C-TAB.V1 intermediate vector using the primers:

forward: 5'- caccCCATTGatggtaacaggagtatttaaagga (SEQ ID NO: 15)

reverse: 5' - CTCGAGctattcactaatcactaattgagctg (SEQ ID NO: 16).

[0171] PCR reactions were performed using PCR Super mix (Invitrogen). The cycle condition was 95 °C for 2 minutes, 95 °C for 45 seconds, 55 °C for 50 seconds, 68 °C for 4 minutes (30 cycles) and 72 °C for 10 minutes. The PCR products were purified with Quick gene extraction kit (Invitrogen) and ligated into the PCR2.1-TOPO vector (Invitrogen). The ligation mixtures were used to transform *E. coli* Mech-1 cells by heat shock. The transformants were plated on plates of ImMedia Amp Blue (Invitrogen). White colonies were picked and cultured in 15 ml tubes with 4 ml of LB medium containing 100 μg/ml ampicillin. Cultures were incubated overnight at 37 °C and plasmids were extracted with Quick plasmid mini-prep kit (Invitrogen). The C-TAB.G5 fusion gene in the PCR 2.1-TOPOTA vector was digested with *NcoI* and *XhoI* restriction enzyme. These C-TAB fragments were ligated into the pET28 expression vector digested with the same restriction enzymes. This resulting construct encodes the toxin A C-terminal domain from amino acids 2272 to 2710 fused to toxin B C-terminal domain from amino acids 1851 to 2366. The pET28/C-TAB.G5 construct was transformed into *E. coli* BL21 (DE3) for expression. Five colonies containing the C-TAB.G5 fusion gene were selected for analysis.

[0172] The C-TAB.G5.1 coding sequence was obtained by codon optimization for improved expression within an *E. coli* host cells. The codon usage was adapted to the codon bias of *E. coli* genes. In addition, GC content was adjusted to prolong mRNA half life; a region of very high (>80 %) or very low (<30 %) GC content have been avoided. Therefore, the optimized gene allows high and stable expression rates in *E. coli.* The codon optimized C-TAB.G5.1 gene was synthesized *in situ* and subcloned into the expression vector pET-28b(+).

[0173] *DNA Sequencing:* Plasmid DNA sequences were confirmed using dye terminator cycle sequencing chemistry with d-Rhodamine dyes. Sequencing data were analyzed using Jellyfish software.

Example 1.2: Expression of the recombinant C-TAB.G5 or C-TAB.G5.1 fusion proteins in *E. coli.*

[0174] Expression of C-TAB.G5 and C-TAB.G5.1 gene constructs may be done using standard procedure for expression in *E. coli.*

[0175] *Screening colonies for expression of the recombinant C-TAB fusion protein:* For the purpose of screening, colonies were picked and grown in 15 ml Falcon tubes with 4 ml of LB media with 50 μg/ml kanamycin. The tubes were cultured overnight at 37°C with mixing at 250 rpm. Following initial growth phase, 1 ml of culture from each tube was transferred to a 24-well tissue culture plate and expression was induced with 1 mM isopropyl-β-D-1-thiogalacto-puranoside (IPTG) for 3 h at 30°C. The cell pellets were collected by centrifugation at 12,000 g for 1 min in microcentrifuge. Cell pellet lysates were prepared, and the soluble fraction was assayed by SDS-PAGE and Western Blot analysis for expression of C-TAB fusion protein. Positive clones were selected for further evaluation.

[0176] *Batch fermentation for C-TAB.G5 expression:* Seed cultures were grown in five 500 ml shake flasks each containing 150 ml Super Broth medium supplemented with 30 μg/ml kanamycin. Cultures were grown for 12 h at 28°C with continuous agitation at 275 rpm until OD_{600} reached 2 - 2.5. The shake flasks were used to inoculate a fermenter containing 10 L Super Broth. The culture was grown approximately 4.5 h at 37°C to OD_{600} = 3.5 - 4. For induction of the product expression 0.1 mM IPTG was added and growth continued for additional 4 h at 25°C. Then the cells were harvested by centrifugation and the cell paste stored frozen at - 70°C. A typical product specific expression rate achieved by this fermentation process was about 200 mg/ml.

[0177] *Fed-batch fermentation for C-TAB.G5.1 preparation:* An aliquot of 500 μl of the glycerol stock of a seed bank (stored at -75°C) was used to inoculate 100 ml pre-culture medium supplemented with 30 μg/ml kanamycin in a 1 L shake flask. The pre-culture was incubated at 37°C under constant agitation at ~150 rpm for approximately 7 h until it reached OD_{600} = 1.0-2.0. 25 mL of pre-culture was used to inoculate 7 L batch fermentation medium in a standard industry 15 L fermenter equipped with process control system, able to perform fed-batch fermentations. 7 L batch culture phase was carried for 12 h at 37°C (OD_{600} = 12-15) until glucose was exhausted. Glucose feed phase (biomass production) was then initiated by an exponential feed mode at a specific growth rate constant μ = 0.25/h at 37°C for 6 h (OD_{600} = 40-50). One hour before switching to a constant feed phase and induction with a final concentration of 1 mM IPTG (product production), temperature was reduced to 30°C to lower the risk of inclusion body formation. Product expression phase was continued for another 5 h with constant feed at 30°C (OD_{600} = ~100), resulting in a total fermentation process time of 23 h and a final culture volume of ~8.2 L. A wet cell biomass of about 1.2 kg was harvested by centrifugation and stored at ≤-70°C. A typical product specific expression rate reached by such fed-batch fermentation was up to 1.3 g/L.

Example 1.3: Purification of the recombinant C-TAB.G5 or C-TAB.G5.1 fusion proteins.

[0178] *Purification of C-TAB.G5 analytical sample:* Frozen cell paste was thawed and resuspended in 10 mM citric acid/NaOH buffer at pH 5.6, and the cell slurry was passed two times through a homogenizer (GEA Niro Soavi homogenizer) at 550 bar. The suspension was centrifuged two times: once at 13500 rpm for 30 minutes and the second time at 18000 rpm in an ultracentrifuge for one hour. The supernatants were pooled, and the pH adjusted to 5.6 with 50 mM citric acid buffer pH 3. Clarified cell lysates were passed over a SP fast flow column with 10 mM citric acid/NaOH buffer at pH 5.6. Proteins were eluted with a liner gradient of sodium chloride increasing from 0 to 500 mM in 20 mM NaPi. Fractions containing the C-TAB.G5 were pooled. The conductivity was adjusted down to 5 mS/cm with distilled H_2O. Tris was added to a 25 mM final concentration. The pooled fractions were passed over a DEAE fast flow column. Protein was eluted with a linear gradient of sodium chloride increasing from 50 to 500 mM in 25 mM Tris. Again, fractions containing C-TAB.G5 were pooled and 1.5 M Na-Citrate, pH 7.5 was added to a final concentration of 0.4 M. The C-TAB.V1 pool was loaded onto a phenol Sepharose HP column equilibrated with 25 mM Tris, 0.4 M Na-Citrate pH 7.5. C-TAB.G5 fusion protein was eluted with a reducing salt concentration in a liner gradient using 5 mM Tris, pH 7.5. All columns were monitored by an AKTA Prime chromatography system. Purified C-TAB fusion protein was buffer exchanged to PBS using a 50 K membrane.

[0179] *Purification of C-TAB.G5.1 bulk preparation:* Biomass was stored at -80°C until processing. 450 g frozen cell paste (equivalent to 2.90 L fermenter) is diluted with 4 volumes of lysis buffer (20 mM Hepes, pH 7.5, ~0.6 mS/cm) (e.g. 450 g paste + 1800 mL buffer) and thawed by this way for ~1 h±0.5 h under mechanical agitation. Optional, remaining clumps can be resuspended using an Ultraturrax (e.g. 5 min at 8000 rpm). Cell lysis is done on a Niro Soavi Panda high homogenizer (640±25 bar, 3 cycles). The lysate is cooled down to < 10°C using a heat exchanger and kept at this temperature until centrifugation. The crude cell lysate is submitted to a batch centrifugation step (Beckmann Avanti JLA 60.25) operated at 14000 rpm (30000 g) at 4°C for 30 min. The supernatants are collected and pooled. The semi-liquid part of the pellet is discarded too, to decrease the risk of clogging the filtration step. The pooled supernatants are then filtered through a Supercap PDH4 100/5 inch depth filter capsule (Pall) (250 cm^2 effective filtration area). The remaining lysate in the filter housing is flushed out with lysis buffer. After clarification, an aliquot of 1M Tris stock solution, pH 7.5 is added to the lysate to a final concentration of 25 mM. The buffer composition of final lysate is 20 mM Hepes, 25 mM Tris, pH 7.5, conductivity ~6 mS/cm). The lysate might be still slightly turbid after filtration, but this does not affect the following capture step. Capture step is performed at room temperature with DEAE Sepharose FF (GE Healthcare) in a XK50/30 column (GE Healthcare) of following dimensions: diameter 50 mm, packed bed height 20 cm, packed bed volume ~400 mL. The loading density is approx. 0.8 to 1.2 g biomass/mL gel. The process is run by an Äkta Explorer system (GE Healthcare) and monitored at 280 nm. Equilibration is performed at 100 cm/h with approx. 5 CV of 25 mM Tris, 20 mM Hepes, 25 mM NaCl, pH 7.5, conductivity ~5 mS/cm until pH, conductivity and 280 nm absorbance are stable. The lysate is loaded onto the column at 75 cm/h and the flow through is discarded. When all filtrated lysate is loaded, flow is resumed with approx. 5 CV of equilibration buffer until the 280nm absorbance is stabilized. Impurities are removed from the column during wash step 2 with 5 CV of 25 mM Tris, 175 mM NaCl, pH 7.5, conductivity 19 mS/cm. The C-TAB protein is eluted from the column by step elution with 3 CV of 25 mM Tris, 375 mM NaCl, pH 7.5, conductivity 36 mS/cm. The collection of the C-TAB containing fractions begins when 280 nm absorbance starts to increase (usually after 1 CV) and lasts for about 0.5 to 1.0 CV. The pooled fractions containing C-TAB can be stored at 2-8°C over night. Intermediate purification step is done with SP-Sepharose FF (GE Healthcare) in a XK50/30 column (GE Healthcare) at room temperature with the following dimensions: diameter 50 mm, packed bed height 20 cm, packed bed volume ~400 mL. The maximum loading density is approx. 4-5 mg C-TAB/mL gel. The process is run by an Äkta Explorer system (GE Healthcare) and monitored at 280 nm. Equilibration, washing and linear gradient elution steps are performed at a maximum flow rate of 200 cm/h (65 mL/min) unless exceeding back pressure (>4 bar) prevents it. Equilibration is performed with approx. 5-10 CV of buffer G at 200 cm/h until pH, conductivity and 280 nm absorbance are stable. Before loading, the DEAE pool has to be adjusted to allow binding of C-TAB on SP-FF resin. DEAE pool is diluted 25 fold with SP-FF equilibration buffer (10 mM citric acid, 2 mM EDTA, pH 5.5±0.1, conductivity ~2 mS/cm) to a final conductivity of not more than 3.5 mS/cm, pH 5.5 ± 0.1. If necessary additional MilliQ water is added to achieve the desired conductivity. Note that low conductivity is very critical to allow binding of C-TAB onto SP-FF. The sample is loaded onto the column at 150 cm/h and the flow through is discarded. After loading the sample, flow is resumed with approx. 5 CV of equilibration buffer at 200 cm/h until the 280 nm absorbance is stabilized. Elution is done by linear gradient at 100 cm/h from 0 % equilibration buffer to 30 % 20 mM sodium phosphate, 500 mM NaCl, pH 7.0 over 10 CV. Fractions are collected and pooling is performed by UV 280 nm absorbance. Pooling starts at 15 % of peak maximum and ends at 15 % of peak maximum. The pool is immediately adjusted to 400 mM citrate (final pH 7, approx. 49 mS/cm) using a 1.5 M citrate stock solution, pH 8.0. The adjusted SPFF pool should have pH 7 and approx. 49 mS/cm and is stored at 2-8°C over night.

[0180] Polishing chromatography step is performed with Phenyl-Sepharose HP (GE Healthcare) in a XK50/30 column (GE Healthcare) at room temperature with the following dimensions: diameter 50mm, packed bed height 15 cm, packed bed volume ~300 mL. The loading density is approx. 4-5 mg C-TAB/mL gel. The process is run by an Äkta Explorer

system (GE Healthcare) and monitored at 280nm. Equilibration, loading, washing and elution steps are performed at a maximum flow rate of 100 cm/h (33 mL/min) unless exceeding back pressure (>4 bar) prevents it. In such a case the flow rate has to be reduced. Equilibration is performed with approx. 5-10 CV of 25 mM Tris, 400 mM sodium citrate, pH 7.5, 46 mS/cm at 100 cm/h until pH, conductivity and 280 nm absorbance are stable. The sample is loaded onto the column at 100 cm/h and the flow through is discarded. After loading the sample, flow is resumed with approx. 5 CV of equilibration buffer at 100 cm/h until the 280 nm absorbance is stabilized. Elution is done by linear gradient at 100 cm/h from 100 % equilibration buffer / 0% 5 mM Tris, pH 7.5, 0.5 mS/cm to 100 % 5 mM Tris, pH 7.5, 0.5 mS/cm over 20 CV. Fractions are collected and pooling is performed by UV280nm absorbance. Pooling starts at approx. 10-15 % of peak maximum and ends at approx. 20 % of peak maximum. The adjusted pool is stored at 2-8°C over night. Preparation of final C-TAB drug substance protein solution is achieved by 30 kDa cut-off tangential flow filtration (TFF, Pellicon 2 membrane, Millipore) operated at room temperature. The protein solution is diafiltered against formulation buffer (20 mM Histidine, 75 mM NaCl, 5 % Sucrose, 0.025 % Tween®80, pH 6.5) until the permeate pH equals 6.5±0.2).

[0181] Final protein concentration is adjusted to 2 mg/mL according to UV measurement at 280 nm using 1.566 as the specific extinction coefficient at 280 nm for C-TAB (protein conc. 1 mg/mL, 1 cm cuvette).

[0182] *SDS-PAGE and Western Blot Analysis:* Whole cell lysates and purified C-TAB.G5 or C-TAB. G5.1 fusion protein were resuspended in Nu-Page sample buffer containing betamercaptoethanol and boiled for 10 min. Samples (25 µl) were loaded onto 3 - 8 % Tris-Acetate gel. Following electrophoresis (150 V for 1 h), proteins were visualized by staining the gels with simply blue stain or used for Western blot analysis.

[0183] C-TAB.G5 or C-TAB.G5.1 specific expression was determined by Western blot analysis using toxin-specific antibodies. Proteins were transferred at 23 V for 60 min onto a PVDF membrane using 1x Transfer buffer in 10 % methanol. Membranes were blocked for 1 h at room temperature with 0.5 % casein in phosphate buffered saline (PBS). Transfer membranes were incubated for 2 hrs at room temperature with either a monoclonal antibody against Toxin B (GenWay; clone B426M) or an in-house derived Guinea Pig polyclonal antibody against Toxin A (List Biological Labs). Washed membranes were incubated with horseradish peroxidase conjugated anti-guinea pig IgG or anti-mouse IgG. The blots were washed and AEC substrates were added. The blots were incubated with gentle mixing for 5-10 minutes. The blots were rinsed with water to stop color development.

[0184] *RBC hemagglutination:* The cell binding domain of toxin A but not toxin B has been shown to be capable of agglutinating rabbit red blood cells (RBCs). The agglutination process is the result of the binding of toxin A to a glycan sequence found on blood antigens on rabbit RBCs. Samples (C-TAB.G5 and native toxin A) are diluted to 100 µg/ml in PBS. In a V-bottom microtiter plate, two-fold serial dilutions are prepared in duplicate across the plate, starting at 100 µg/ml and leaving 50 µl of the dilution in each well. Fifty microliters of a 0.75 % rabbit RBC/PBS suspension is added to each well of the microtiter plate and the plate is incubated for 1 h at room temperature. Hemagglutination is indicated by the failure to form a pellet of RBCs on the bottom of the plate. The hemagglutination titer of a sample is represented by the concentration of protein present in the well with the highest sample dilution in which no RBC pellet is observed.

Example 2: Dose titration of the recombinant C-TAB.G5 fusion protein in the presence and absence of alum in mice.

[0185] This study was to determine the feasibility of an *in vivo* dose titration of C-TAB.G5 with and without alum adjuvant as a C-TAB potency assay. The alum utilized was Alydragel, (alum hydroxide, Brenntag). C57BL/6 female mice (Charles River Labs.), aged between 8 and 9 weeks, were utilized for immunization. All animals received a first immunization by intramuscular (IM) injection (50 µl) into the right thigh muscle on day 0. The second immunization was done by IM injection into the left thigh muscle on day 14. A total of 72 mice were divided into 12 groups vaccinated as follows:

- Group 1: PBS only
- Group 2: 100 (154) ng C-TAB.G5
- Group 3: 300 (462) ng C-TAB.G5
- Group 4: 1,000 (1,540) ng C-TAB.G5
- Group 5: 3,000 (4,620) ng C-TAB.G5
- Group 6: 10,000 (15,400) ng C-TAB.G5
- Group 7: PBS with 50 µg alum
- Group 8: 10.0 (15.4) ng C-TAB.G5 with 50 µg alum OH
- Group 9: 30.0 (46.2) ng C-TAB.G5 with 50 µg alum OH
- Group 10: 100 (154) ng C-TAB.G5 with 50 µg alum OH
- Group 11: 300 (462) ng C-TAB.G5 with 50 µg alum OH
- Group 12: 1,000 (1,540) ng C-TAB.G5 with 50 µg alum OH

[0186] In this study the protein concentration was firstly determined according to the standard protocol Quick Start™

Bradford Protein Assay (Bio-Rad). Lately, the protein concentration (shown in parentheses) was re-determined by UV measurement at 280 nm according to the procedure described in Example 1.3. In all follow-up studies the protein concentration was measured by UV method.

[0187] Blood samples were collected from all animals two weeks after the first immunization (study day 14) and two weeks after the second immunization (study day 28). The serum was stored at -20°C until analyzed.

[0188] *Serum IgG ELISA:* Serum antibodies elicited to C-TAB.G5 or C-TAB.G5.1 (referred as C-TAB), toxin A and toxin B or toxoids thereof were evaluated in an enzyme linked immunosorbent assay (ELISA). Briefly, stock solutions of 1.0 μg/ml of toxin A, toxin B or the C-TAB.G5 isolated polypeptide were prepared in PBS and 100 μl were added to each well of a 96-well plates. After overnight incubation at 4°C, the plates were washed and blocked with 0.5 % casein blocking buffer. Plates were washed again and serial, two-fold dilutions of test sera added to the plates. After a second overnight incubation at 4°C, plates were washed and incubated with peroxidase-conjugated anti-mouse IgG (H + L). After a 2 hours incubation at room temperature, the plates were again washed, peroxidase substrate (2,2'-azinobis(3-ethylbenzthiazoline-6-sulfonate) added and color allowed to develop for 2 h at room temperature. The reaction was stopped by adding 50 μl of 2 % SDS to the wells. Plates are read with an ELISA plate reader at an absorbance of 405 nm. Serum antibody titers are reported as the geometric mean of ELISA Units, which are the serum dilutions that results in an OD 405 nm reading of 1.0. As a negative control a pool sample of pre-immune serum obtained from animals pre-bled before the first immunization was used to evaluate an antibody response.

[0189] Animals receiving C-TAB.G5 demonstrated a dose dependent increase in antibody titers, with the alum adjuvant allowing for significantly improved antibody titers at a lower dose of C-TAB. G5. Figure 3 shows the titers for anti C-TAB, anti-toxin A and anti-toxin B IgG. Figure 4 shows a graphical comparison of antibody titers in the presence or absence of alum.

Example 3: Immunogenicity and protective efficacy of C-TAB.G5 in mice.

[0190] This study was to evaluate the immunogenicity and protective efficacy of C-TAB.G5 in vaccinated mice receiving a lethal challenge of *C. difficile* toxin A or toxin B. Female C57BL/6 mice (Charles River Labs.), aged 6-7 weeks, were utilized for this study. All animals received the first vaccination by intramuscular (IM) injection (50 μl) into the right thigh muscle on day 0. The second vaccination was done by IM injection into the left thigh muscle on day 14. 116 mice were divided into groups vaccinated as follows:

- Group 1: PBS only
- Group 2: 3 μg C-TAB.G5
- Group 3:10 μg C-TAB.G5
- Group 4: 30 μg C-TAB.G5
- Group 5: 3 μg C-TAB.G5 + 50 μg alum OH
- Group 6: 10 μg C-TAB.G5 + 50 μg alum OH
- Group 7: 30 μg C-TAB.G5 + 50 μg alum OH
- Group 8:: PBS only
- Group 9: 3 μg C-TAB.G5
- Group 10: 10 μg C-TAB.G5
- Group 11: 30 μg C-TAB.G5
- Group 12: 3 μg C-TAB.G5 + 50 μg alum OH
- Group 13: 10 μg C-TAB.G5 + 50 μg alum OH
- Group 14: 30 μg C-TAB.G5 + 50 μg alum OH

[0191] Blood samples were collected from all animals two weeks after the second immunization (study day 28). The serum was stored at -20°C until analyzed. Serum antibody titers to C-TAB, toxin A and toxin B were then determined by ELISA and reported as ELISA Units (EU).

[0192] Figure 5 shows serum antibody titers to C-TAB, toxin A and toxin B in mice evaluated two weeks after the second immunization (study day 28). This study demonstrated that the C-TAB.G5 fusion protein is highly immunogenic in mice and is able to induce strong antibody response against both toxin A and toxin B even without adding an adjuvant. The C-TAB.G5 immunogenicity can be significantly augmented (more than a one log) by co-delivery with alum hydroxide. The animals receiving C-TAB.G5 with or without alum demonstrated 2-fold increased antibody response over a one log dose range.

[0193] Besides evaluating antibody titers, the antibodies generated by immunization with C-TAB.G5 were assessed for their ability to neutralize native toxin A and B in *in vitro* toxin neutralization assay (TNA).

[0194] *Toxin Neutralizing Antibody Assay (TNA).* For *in vitro* analysis, 125 μl of either toxin A (5 ng/ml) or toxin B (1 ng/ml) was incubated with 125 μl of serial dilutions of anti-sera obtained from immunized mice. After one hr of incubation

at 37°C, the toxin:serum mixture was added to microtiter wells containing Vero cells (monkey kidney cells), and the microtiter plates incubated for 18 hr. Incubation of either toxin A or B with Vero cells resulted in a change in cell morphology and a loss of cell adherence which was measured by neutral red staining of toxin treated cells after removal of non-adherent cells. The toxin neutralization titer of a serum is reported as the serum dilution which gives a 50 % reduction in toxin activity.

[0195] The results of the TNA assay are shown in Figure 6. The data indicate that antibodies generated following immunization with the C-TAB.G5 alone are capable of neutralizing the toxic activity of native toxin A but not toxin B. When the C-TAB. G5 was co-delivered with alum, TNA titers were augmented with approximately 6-fold increase in anti-toxin A TNA and only 2-fold lower titers in anti-toxin B TNA. This data indicates that the C-TAB.G5 isolated polypeptide not only retains the antibody recognition antigenic epitopes present in the native toxins, but comprises critical antigenic epitopes required for the generation of functional toxin neutralizing antibody. Thus, C-TAB.G5 is effective in neutralizing toxic effects of C. *difficile* toxin A and toxin B and, therefore, is useful in vaccination.

[0196] In addition to assessing antibody response, the ability of C-TAB.G5 immunization to protect mice from a lethal challenge of native toxins was determined. Three weeks after the second vaccination (study day 35) animals in vaccinated and non-vaccinated groups (N=8) received intraperitoneally (IP) a lethal dose of either 25 ng of toxin A or 50 ng of toxin B. Survival of the mice was monitored over the following 9 days and the results are shown in Figure 6. This experiment demonstrated that immunization of mice with C-TAB.G5 in the absence of the alum adjuvant was capable of conferring 100 % protection against a lethal challenge with native toxin A and 50 % protection against toxin B challenge. Co-delivery of C-TAB.G5 with Alum enhanced the protective immunity to toxin B up to 100 % protection. This data indicates that C-TAB.G5 vaccination induces an immune response sufficient to protect mice from the toxic effects of both toxin A and B in the lethal challenge model.

Example 4: Evaluation of the immunogenicity and protective efficacy of C-TAB.G5 in young and aged mice.

[0197] This study was to compare the immune response mounted against C-TAB.G5 in young and aged mice. Female C57BL/6 mice (Charles River Labs.), aged 6-7 weeks and 18 months, respectively, were utilized for this study. All animals received the first vaccination by intramuscular (IM) injection (50 μl) into the right thigh muscle on day 0. The second vaccination was done by IM injection into the left thigh muscle on day 14. 192 mice were divided into groups vaccinated as follows:

- Group 1: PBS to young mice
- Group 2: PBS to aged mice
- Group 3: 10 μg C-TAB.G5 to young mice
- Group 4: 30 μg C-TAB.G5 to young mice
- Group 5: 10 μg C-TAB.G5 to aged mice
- Group 6: 30 μg C-TAB.G5 to aged mice
- Group 7: 10 μg C-TAB.G5 + 50 μg alum OH to young mice
- Group 8: 30 μg C-TAB.G5 + 50 μg alum OH to young mice
- Group 9: 10 μg C-TAB.G5 + 50 μg alum OH to aged mice
- Group 10: 30 μg C-TAB.G5 + 50 μg alum OH to aged mice
- Group 11: PBS to young mice
- Group 12: PBS to aged mice
- Group 13: 10 μg C-TAB.G5 to young mice
- Group 14: 30 μg C-TAB.G5 to young mice
- Group 15: 10 μg C-TAB.G5 to aged mice
- Group 16: 30 μg C-TAB 5^{th} to aged mice
- Group 17: 10 μg C-TAB.G5 + 50 μg alum OH to young mice
- Group 18: 30 μg C-TAB.G5 + 50 μg alum OH to young mice
- Group 19: 10 μg C-TAB.G5 + 50 μg alum OH to aged mice
- Group 20: 30 μg C-TAB.G5 + 50 μg alum OH to aged mice

[0198] Three weeks after the second vaccination (study day 35) animals in vaccinated and non-vaccinated groups (N=6) received a lethal challenge by intraperitoneal (IP) injection with 25 ng toxin A or 50 ng toxin B. Survival of the mice was monitored over the following 9 days.

[0199] Blood samples were collected from all animals two weeks after the first immunization (study day 14) and two weeks after the second immunization (study day 28). The serum was stored at -20°C until analyzed. Serum antibody titers to C-TAB, toxin A and toxin B were then determined by and reported as ELISA Units (EU). Toxin A and toxin B neutralizing antibodies (TNA) were determined using Vero cells treated with a cytotoxic amount of recombinant toxin A

and toxin B.

[0200] Young animals receiving the C-TAB.G5 vaccine demonstrated significantly higher levels of all antibodies tested, as compare to old animals. Especially high antibody titers were obtained in young mice vaccinated with C-TAB.G5 in the presence of alum hydroxide (Figure 7). Particularly significant improvement was achieved in toxin B TNA titer. At the same time, there was no big difference between young and aged mice in ability to withstand the toxin A and toxin B challenges. However, both groups demonstrated improved protection rate when vaccinated in the presence of alum. Figure 7 shows a comparison of C-TAB.G5 immunogenicity and protective efficacy in young vs. old mice. Figure 8 shows the kinetics of anti-C-TAB antibody development in young and old mice.

Example 5: Comparison of the immunogenicity and protective efficacy of C-TAB.G5.1 and toxoid A and B.

[0201] This study was to compare the immunogenicity and protective efficacy of C-TAB.G5.1, vs. toxoid A/B. The toxoid A/B used was the mixture of equal parts (1:1) of toxoid A (lot #1009132) and toxoid B (lot # 1009133). Toxoid was prepared by formalin fixation and provided by TechLab. Female C57BL/6 mice (Charles River Labs.), aged 6-7 weeks, were utilized for this study. All animals received the first vaccination by intramuscular (IM) injection (50 μl) into the right thigh muscle on day 0. The second vaccination was done by IM injection into the left thigh muscle on day 14. 180 mice were divided into groups vaccinated as follows:

- Group 1: PBS only
- Group 2: 10 μg C-TAB.G5.1
- Group 3: 30 μg C-TAB.G5.1
- Group 4: 10 μg C-TAB.G5.1 + 50 μg alum OH
- Group 5: 10 μg C-TAB.G5.1 + 50 μg alum OH
- Group 6: 30 μg toxoid A/B
- Group 7: 10 μg toxoid A/B
- Group 8: 30 μg toxoid A/B + 50 μg alum OH
- Group 9: 30 μg toxoid A/B + 50 μg alum OH
- Group 10: PBS
- Group 11: 10 μg C-TAB.G5.1
- Group 12: 30 μg C-TAB.G5.1
- Group 13: 10 μg C-TAB.G5.1 + 50 μg alum OH
- Group 14: 30 μg C-TAB.G5.1 + 50 μg alum OH
- Group 15: 10 μg toxoid A/B
- Group 16: 30 μg toxoid A/B
- Group 17: 10 μg toxoid A/B + 50 μg alum OH
- Group 18: 30 μg toxoid A/B + 50 μg alum OH

[0202] Three weeks after the second vaccination (study day 35) animals in vaccinated and non-vaccinated groups (N=6) received a lethal challenge by intraperitoneal (IP) injection with 28 ng toxin A or 50 ng of toxin B. Survival of the mice was monitored over the following 9 days.

[0203] Blood samples were collected from all animals two weeks after the first immunization (study day 14) and two weeks after the second immunization (study day 28). The serum was stored at -20°C until analyzed. Serum antibody titers to C-TAB, toxin A and toxin B were then determined by ELISA and reported as ELISA Units (EU). Toxin A and toxin B neutralizing antibodies (TNA) were determined using Vero cells treated with a cytotoxic amount of recombinant toxin A and toxin B.

[0204] This study demonstrates immunogenicity and protective efficacy of C-TAB.G5.1 and toxoid A/B in mice after two vaccinations. Animals receiving C-TAB.G5.1 showed lower but significant anti-C-TAB antibody titers, as compare to animals receiving toxoid A/B. Also, co-delivery of alum greatly augmented all tested antibody responses. As a result, the level of anti-C-TAB and anti-toxin A antibodies achieved in animals immunized either with C-TAB.G5.1 or with toxoid A/B in the presence of alum are similar. The only lower antibody titer was observed for anti-toxin B antibody when mice were immunized with C-TAB.G5.1, as compare to mice immunized with toxoid A/B. Noteworthy, unlike the antibodies generated against C-TAB.G5.1 recognizing epitopes in the C-terminal portion of the toxin molecules, antibodies induced with toxoid immunization were specific to the N-terminal portion of the toxin molecules, which was read out in the anti-toxin ELISA. Thus, anti-toxin A and anti-toxin B antibodies generated in mice immunized with C-TAB.G5.1 and toxoid A/B were antibodies of different specificity and, therefore, can not be compared directly. However, the data indicates that antibody response to C-TAB. G5.1 immunization is significantly high, like in case of immunization with toxoids. In addition, the toxin challenge study demonstrated that ability of C-TAB.G5.1 immunization to protect mice against from a lethal challenge is comparable to protection efficacy of toxoid A and B. Figure 9 shows a comparison of the immuno-

genicity of C-TAB.G5.1 and toxoid A/B. Figure 10 shows the toxin neutralization and protection data for mice immunized with C-TAB.G5.1 as compared to those immunized with toxoid A/B.

Example 5.1: Comparison of antibody titers and protective efficacy of C-TAB.G5.1 in different immunization regimens.

[0205] This study was to compare the immunogenicity and protective efficacy of C-TAB.G5.1 in mice receiving three doses of the vaccine in different immunization regimens. Female C57BL/6 mice (Charles River Labs.), aged 6-7 weeks, were utilized for this study. 135 mice were divided into 14 groups. All animals in groups no. 2-13 received three vaccinations by intramuscular (IM) injection (50 μl) into the right thigh muscle or left thigh muscle in days indicated below. Mice in groups 1 and 8 did not received vaccination; they served as a negative control. The immunization was performed as follows:

Group	Vaccine	Immunization day (route)
1	-	-
2	10 μg G5.1	0 (R), 3 (L), 14 (R)
3	10 μg G5.1 + 50 μg Alum	0 (R), 3 (L), 14 (R)
4	10 μg G5.1	0 (R), 7 (L), 21 (R)
5	10 μg G5.1 + 50 μg Alum	0 (R), 7 (L), 21 (R)
6	10 μg G5.1	0 (R), 14 (L), 28 (R)
7	10 μg G5.1 + 50 μg Alum	0 (R),14 (L), 28 (R)
8	-	-
9	10 μg G5.1	0 (R), 3 (L), 14 (R)
10	10 μg G5.1 + 50 μg Alum	0 (R), 3 (L), 14 (R)
11	10 μg G5.1	0 (R), 7 (L), 21 (R)
12	10 μg G5.1 + 50 μg Alum	0 (R), 7 (L), 21 (R)
13	10 μg G5.1	0 (R), 14 (L), 28 (R)
14	10 μg G5.1 + 50 μg Alum	0 (R), 14 (L), 28 (R)
R) = right thigh muscle (L) - left thigh muscle		

[0206] Blood samples were collected from all animals on study day 0, 3, 7, 14, 21, 28, 35 and 42. The serum was stored at -20°C until analyzed. Serum antibody titers to C-TAB, toxin A and toxin B were determined by ELISA and reported as ELISA Units (EU). Toxin A and toxin B neutralizing antibodies (TNA) were determined on study day 42 using Vero cells treated with a cytotoxic amount of recombinant toxin A and toxin B.

[0207] Three weeks after the last vaccination (study day 49) animals in vaccinated and non-vaccinated groups (N=8) received a lethal challenge by intraperitoneal (IP) injection with 28 ng toxin A or 50 ng toxin B. Survival of the mice was monitored over the following 9 days.

[0208] This study demonstrates that all antibody titers measured two weeks after the third vaccination or on study day 35 and 42 are comparable in all immunization regimens, although the immunization regimen of 0/14/28 shows the best antibody responses. If compare antibody titers measured two weeks after the second vaccination, then the immunization regimen of 0/14/28 is better than the immunization regimen of 0/7/21, and much better than the immunization regimen of 0/3/14. This study confirmed that anti-toxin A/B antibody titers are significantly enhanced when the antigen is co-injected with aluminum hydroxide (data not shown). The study also shows that even two doses of the vaccine with alum administered in two-week interval can elicit high antibody level, comparable to the level obtained after three dose vaccinations.

[0209] The level of toxin A/B neutralizing antibodies is much higher in the immunization regimen of 0/7/21 and 0/14/28 than in the immunization regimen of 0/3/14.

[0210] Complete protection against challenge with toxin A does not require alum in the immunization regimen of 0/7/21 and 0/14/28 but not in 0/3/14. The immunization regimen of 0/14/28/ with alum induced a highest level of protection (87.5%) against toxin B challenge, while the immunization regimen of 0/7/21 provides 37.5 % of protection and 0/3/14 shows 28.6 % of protection. Results of this study are shown in Figure 20.

Example 6: Evaluation of the immunogenicity and protective efficacy of the recombinant C-TAB.G5.1 fusion protein in hamsters.

[0211] This study was to further evaluate the immunogenicity of the recombinant fusion protein C-TAB.G5.1 administered with or without adjuvant in a different animal model.

[0212] Female hamsters (Harlan), aged over 7 weeks and weighing between 80 and 90 g were utilized for this study. All animals received the first vaccination by bolus (50 μl) intramuscular (IM) injection into the right thigh muscle on day 0. The second vaccination was by IM injection into the left thigh muscle on day 14 and the third vaccination was by IM injection on day 28. Hamsters were divided into groups (N=6) and vaccinated as follows:

- Group 1: Formulation buffer only
- Group 2: 10 μg C-TAB.G5.1
- Group 3: 10 μg C-TAB.G5.1 + 100 μg alum OH
- Group 4: 30 μg C-TAB.G5.1
- Group 5: 30 μg C-TAB.G5.1 + 100 μg alum OH
- Group 6: 100 μg C-TAB.G5.1
- Group 7: 100 μg C-TAB.G5.1 + 100 μg alum OH
- Group 10: Formulation buffer only
- Group 11. 10 μg C-TAB.G5.1
- Group 12. 10 μg C-TAB.G5.1 + 100 μg alum OH
- Group 13. 30 μg C-TAB.G5.1
- Group 14. 30 μg C-TAB.G5.1 + 100 μg alum OH
- Group 15 100 μg C-TAB.G5.1
- Group 16. 100 μg C-TAB.G5.1 + 100 μg alum OH

[0213] Two weeks after the third vaccination (study day 42) animals in vaccinated and non-vaccinated groups (N=6) received a lethal challenge by intraperitoneal (IP) injection with 75 ng toxin A or 125 ng toxin B. An extra 12 hamsters were used for a dose titration of toxin A or toxin B challenge on the day 44. Survival of the hamsters was monitored over the following 8 days.

[0214] Blood samples were collected from all animals two weeks after the first immunization (study day 14), after the second immunization (study day 28) and the third immunization (study day 35). The serum was stored at -20°C until analyzed. Serum antibody titers to C-TAB, toxin A and toxin B were then determined by ELISA and reported as ELISA Units (EU). Toxin A and toxin B neutralizing antibodies (TNA) were determined using Vero cells treated with a cytotoxic amount of recombinant toxin A and toxin B.

[0215] This study demonstrated that hamsters, similarly to mice, were able positively respond to the C-TAB.G5.1 vaccination. Animals receiving C-TAB.G5.1 demonstrated a dose dependent increase in all tested antibody titers, while the alum adjuvant significantly improved antibody titers at all doses of C-TAB.G5. The highest antibody titers were observed two weeks after the second shots (study day 28). Figure 11 (A-C) shows antibody titers for each group of immunized hamsters. Figure 12 shows the kinetics of anti-C-TAB antibody development in hamsters immunized with C-TAB. G5 in the presence or absence of alum hydroxide.

[0216] The results of the TNA assay are shown in Figure 13. These results are similar to those obtained for mice and indicate that antibody generated against the C-TAB.G5.1 fusion protein in hamsters are effective in neutralizing toxic effects of *C. difficile* toxin A and toxin B.

[0217] Figure 13 also shows protection data for hamsters immunized with C-TAB.G5.1 following a lethal toxin challenge. High protection was achieved even by vaccination with C-TAB. G5.1 in the absence of the adjuvant. The protection level was improved to 100 % by adding alum to the vaccine.

Example 7: The protective efficacy of the C-TAB.G5.1 fusion protein against a *C. difficile* spore challenge in clindamycin-treated hamsters.

[0218] Following antibiotic treatment *C. difficile* can colonize the gut and, if toxigenic, may cause an antibiotic associated diarrhea. *C. difficile* associated disease (CDAD) of humans is modeled in hamsters using clindamycin to make the animals susceptible to colonization, diarrhea and death, usually within a few days after seeding with a toxigenic strain. To assess the efficacy of the C-TAB.G5.1 vaccine, vaccinated and non-vaccinated hamsters were challenged with clindamycin and *C. difficile* strain 630. 100 μg of C-TAB.G5.1 was mixed with 125 μg alum-hydroxide adjuvant. Female adult hamsters weighting ~100 g received 3 vaccinations by intramuscular (IM) injection on days 0, 14 and 28. The placebo was PBS. 48 hamsters were divided into groups of 8 as vaccinated as follows:

Group 1: PBS only + 10^2 spore challenge
Group 2: C-TAB.G5.1 + 10^2 spore challenge
Group 3: PBS only + 10^3 spore challenge
Group 4: C-TAB.G5.1 + 10^2 spore challenge
Group 5: PBS only + 10^4 spore challenge
Group 6: C-TAB.G5.1 + 10^4 spore challenge

[0219] On day 42 all animals in all groups received an oral dose of 10 mg clindamycin phosphate/ kg body weight. On day 43 all animals in all groups were dosed by oral gavage with washed spores of *C. difficile* strain 630. Three levels of spore challenge were used (~10^2, 10^3 and 10^4). Observation, but no treatment, continued until day 54. At study termination, all surviving animals were disease free for ≥ 5 days.

[0220] Blood samples were drawn to obtain serum for serological studies on day 0, 14, 28, 42 and day 54 (end of study). Feces were collected on days 1 and 42, directly from the anus of the hamsters, or if needed, from among the bedding.

[0221] Results are shown on Figure 14 demonstrating survival curves after spore challenge in hamsters. Survival data was plotted as Kaplan-Meier survival fit curves and statistical analysis was done using a log rank analysis. At all spore doses, 100 % survival of hamsters in the vaccinated group was observed and survival was significantly enhanced when compared to the placebo group: p=0.0245 at 10^2 spores, p=0.0006 at 10^3 spores, p<0.0001 at 10^4 spores.

Example 8: Immunogenicity and protection efficacy of C-TAB.G5.1 in monkeys.

[0222] This study was to evaluate the immunogenicity and protection of C-TAB.G5.1 in cynomolgus monkeys. Six female cynonomolgus monkeys, aged between 4 and 6 years and weighing between 2 and 4 kg, were used for this study. Two groups of three monkeys were arranged, the first group (Group 1) receiving 200 μg of C-TAB.G5.1 and the second (Group 2) receiving 200 μg of C-TAB.G5.1 and 250 μg alum. As alum adjuvant Rehydragel (Reheis, lot#534401, dilute in PBS to 2 mg/ml) was used. Before blood collection or immunization, animals were shaved (if necessary).

[0223] The 1st (study day 0) and 3rd (study day 28) immunizations were injected on the left arm (deltoid), the 2nd immunization (study day 14) was injected to the right arm (deltoid). Group 1 received 200 μg C-TAB.G5.1 alone in 0.5ml 1xPBS by IM injection and Group 2 received 200 μg C-TAB.G5.1 with 250 μg alum in 0.5ml 1xPBS by IM injection.

[0224] At the established time points (study days 0, 14, 28 and 42), 2-3 mL of whole blood was obtained by standard methods into serum separator tubes. Serum samples were frozen at approximately -20°C. ELISA method was then used to assess anti-C-TAB, anti-toxin A and anti-toxin B IgG titers. Antibody titers were presented in ELISA Units (EU).

[0225] Figure 15 shows that increased doses of C-TAB. G5.1 lead to increased antibody production recognizing all three proteins, while the presence of alum significantly improved antibody levels. The highest antibody titers were observed with two vaccinations on day 42. These data clearly indicate feasibility of using the recombinant C-TAB.G5 or C-TAB.G5.1 fusion proteins for vaccination subjects in need thereof.

Example 9: Comparison of the immunogenicity of C-TAB.G5 and C-TAB.G5.1.

[0226] This study was to compare the immunogenicity of C-TAB.G5 and C-TAB.G5.1 as well as the effect of two different buffers in which the C-TAB was delivered in. C57BL/6 female mice (Charles River Labs.), aged between 8 and 9 weeks, were utilized for immunization. All animals received the first immunization by intramuscular (IM) injection (50 μl) into the right thigh muscle on day 0. The second immunization was done by IM injection into the left thigh muscle on day 14. A total of 72 mice were divided into 12 groups vaccinated as follows:

Group 1: 1 μg C-TAB.G5 in PBS
Group 2: 3 μg C-TAB.G5 in PBS
Group 3: 10 μg C-TAB.G5 in PBS
Group 4: 30 μg C-TAB.G5 in PBS
Group 5: 1 μg C-TAB.G5 in histidine buffer
Group 6: 3 μg C-TAB.G5 in histidine buffer
Group 7: 10 μg C-TAB.G5 in histidine buffer
Group 8: 30 μg C-TAB.G5 in histidine buffer
Group 9: 1 μg C-TAB.G5.1 in histidine buffer
Group 10: 3 μg C-TAB.G5.1 in histidine buffer
Group 11: 10 μg C-TAB.G5.1 in histidine buffer
Group 12: 30 μg C-TAB.G5.1 in histidine buffer

[0227] Blood samples were collected from all animals two weeks after the second immunization (study day 28). The serum was stored at -20°C until analyzed. Serum antibody titers to C-TAB, toxin A and toxin B were determined by ELISA and reported as ELISA Units.

[0228] Figure 16 shows that all antibody titers (anti-C-TAB, anti-toxin A and anti-toxin B) were not significantly different (as revealed by T-test analysis) over 1-30 μg dose range for three vaccine formulations. Slightly higher antibody production was achieved with C-TAB.G5 formulation in histidine buffer, as compare to PBS. No significant difference was observed between immunization with C-TAB.G5 and C-TAB.G5.1 histidine formulations. Thus, this study demonstrates the equal immunogenicity of C-TAB.G5 and C-TAB.G5.1 constructs.

Example 10: Preparation and evaluation of the alternative C-TABNCTB and C-TADCTB fusion proteins.

[0229] This Example describes the preparation of two other fusion proteins comprising one portion of the C-terminal domain of CTA and two portions of the C-terminal domain of CTB derived from *C. difficile* VPI-10463 strain. The C-TABNCTB fusion protein (SEQ ID NO: 18) comprises, like C-TAB.G5, 19 repeating units of CTA (amino acids 2272-2710), 23 repeating units of CTB (amino acids 1850-2366), plus additional 10 repeats of CTB (amino acids 1834-2057) fused to the C-terminus of CTB. The C-TADCTB fusion protein (SEQ ID NO: 20) comprises C-TAB.G5 sequence (19 repeats of CTA and 23 repeats of CTB) plus additional 24 repeating units of CTB (amino acids 1834-2366) fused to the C-terminus of C-TAB.G5. Thus, C-TADCTB comprises a double portion of repeating units of CTB. Cloning of the C-TABNCTB and C-TADCTB gene constructs was done in a way similar to that described in Example 1. 1 The recombinant fusion proteins were expressed in *E. coli* cells and purified using standard procedure as described in Example 1.2. The isolated polypeptides were evaluated in the immunogenicity and protection studies in animals.

Example 10.1: Comparison of the immunogenicity and protective efficacy of C-TAB.G5, C-TABNCTB and C-TADCTB in mice.

[0230] This study was to compare the immunogenicity and protective efficacy of C-TAB.G5, C-TABNCTB and C-TADCTB in mice vaccinated with five antigen doses over a two log range. Female C57BL/6 mice (Charles River Labs.), aged 6-7 weeks, were utilized for this study. All animals received two vaccinations: the first one by intramuscular (IM) injection (50 μl) into the right thigh muscle on day 0. The second vaccination was done by IM injection into the left thigh muscle on day 14. All immunizations were done in the absence of alum. Blood samples were collected two weeks after the second immunization (study day 28). The serum was stored at-20°C until analyzed. Serum antibody titers to C-TAB, toxin A and toxin B were determined by ELISA and reported as ELISA Units (EU) shown in Figure 17.

[0231] This study demonstrated that the alternative fusion proteins C-TADCTB and C-TABNCTB, as well as C-TAB.G5, are highly immunogenic and able to induce strong antibody response against both toxin A and toxin B even without adding an adjuvant.

[0232] In addition to assessing antibody response, the ability of C- TADCTB and C-TABNCTB immunization to protect mice from a lethal challenge of native toxin B was determined. Three weeks after the second vaccination (study day 35) animals in vaccinated and non-vaccinated groups (N=6) received intraperitoneally (IP) a lethal dose of 50 ng of toxin B. Survival of the mice was monitored over the following 9 days and the results are shown in Figure 18. This experiment demonstrated that immunization of mice with 33 μg of C-TADCTB in the absence of alum was capable of conferring 100 % protection against a lethal challenge with native toxin B, while the same dose of C-TAB.G5 and C-TABNCTB induces only partial protection. This data indicates that, similarly to C-TAB.G5, two other fusion proteins C-TADCTB and C-TABNCTB may be protective against the lethal challenge with the native toxin.

Example 10.2: Comparison of the immunogenicity and protective efficacy of C-TAB.G5.1 and C-TADCTB in hamsters.

[0233] This study was to further evaluate the immunogenicity of the alternative fusion protein C-TADCTB administered with or without alum adjuvant in a different animal model.

[0234] The study was designed as described in Example 6: female hamsters were vaccinated three times by IM injection (study day 0, 14 and 28) in the presence or absence of 100 μg alum hydroxide. Two weeks after the third vaccination (study day 42) all animals received a lethal challenge by intraperitoneal (IP) injection with 75 ng toxin A or 125 ng toxin B. Blood samples were collected on study day 14, 28 and 35 and serum antibody titers to C-TAB, toxin A and toxin B were determined by ELISA. Toxin A and toxin B neutralizing antibodies (TNA) were measured in day 35 sera. Survival of the hamsters was monitored and reported as % of protection.

[0235] This study demonstrated that the fusion protein C-TADCTB can induce anti-toxin antibody response in hamsters, similarly to mice. The alum adjuvant significantly improved all tested antibody titers. The results of the TNA assay shown in Figure 19 indicate that antibody generated against C-TADCTB are effective in neutralizing toxic effects of *C. difficile* toxin A and toxin B. Figure 19 also demonstrates comparison of protection data for hamsters immunized either with C-

TAB.G5.1 or with C-TADCTB. High protection was achieved by vaccination with both recombinant fusion proteins.

Example 11: An open-label phase 1 study assessing the safety, immunogenicity and dose response of a pharmaceutical composition comprising C-TAB. G5.1

[0236] The pharmaceutical composition comprising C-TAB.G5.1, a recombinant fusion protein consisting of truncated *Clostridium difficile* (*C. difficile*) Toxin A and Toxin B, which will be administered at three different doses: 20 μg with $Al(OH)_3$ (alum), 75 and 200 μg without or with $Al(OH)_3$, respectively, intramuscular (IM) injection, three vaccinations on Day 0, 7 and 21.

STUDY OBJECTIVES

Primary:

[0237]

- To investigate the safety and tolerability of a pharmaceutical composition comprising C-TAB. G5.1 up to 6 months after the third vaccination.

Secondary:

[0238]

- To investigate the immune response measured against the vaccine antigen C-TAB.G5.1 and the native Toxins A and B of *C. difficile* to three different doses and two formulations on Days 0, 7, 14, 21, 28, 113, 201 after the first vaccination to obtain a first indication of the optimal dose and formulation.

- To investigate the capacity of C-TAB.G5.1 vaccine-induced IgG antibodies to neutralize *C. difficile* Toxins A and B in vitro.

STUDY DESIGN

[0239] This is an open-label, partially randomized, dose escalation Phase 1 study which will consist of a part A in healthy adults aged between ≥18 and <65 years and a part B in healthy elderly ≥65 years, the latter age group being the most vulnerable population to suffer from *C. difficile* infections. Part A will be conducted with vaccination schedule Day 0, 7 and 21 in five treatment groups of 12 healthy adult subjects to study safety and dose response to 20 μg C-TAB.G5.1 vaccine with adjuvant, and to 75 μg and 200 μg of C-TAB.G5.1 vaccine with or without adjuvant, respectively. Safety and immunogenicity will be analyzed after all adult subjects of part A have received the third vaccination, all safety data will be reviewed by a Data Safety Monitoring Board (DSMB) prior to enrollment of subjects from part B. In case non-safe or futile treatment groups (i.e., doses that do not induce considerable IgG responses) are identified during the interim analysis, these treatment groups will be dropped and not carried forward to part B.

[0240] Part B of the study will seek dose confirmation in the elderly population. Accordingly, Part B will be conducted in 5 treatment groups of 20 elderly healthy subjects per group. Vaccination schedule Day 0, 7 and 21 will be applied. This study design will allow to compare dose responses in both adults and elderly. The latter age group will be the major target population for a *C. difficile* vaccine, representing the most vulnerable population for the two target indications in the development pathway of a *C. difficile* vaccine, i.e. prevention of recurrent *C. difficile* diarrhea and prevention of primary *C. difficile* infection in an age-based or age-risk based preventive vaccination approach. However, elderly subjects might be less responsive to vaccination than young adults; thus, dose confirmation in the elderly target population from an early development stage on is required. An interim analysis after all adults from part A have been vaccinated will allow to drop non-safe or doses/formulations which do not induce considerable IgG responses in adults in order to mitigate the risk of exposing subjects in the elderly group to potentially unsafe or futile doses (e.g. lowest dose) and/or formulations (e.g. non-adjuvanted formulation) of the vaccine.

[0241] The C-TAB.G5.1 vaccine is an aqueous solution of C-TAB.G5.1 in 20 mM L-Histidine, 75mM NaCl, 5% Sucrose, 0.025% Tween®80; pH6.5 produced by standard methods.

SEQUENCES:

Name	SEQ ID NOs:	Sequences
C-TAB.G5 (nucleic acid sequence)	1	ATGGTAACAGGAGTATTTAAAGGACCTAATGGATTTGAGTATTTTGC ACCTGCTAATACTCACAATAATAACATAGAAGGTCAGGCTATAGTTT ACCAGAACAAATTCTTAACTTTGAATGGCAAAAAATATTATTTTGAT AATGACTCAAAAGCAGTTACTGGATGGCAAACCATTGATGGTAAAA AATATTACTTTAATCTTAACACTGCTGAAGCAGCTACTGGATGGCAA ACTATTGATGGTAAAAAATATTACTTTAATCTTAACACTGCTGAAGC AGCTACTGGATGGCAAACTATTGATGGTAAAAAATATTACTTTAATA CTAACACTTTCATAGCCTCAACTGGTTATACAAGTATTAATGGTAAA CATTTTTATTTTAATACTGATGGTATTATGCAGATAGGAGTGTTTAAA GGACCTAATGGATTTGAATACTTTGCACCTGCTAATACTCATAATAA CAACATAGAAGGTCAAGCTATACTTTACCAAAATAAATTCTTAACTT TGAATGGTAAAAAATATTACTTTGGTAGTGACTCAAAAGCAGTTACC GGATTGCGAACTATTGATGGTAAAAAATATTACTTTAATACTAACAC TGCTGTTGCAGTTACTGGATGGCAAACTATTAATGGTAAAAAATACT ACTTTAATACTAACACTTCTATAGCTTCAACTGGTTATACAATTATTA GTGGTAAACATTTTTATTTTAATACTGATGGTATTATGCAGATAGGAG TGTTTAAAGGACCTGATGGATTTGAATACTTTGCACCTGCTAATACA GATGCTAACAATATAGAAGGTCAAGCTATACGTTATCAAAATAGATT CCTATATTTACATGACAATATATATTATTTTGGTAATAATTCAAAAGC AGCTACTGGTTGGGTAACTATTGATGGTAATAGATATTACTTCGAGC CTAATACAGCTATGGGTGCGAATGGTTATAAAACTATTGATAATAAA AATTTTTACTTTAGAAATGGTTTACCTCAGATAGGAGTGTTTAAAGG GTCTAATGGATTTGAATACTTTGCACCTGCTAATACGGATGCTAACA ATATAGAAGGTCAAGCTATACGTTATCAAAATAGATTCCTACATTTA CTTGGAAAAATATATTACTTTGGTAATAATTCAAAAGCAGTTACTGG ATGGCAAACTATTAATGGTAAAGTATATTACTTTATGCCTGATACTG CTATGGCTGCAGCTGGTGGACTTTTCGAGATTGATGGTGTTATATATT

(continued)

Name	SEQ ID NOs:	Sequences
		TCTTTGGTGTTGATGGAGTAAAAGCCCCTGGGATATATGGCAGATCT ATGCATAATTTGATAACTGGATTTGTGACTGTAGGCGATGATAAATA CTACTTTAATCCAATTAATGGTGGAGCTGCTTCAATTGGAGAGACAA TAATTGATGACAAAAATTATTATTTCAACCAAAGTGGAGTGTTACAA ACAGGTGTATTTAGTACAGAAGATGGATTTAAATATTTTGCCCCAGC TAATACACTTGATGAAAACCTAGAAGGAGAAGCAATTGATTTTACTG GAAAATTAATTATTGACGAAAATATTTATTATTTTGATGATAATTATA GAGGAGCTGTAGAATGGAAAGAATTAGATGGTGAAATGCACTATTTT AGCCCAGAAACAGGTAAAGCTTTTAAAGGTCTAAATCAAATAGGTG ATTATAAATACTATTTCAATTCTGATGGAGTTATGCAAAAAGGATTT GTTAGTATAAATGATAATAAACACTATTTTGATGATTCTGGTGTTATG AAAGTAGGTTACACTGAAATAGATGGCAAGCATTTCTACTTTGCTGA AAACGGAGAAATGCAAATAGGAGTATTTAATACAGAAGATGGATTT AAATATTTTGCTCATCATAATGAAGATTTAGGAAATGAAGAAGGTGA AGAAATCTCATATTCTGGTATATTAAATTTCAATAATAAAATTTACTA TTTTGATGATTCATTTACAGCTGTAGTTGGATGGAAAGATTTAGAGG ATGGTTCAAAGTATTATTTTGATGAAGATACAGCAGAAGCATATATA GGTTTGTCATTAATAAATGATGGTCAATATTATTTTAATGATGATGGA ATTATGCAAGTTGGATTTGTCACTATAAATGATAAAGTCTTCTACTTC TCTGACTCTGGAATTATAGAATCTGGAGTACAAAACATAGATGACAA TTATTTCTATATAGATGATAATGGTATAGTTCAAATTGGTGTATTTGA TACTTCAGATGGATATAAATATTTTGCACCTGCTAATACTGTAAATG ATAATATTTACGGACAAGCAGTTGAATATAGTGGTTTAGTTAGAGTT GGGGAAGATGTATATTATTTTGGAGAAACATATACAATTGAGACTGG ATGGATATATGATATGGAAAATGAAAGTGATAAATATTATTTCAATC CAGAAACTAAAAAAGCATGCAAAGGTATTAATTTAATTGATGATATA AAATATTATTTTGATGAGAAGGGCATAATGAGAACGGGTCTTATATC ATTTGAAAATAATAATTATTACTTTAATGAGAATGGTGAAATGCAAT TTGGTTATATAAATATAGAAGATAAGATGTTCTATTTTGGTGAAGAT GGTGTCATGCAGATTGGAGTATTTAATACACCAGATGGATTTAAATA CTTTGCACATCAAAATACTTTGGATGAGAATTTTGAGGGAGAATCAA TAAACTATACTGGTTGGTTAGATTTAGATGAAAAGAGATATTATTTT ACAGATGAATATATTGCAGCAACTGGTTCAGTTATTATTGATGGTGA GGAGTATTATTTTGATCCTGATACAGCTCAATTAGTGATTAGTGAATA G

(continued)

Name	SEQ ID NOs:	Sequences
C-TAB.G5 (amino acid sequence)	2	MVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDN DSKAVTGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAA TGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPN GFEYFAPANTHNNNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTI DGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYF NTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIY YFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLP QIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNS KAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGI YGRSMHNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGV LQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYR GAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFV SINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKY FAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSK YYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGII ESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAV EYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGI NLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYF GEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE
C-TAB.G5.1	3	CCATGGTTACAGGTGTTTTCAAAGGTCCGAACGGCTTTGAATATTTTG

(continued)

Name	SEQ ID NOs:	Sequences
(nucleic acid sequence)		CACCGGCAAATACCCACAATAATAATATTGAAGGCCAGGCCATCGTG TATCAGAATAAATTTCTGACCCTGAACGGCAAAAAATACTATTTCGA TAACGATAGCAAAGCAGTTACCGGTTGGCAAACCATTGATGGCAAA AAATATTACTTCAACCTGAATACCGCAGAAGCAGCAACCGGCTGGCA GACGATCGACGGTAAAAAGTACTATTTTAACCTGAACACAGCCGAA GCCGCTACAGGCTGGCAGACAATAGATGGGAAGAAGTATTATTTTAA TACCAATACCTTTATTGCCAGCACCGGCTATACCAGCATTAATGGCA AACACTTCTATTTTAACACCGATGGTATTATGCAGATCGGTGTGTTTA AGGGCCCTAATGGTTTTGAGTACTTCGCTCCGGCTAATACCGATGCA AATAACATCGAAGGTCAGGCAATTCTGTACCAGAACAAATTTTTAAC GCTGAACGGTAAGAAATATTACTTTGGTAGCGATTCAAAAGCCGTTA CCGGTCTGCGTACGATCGACGGCAAGAAATATTATTTCAATACAAAC ACCGCAGTTGCCGTGACAGGTTGGCAGACGATAAATGGTAAGAAGT ACTACTTCAACACCAATACCAGCATTGCAAGTACCGGTTATACCATT ATCAGCGGCAAACACTTTTACTTCAATACAGACGGCATTATGCAGAT TGGCGTTTTCAAAGGTCCGGATGGTTTCGAGTACTTTGCCCCTGCAA ATACAGATGCAAACAATATTGAGGGACAGGCAATTCGCTATCAGAA TCGTTTTCTGTATCTGCACGATAACATCTATTACTTCGGCAATAATTC AAAAGCAGCCACCGGTTGGGTTACAATTGATGGTAATCGTTATTACT TTGAGCCGAATACCGCAATGGGTGCAAATGGTTATAAAACCATCGAT AACAAAAATTTTTATTTCCGCAACGGTCTGCCGCAGATTGGTGTTTTT AAGGGTAGCAATGGCTTCGAGTATTTTGCGCCAGCCAACACCGATGC CAACAACATTGAAGGCCAAGCGATTCGTTATCAAAACCGCTTTCTGC ATCTGCTGGGCAAAATTTATTACTTTGGCAACAATAGCAAAGCGGTG ACGGGCTGGCAAACCATTAACGGTAAAGTTTATTATTTCATGCCGGA TACCGCTATGGCAGCAGCCGGTGGTCTGTTTGAAATTGATGGCGTGA TTTATTTTTTTGGCGTGGATGGTGTTAAAGCACCGGGTATTTATGGTC GTAGCATGCATAATCTGATTACCGGTTTTGTTACCGTGGGCGACGAT AAATACTACTTTAATCCGATTAATGGTGGTGCAGCAAGCATTGGTGA AACCATTATCGATGACAAAAACTATTATTTTAACCAGAGCGGTGTTC TGCAGACAGGTGTTTTTAGCACCGAAGATGGCTTCAAATATTTTGCT CCTGCGAATACACTGGATGAAAATCTGGAAGGTGAAGCAATTGATTT TACCGGCAAACTGATCATCGACGAGAACATCTACTATTTTGATGATA ATTATCGCGGTGCCGTGGAATGGAAAGAACTGGATGGTGAAATGCA CTATTTTAGTCCGGAAACCGGTAAAGCCTTTAAAGGTCTGAATCAGA TCGGCGATTACAAGTATTACTTTAATTCAGATGGCGTGATGCAGAAA GGCTTTGTGAGCATTAACGACAACAAACACTATTTTGACGACAGCGG TGTGATGAAAGTGGGTTATACCGAAATCGACGGGAAACATTTTTATT TTGCCGAAAACGGCGAAATGCAGATTGGAGTATTTAATACCGAGGA CGGCTTTAAATACTTTGCCCATCATAATGAAGATCTGGGTAATGAAG AAGGCGAAGAAATTAGCTATAGCGGCATTCTGAATTTTAATAACAAG ATCTATTATTTCGATGATAGCTTCACCGCAGTTGTTGGTTGGAAAGAT CTGGAAGATGGCAGCAAATATTATTTTGATGAAGATACCGCAGAGGC CTATATTGGTCTGAGCCTGATTAATGATGGCCAGTATTATTTCAACGA TGATGGTATCATGCAGGTTGGTTTTGTGACCATCAACGATAAAGTGT TCTATTTCAGCGATAGCGGCATTATTGAAAGCGGTGTTCAGAACATC GACGATAACTATTTCTACATCGATGATAACGGTATTGTTCAGATTGG CGTGTTTGATACCTCCGATGGTTATAAATATTTCGCACCAGCCAATAC CGTGAACGATAATATTTATGGTCAGGCAGTTGAATATTCAGGTCTGG TTCGTGTTGGCGAAGATGTTTATTATTTTGGCGAAACCTATACCATTG AAACCGGCTGGATCTATGATATGGAAAACGAGAGCGACAAGTACTA TTTCAATCCGGAAACGAAAAAAGCCTGCAAAGGCATTAATCTGATCG ACGATATTAAGTACTACTTTGACGAAAAAGGCATTATGCGTACCGGT CTGATTAGCTTTGAGAACAACAACTATTACTTCAATGAGAACGGTGA GATGCAGTTTGGCTATATCAACATCGAGGACAAAATGTTTTATTTTG GTGAGGACGGTGTGATGCAGATAGGGGTTTTTAATACACCGGATGGG

(continued)

Name	SEQ ID NOs:	Sequences
		TTTAAGTATTTTGCACATCAGAACACCCTGGATGAAAACTTTGAAGG CGAAAGCATTAATTATACCGGTTGGCTGGATCTGGATGAGAAACGTT ATTATTTCACCGACGAATACATTGCAGCAACCGGTAGCGTTATTATT GATGGTGAGGAATATTACTTCGATCCGGATACAGCACAGCTGGTTAT TAGCGAATAACTCGAG
C-TAB.G5.1 (amino acid sequence)	4	VTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDND SKAVTGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAAT GWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPNG FEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTID GKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFN TDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYY FGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLPQ IGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNSK AVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGIY GRSMHNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVL QTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRG AVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSI NDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYF AHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKY YFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIE SGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVE YSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGI NLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYF GEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE

(continued)

Name	SEQ ID NOs:	Sequences
Nucleic acid sequence of *trdA (strain 630)*	5	atgtctttaatatctaaagaagagttaataaaactcgcatatagcattagaccaagagaaaatgagtataaaactatac taactaatttagacgaatataataagttaactacaaacaataatgaaaataaatatttacaattaaaaaaactaaatgaa tcaattgatgtttttatgaataaatataaaacttcaagcagaaatagagcactctctaatctaaaaaaagatatattaaaa gaagtaattcttattaaaaattccaatacaagccctgtagaaaaaaatttacattttgtatggataggtggagaagtcag tgatattgctcttgaatacataaaacaatgggctgatattaatgcagaatataatattaaactgtggtatgatagtgaag cattcttagtaaatacactaaaaaaggctatagttgaatcttctaccactgaagcattacagctactagaggaagagat tcaaaatcctcaatttgataatatgaaattttacaaaaaaaggatggaatttatatatgatagacaaaaaaggtttataaa ttattataaatctcaaatcaataaacctacagtacctacaatagatgatattataaagtctcatctagtatctgaatataat agagatgaaactgtattagaatcatatagaacaaattctttgagaaaaataaatagtaatcatgggatagatatcagg gctaatagtttgtttacagaacaagagttattaaatatttatagtcaggagttgttaaatcgtggaaatttagctgcagca tctgacatagtaagattattagccctaaaaaattttggcggagtatatttagatgttgatatgcttccaggtattcactctg atttatttaaaacaatatctagacctagctctattggactagaccgttgggaaatgataaaattagaggctattatgaag tataaaaaatatataaataattatacatcagaaaacttttgataaacttgatcaacaattaaaagataatttttaaactcatta tagaaagtaaaagtgaaaaatctgagatattttctaaattagaaaatttaaatgtatctgatcttgaaattaaaatagcttt cgctttaggcagtgttataaatcaagccttgatatcaaaacaaggttcatatcttactaacctagtaatagaacaagtaa aaaatagatatcaatttttaaaccaacaccttaacccagccatagagtctgataataacttcacagatactactaaaatt tttcatgattcattatttaattcagctaccgcagaaaactctatgtttttaacaaaaatagcaccatacttacaagtaggttt tatgccagaagctcgctccacaataagtttaagtggtccaggagcttatgcgtcagcttactatgatttcataaatttac aagaaaatactatagaaaaaactttaaaagcatcagatttaatagaatttaaattcccagaaaataatctatctcaattg acagaacaagaaataaatagtctatggagctttgatcaagcaagtgcaaaatatcaatttgagaaatatgtaagagat tatactggtggatctctttctgaagacaatggggtagactttaataaaaatactgccctcgacaaaaactatttattaaat aataaaattccatcaaacaatgtagaagaagctggaagtaaaaattatgttcattatatcatacagttacaaggagatg atataagttatgaagcaacatgcaatttatttttctaaaaatcctaaaaatagtattattatacaacgaaatatgaatgaaa gtgcaaaaagctacttttttaagtgatgatggagaatctatttagaattaaataaatataggatacctgaaagattaaaa aataaggaaaaagtaaaagtaaccttattggacatggtaaagatgaattcaacacaagcgaatttgctagattaagt gtagattcactttccaatgagataagttcatttttagataccataaaattagatatatcacctaaaaatgtagaagtaaac ttacttggatgtaatatgtttagttatgattttaatgttgaagaaacttatcctgggaagttgctattaagtattatggacaa aattacttccactttacctgatgtaaataaaaattctattactataggagcaaatcaatatgaagtaagaattaatagtga gggaagaaaagaacttctggctcactcaggtaaatggataaataaagaagaagctattatgagcgatttatctagta aagaatacatttttttttgattctatagataataagctaaaagcaaagtccaagaatattccaggattagcatcaatatca

(continued)

Name	SEQ ID NOs:	Sequences
		gaagatataaaaacattattacttgatgcaagtgttagtcctgatacaaaatttatttttaaataatcttaagcttaatattga atcttctattggtgattacatttattatgaaaaattagagcctgttaaaaatataattcacaattctatagatgatttaataga tgagttcaatctacttgaaaatgtatctgatgaattatatgaattaaaaaaattaaataatctagatgagaagtatttaata tcttttgaagatatctcaaaaaataattcaacttactctgtaagatttattaacaaaagtaatggtgagtcagtttatgtag aaacagaaaaagaaatttttttcaaaatatagcgaacatattacaaaagaaataagtactatataaagaatagtataattac agatgttaatggtaatttattggataatatacagttagatcatacttctcaagttaatacattaaacgcagcattctttattc aatcattaatagattatagtagcaataaagatgtactgaatgatttaagtacctcagttaaggttcaacttttatgctcaac tatttagtacaggtttaaatactatatatgactctatccaattagtaaatttaatatcaaatgcagtaaatgatactataaat gtactacctacaataacagaggggatacctattgtatctactatattagacggaataaacttaggtgcagcaattaag gaattactagacgaacatgacccattactaaaaaaagaattagaagctaaggtgggtgttttagcaataaatatgtca ttatctatagctgcaactgtagcttcaattgttggaataggtgctgaagttactattttcttattacctatagctggtatatct gcaggaataccttcattagttaataatgaattaatattgcatgataaggcaacttcagtggtaaactattttaatcatttgt ctgaatctaaaaaatatggccctcttaaaacagaagatgataaaattttagttcctattgatgatttagtaatatcagaaa tagattttaataataattcgataaaactaggaacatgtaatatattagcaatggaggggggatcaggacacacagtga ctggtaatatagatcacttttctcatctccatctataagttctcatattccttcattatcaatttattctgcaataggtataga aacagaaaatctagatttttcaaaaaaaataatgatgttacctaatgctccttcaagagtgttttggtgggaaactgga gcagttccaggtttaagatcattggaaaatgacggaactagattacttgattcaataagagatttatacccaggtaaat tttactggagattctatgctttttcgattatgcaataactacattaaaaccagtttatgaagacactaatattaaaattaaa ctagataaagatactagaaacttcataatgccaactataactactaacgaaattagaaacaaattatcttattcatttgat ggagcaggaggaacttactctttattattatcttcatatccaatatcaacgaatataaatttatctaaagatgatttatgga tatttaatattgataatgaagtaagagaaatatctatagaaaatggtactattaaaaaaggaaagttaataaaagatgttt taagtaaaattgatataaataaaaataaacttattataggcaatcaaacaatagatttttcaggcgatatagataataaa gatagatatatattcttgacttgtgagttagatgataaaattagtttaataatagaaataaatcttgttgcaaaatcttatag tttgttattgtctggggataaaaattatttgatatccaatttatctaatattattgagaaaatcaatactttaggcctagatag taaaaatatagcgtacaattacactgatgaatctaataataaatatttttggagctatatctaaaacaagtcaaaaaagca taatacattataaaaaagacagtaaaaatatattagaatttttataatgacagtacattagaatttaacagtaaagatttttat tgctgaagatataaatgtatttatgaaagatgatattaatactataacaggaaaatactatgttgataataatactgataa aagtatagatttctctatttctttagttagtaaaaatcaagtaaaagtaaatggattatatttaaatgaatccgtatactcat cttaccttgattttgtgaaaaattcagatggacaccataatacttctaattttatgaatttatttttggacaatataagtttctg gaaattgtttgggtttgaaaatataaattttgtaatcgataaatactttacccttgttggtaaaactaatcttggatatgtag aatttatttgtgacaatataaaaatatagatatatattttggtgaatggaaaacatcgtcatctaaaagcactatatttag cggaaatggtagaaatgttgtagtagagcctatatataatcctgatacgggtgaagatatatctacttcactagatttttt cctatgaacctctctatggaatagatagatatatcaataaagtattgatagcacctgatttatatacaagtttaataaatat taataccaattattattcaaatgagtactaccctgagattatagttcttaacccaaatacattccacaaaaaagtaaatat aaatttagatagttcttcttttgagtataaatggtctacagaaggaagtgactttatttttagttagatacttagaagaaagt aataaaaaaatattacaaaaaataagaatcaaaggtatcttatctaatactcaatcatttaataaaatgagtatagatttt aaagatattaaaaaaactatcattaggatatataatgagtaattttaaatcatttaattctgaaaatgaattagatagagatc atttaggatttaaaataatagataataaaacttattactatgatgaagatagtaaattagttaaaggattaatcaatataaa taattcattattctatttttgatcctatagaatttaacttagtaactggatggcaaactatcaatggtaaaaaatattattttga tataaatactggagcagctttaattagttatataaaattattaatggtaaacacttttatttttaataatgatggtgtgatgcagt tgggagtatttaaaggacctgatggatttgaatattttgcacctgccaatactcaaaataataacatagaaggtcaggc tatagtttatcaaagtaaattcttaactttgaatggcaaaaaatattattttgataatgactcaaaagcagtcactggatg gagaattattaacaatgagaaatattacttttaatcctaataatgctattgctgcagtcggattgcaagtaattgacaataa taagtattatttcaatcctgacactgctatcatctcaaaaggttggcagactgttaatggtagtagatactactttgatact gataccgctattgcctttaatggttataaaactattgatggtaaacacttttattttgatagtgattgtgtagtgaaaatag gtgtgtttagtacctctaatggatttgaatattttgcacctgctaatacttataataataacatagaaggtcaggctatagt ttatcaaagtaaattcttaactttgaatggtaaaaaatattactttgataataactcaaaagcagttaccggatggcaaa ctattgatagtaaaaaatattactttaatactaacactgctgaagcagctactggatggcaaactattgatggtaaaaa atattacttttaatactaacactgctgaagcagctactggatggcaaactattgatggtaaaaaatattacttttaatactaa cactgctatagcttcaactggttatacaattattaatggtaaacatttttatttttaatactgatggtattatgcagataggag tgtttaaaggacctaatggatttgaatattttgcacctgctaatacggatgctaacaacatagaaggtcaagctatactt taccaaaatgaattcttaactttgaatggtaaaaaatattactttggtagtgactcaaaagcagttactggatggagaat tattaacaataagaaatattacttttaatcctaataatgctattgctgcaattcatctatgcactataaataatgacaagtatt actttagttatgatggaattcttcaaaatggatatattactattgaaagaaataatttctatttttgatgctaataatgaatcta aaatggtaacaggagtatttaaaggacctaatggatttgagtattttgcacctgctaatactcacaataataacataga aggtcaggctatagtttaccagaacaaattcttaactttgaatggcaaaaaatattattttgataatgactcaaaagcag

(continued)

Name	SEQ ID NOs:	Sequences
		ttactggatggcaaaccattgatggtaaaaaatattactttaatcttaacactgctgaagcagctactggatggcaaac tattgatggtaaaaaatattactttaatcttaacactgctgaagcagctactggatggcaaactattgatggtaaaaaat attactttaatactaacactttcatagcctcaactggttatacaagtattaatggtaaacatttttattttaatactgatggta ttatgcagataggagtgtttaaaggacctaatggatttgaatactttgcacctgctaatactcataataataacatagaa ggtcaagctatactttaccaaaataaattcttaactttgaatggtaaaaaatattactttggtagtgactcaaaagcagtt accggattgcgaactattgatggtaaaaaatattactttaatactaacactgctgttgcagttactggatggcaaactat taatggtaaaaaatactactttaatactaacacttctatagcttcaactggttatacaattattagtggtaaacatttttatttt aatactgatggtattatgcagataggagtgtttaaaggacctgatggatttgaatactttgcacctgctaatacagatg ctaacaatatagaaggtcaagctatacgttatcaaaatagattcctatatttacatgacaatatatattatttttggtaataat tcaaaagcagctactggttgggtaactattgatggtaatagatattacttcgagcctaatacagctatgggtgcgaatg gttataaaactattgataataaaaatttttactttagaaatggtttacctcagataggagtgtttaaagggtctaatggattt gaatactttgcacctgctaatacggatgctaacaatatagaaggtcaagctatacgttatcaaaatagattcctacattt acttggaaaaatatattactttggtaataattcaaaagcagttactggatggcaaactattaatggtaaagtatattacttt atgcctgatactgctatggctgcagctggtggacttttcgagattgatggtgttatatatttctttggtgttgatggagta aaagcccctgggatatatggctaa

(continued)

Name	SEQ ID NOs:	Sequences
Amino acid sequence of *trdA (strain 630)*	6	MSLISKEELIKLAYSIRPRENEYKTILTNLDEYNKLTTNNNENKYLQLKK LNESIDVFMNKYKTSSRNRALSNLKKDILKEVILIKNSNTSPVEKNLHFV WIGGEVSDIALEYIKQWADINAEYNIKLWYDSEAFLVNTLKKAIVESSTT EALQLLEEEIQNPQFDNMKFYKKRMEFIYDRQKRFINYYKSQINKPTVPT IDDIIKSHLVSEYNRDETVLESYRTNSLRKINSNHGIDIRANSLFTEQELLN IYSQELLNRGNLAAASDIVRLLALKNFGGVYLDVDMLPGIHSDLFKTISR PSSIGLDRWEMIKLEAIMKYKKYINNYTSENFDKLDQQLKDNFKLIIESK SEKSEIFSKLENLNVSDLEIKIAFALGSVINQALISKQGSYLTNLVIEQVKN RYQFLNQHLNPAIESDNNFTDTTKIFHDSLFNSATAENSMFLTKIAPYLQ VGFMPEARSTISLSGPGAYASAYYDFINLQENTIEKTLKASDLIEFKFPEN NLSQLTEQEINSLWSFDQASAKYQFEKYVRDYTGGSLSEDNGVDFNKN TALDKNYLLNNKIPSNNVEEAGSKNYVHYIIQLQGDDISYEATCNLFSK NPKNSIIIQRNMNESAKSYFLSDDGESILELNKYRIPERLKNKEKVKVTFI GHGKDEFNTSEFARLSVDSLSNEISSFLDTIKLDISPKNVEVNLLGCNMFS YDFNVEETYPGKLLLSIMDKITSTLPDVNKNSITIGANQYEVRINSEGRKE LLAHSGKWINKEEAIMSDLSSKEYIFFDSIDNKLKAKSKNIPGLASISEDI KTLLLDASVSPDTKFILNNLKLNIESSIGDYIYYEKLEPVKNIIHNSIDDLI DEFNLLENVSDELYELKKLNNLDEKYLISFEDISKNNSTYSVRFINKSNG ESVYVETEKEIFSKYSEHITKEISTIKNSIITDVNGNLLDNIQLDHTSQVNT LNAAFFIQSLIDYSSNKDVLNDLSTSVKVQLYAQLFSTGLNTIYDSIQLV NLISNAVNDTINVLPTITEGIPIVSTILDGINLGAAIKELLDEHDPLLKKELE AKVGVLAINMSLSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELIL HDKATSVVNYFNHLSESKKYGPLKTEDDKILVPIDDLVISEIDFNNNSIKL GTCNILAMEGGSGHTVTGNIDHFFSSPSISSHIPSLSIYSAIGIETENLDFSK KIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLYPGKFYWR FYAFFDYAITTLKPVYEDTNIKIKLDKDTRNFIMPTITTNEIRNKLSYSFD GAGGTYSLLLSSYPISTNINLSKDDLWIFNIDNEVREISIENGTIKKGKLIK DVLSKIDINKNKLIIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIIEINLVA KSYSLLLSGDKNYLISNLSNIIEKINTLGLDSKNIAYNYTDESNNKYFGAI SKTSQKSIIHYKKDSKNILEFYNDSTLEFNSKDFIAEDINVFMKDDINTITG KYYVDNNTDKSIDFSISLVSKNQVKVNGLYLNESVYSSYLDFVKNSDGH HNTSNFMNLFLDNISFWKLFGFENINFVIDKYFTLVGKTNLGYVEFICDN NKNIDIYFGEWKTSSSKSTIFSGNGRNVVVEPIYNPDTGEDISTSLDFSYE PLYGIDRYINKVLIAPDLYTSLININTNYYSNEYYPEIIVLNPNTFHKKVNI NLDSSSFEYKWSTEGSDFILVRYLEESNKKILQKIRIKGILSNTQSFNKMSI DFKDIKKLSLGYIMSNFKSFNSENELDRDHLGFKIIDNKTYYYDEDSKLV KGLININNSLFYFDPIEFNLVTGWQTINGKKYYFDINTGAALISYKIINGK HFYFNNDGVMQLGVFKGPDGFEYFAPANTQNNNIEGQAIVYQSKFLTL NGKKYYFDNDSKAVTGWRIINNEKYYFNPNNAIAAVGLQVIDNNKYYF NPDTAIISKGWQTVNGSRYYFDTDTAIAFNGYKTIDGKHFYFDSDCVVK IGVFSTSNGFEYFAPANTYNNNIEGQAIVYQSKFLTLNGKKYYFDNNSK

(continued)

Name	SEQ ID NOs:	Sequences
		AVTGWQTIDSKKYYFNTNTAEAATGWQTIDGKKYYFNTNTAEAATGW QTIDGKKYYFNTNTAIASTGYTIINGKHFYFNTDGIMQIGVFKGPNGFEY FAPANTDANNIEGQAILYQNEFLTLNGKKYYFGSDSKAVTGWRIINNKK YYFNPNNAIAAIHLCTINNDKYYFSYDGILQNGYITIERNNFYFDANNES KMVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFD NDSKAVTGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEA ATGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGP NGFEYFAPANTHNNNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRT IDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYF NTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIY YFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLP QIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNS KAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGI YG

(continued)

Name	SEQ ID NOs:	Sequences
Nucleic acid sequence of *trdB(strain 630)*	7	atgagtttagttaatagaaaacagttagaaaaaatggcaaatgtaagatttcgtactcaagaagatgaatatgttgcaa tattggatgctttagaagaatatcataatatgtcagagaatactgtagtcgaaaaatatttaaaattaaaagatataaata gtttaacagatatttatatagatacatataaaaaatctggtagaaataaagccttaaaaaaatttaaggaatatctagtta cagaagtattagagctaaagaataataatttaactccagttgagaaaaatttacattttgtttggattggaggtcaaata aatgacactgctattaattatataaatcaatggaaagatgtaaatagtgattataatgttaatgttttttatgatagtaatgc atttttgataaacacattgaaaaaaactgtagtagaatcagcaataaatgatacacttgaatcatttagagaaaacttaa atgaccctagatttgactataataaattcttcagaaaacgtatggaaataatttatgataaacagaaaaatttcataaact actataaagctcaaagagaagaaaatcctgaacttataattgatgatattgtaaagacatatctttcaaatgagtattca aaggagatagatgaacttaatacctatattgaagaatccttaaataaaattacacagaatagtggaaatgatgttagaa actttgaagaatttaaaaatggagagtcattcaacttatatgaacaagagttggtagaaaggtggaatttagctgctgc ttctgacatattaagaatatctgcattaaaagaaattggtggtatgtatttagatgttgatatgttaccaggaatacaacc agacttatttgagtctatagagaaacctagttcagtaacagtggatttttgggaaatgacaaagttagaagctataatg aaatacaaagaatatataccagaatatacctcagaacattttgacatgttagacgaagaagttcaaagtagttttgaat ctgttctagcttctaagtcagataaatcagaaatattctcatcacttggtgatatggaggcatcaccactagaagttaaa attgcatttaatagtaagggtattataaatcaagggctaatttctgtgaaagactcatattgtagcaatttaatagtaaaa caaatcgagaatagatataaaatattgaataatagtttaaatccagctattagcgaggataatgattttaatactacaac gaataccttattgatagtataatggctgaagctaatgcagataatggtagatttatgatggaactaggaaagtatttaa gagttggtttcttcccagatgttaaaactactattaacttaagtggccctgaagcatatgcggcagcttatcaagatttat taatgtttaaagaaggcagtatgaatatccatttgatagaagctgatttaagaaacttgaaatctctaaaactaatattt ctcaatcaactgaacaagaaatggctagcttatggtcatttgacgatgcaagagctaaagctcaatttgaagaatata aaaggaattatttttgaaggttctcttggtgaagatgataatcttgatttttctcaaaatatagtagttgacaaggagtatct tttagaaaaaatatcttcattagcaagaagttcagagagaggatatatacactatattgttcagttacaaggagataaa attagttatgaagcagcatgtaacttatttgcaaagactccttatgatagtgtactgtttcagaaaaatatagaagattca gaaattgcatattattataatcctggagatggtgaaatacaagaaatagacaagtataaaattccaagtataatttctga tagacctaagattaaattaacatttattggtcatggtaaagatgaatttaatactgatatatttgcaggttttgatgtagatt cattatccacagaaatagaagcagcaatagatttagctaaagaggatatttctcctaagtcaatagaaataaatttatta ggatgtaatatgtttagctactctatcaacgtagaggagacttatcctggaaaattattacttaaagttaaagataaaat atcagaattaatgccatctataagtcaagactctattatagtaagtgcaaatcaatatgaagttagaataaatagtgaa ggaagaagagaattattggatcattctggtgaatggataaataaagaagaaagtattataaaggatatttcatcaaaa gaatatatatcatttaatcctaaagaaaataaaattacagtaaaatctaaaaatttacctgagctatctacattattacaa gaaattagaaataattctaattcaagtgatattgaactagaagaaaaagtaatgttaacagaatgtgagataaatgttat ttcaaatatagatacgcaaattgttgaggaaaggattgaagaagctaagaatttaacttctgactctattaattatataaa agatgaatttaaactaatagaatctatttctgatgcactatgtgacttaaaacaacagaatgaattagaagattctcatttt atatcttttgaggacatatcagagactgatgagggatttagtataagatttattaataaagaaactggagaatctatattt gtagaaactgaaaaaacaatattctctgaatatgctaatcatataactgaagagatttctaagataaaaggtactatatt tgatactgtaaatggtaagttagtaaaaaaagtaaatttagatactacacacgaagtaaatactttaaatgctgcatttt tatacaatcattaatagaatataatagttctaaagaatctcttagtaatttaagtgtagcaatgaaagtccaagtttacgct caattatttagtactggtttaaatactattacagatgcagccaaagttgttgaattagtatcaactgcattagatgaaact atagacttacttcctacattatctgaaggattacctataattgcaactattatagatggtgtaagtttaggtgcagcaatc aaagagctaagtgaaacgagtgacccattattaagacaagaaatagaagctaagataggtataatggcagtaaattt aacaacagctacaactgcaatcattacttcatctttggggatagctagtggatttagtatacttttagttccttagcagg aatttcagcaggtataccaagcttagtaaacaatgaacttgtacttcgagataaggcaacaaaggttgtagattatttta

(continued)

Name	SEQ ID NOs:	Sequences
		aacatgtttcattagttgaaactgaaggagtatttacttttattagatgataaaataatgatgccacaagatgatttagtga tatcagaaatagattttaataataattcaatagttttaggtaaatgtgaaatctggagaatggaaggtggttcaggtcat actgtaactgatgatatagatcacttcttttcagcaccatcaataacatatagagagccacacttatctatatatgacgta ttggaagtacaaaaagaagaacttgatttgtcaaaagatttaatggtattacctaatgctccaaatagagtatttgcttg ggaaacaggatggacaccaggtttaagaagcttagaaaatgatggcacaaaactgttagaccgtataagagataa ctatgaaggtgagttttattggagatattttgcttttatagctgatgctttaataacaacattaaaaccaagatatgaagat actaatataagaataaatttagatagtaatactagaagttttatagttccaataataactacagaatatataagagaaaa attatcatattctttctatggttcaggaggaacttatgcattgtctctttctcaatataatatgggtataaatatagaattaag tgaaagtgatgtttggattatagatgttgataatgttgtgagagatgtaactatagaatctgataaaattaaaaaaggtg atttaatagaaggtattttatctacactaagtattgaagagaataaaattatcttaaatagccatgagattaatttttctggt gaggtaaatggaagtaatggatttgtttctttaacattttcaattttagaaggaataaatgcaattatagaagttgatttatt atctaaatcatataaattacttatttctggcgaattaaaaatattgatgttaaattcaaatcatattcaacagaaaatagatt atataggattcaatagcgaattacagaaaaatataccatatagctttgtagatagtgaaggaaaagagaatggttttat taatggttcaacaaaagaaggtttatttgtatctgaattacctgatgtagttcttataagtaaggtttatatggatgatagt aagccttcatttggatattatagtaataatttgaaagatgtcaaagttataactaaagataatgttaatatattaacaggtt attatcttaaggatgatataaaaatctctctttctttgactctacaagatgaaaaaactataaagttaaatagtgtgcattt agatgaaagtggagtagctgagattttgaagttcatgaatagaaaaggtaatacaaatacttcagattctttaatgagc tttttagaaagtatgaatataaaaagtattttcgttaatttcttacaatctaatattaagtttatattagatgctaattttataat aagtggtactacttctattggccaatttgagtttatttgtgatgaaaatgataatatacaaccatatttcattaagtttaata cactagaaactaattatacttttatatgtaggaaatagacaaaatatgatagtggaaccaaattatgatttagatgattctg gagatatatcttcaactgttatcaatttctctcaaaagtatctttatggaatagacagttgtgttaataaagttgtaatttca ccaaatatttatacagatgaaataaatataacgcctgtatatgaaacaaataatacttatccagaagttattgtattagat gcaaattatataaatgaaaaaataaatgttaatatcaatgatctatctatacgatatgtatggagtaatgatggtaatgat tttattcttatgtcaactagtgaagaaaataaggtgtcacaagttaaaataagattcgttaatgtttttaaagataagactt tggcaaataagctatcttttaactttagtgataaacaagatgtacctgtaagtgaaataatcttatcatttacaccttcata ttatgaggatggattgattggctatgatttgggtctagtttctttatataatgagaaatttttatattaataactttggaatgat ggtatctggattaatatatattaatgattcattatattattttaaaccaccagtaaataatttgataactggatttgtgactgt aggcgatgataaatactactttaatccaattaatggtggagctgcttcaattggagagacaataattgatgacaaaaat tattatttcaaccaaagtggagtgttacaaacaggtgtatttagtacagaagatggatttaaatattttgccccagctaat acacttgatgaaaacctagaaggagaagcaattgattttactggaaaattaattattgacgaaaatatttattattttgat gataattatagaggagctgtagaatggaaagaattagatggtgaaatgcactattttagcccagaaacaggtaaagc ttttaaaggtctaaatcaaataggtgattataaatactatttcaattctgatggagttatgcaaaaaggatttgttagtata aatgataataaacactattttgatgattctggtgttatgaaagtaggttacactgaaatagatggcaagcatttctacttt gctgaaaacggagaaatgcaaataggagtatttaatacagaagatggatttaaatattttgctcatcataatgaagatt taggaaatgaagaaggtgaagaaatctcatattctggtatattaaatttcaataataaaatttactattttgatgattcattt acagctgtagttggatggaaagatttagaggatggttcaaagtattattttgatgaagatacagcagaagcatatata ggtttgtcattaataaatgatggtcaatattattttaatgatgatggaattatgcaagttggatttgtcactataaatgataa agtcttctacttctctgactctggaattatagaatctggagtacaaaacatagatgacaattatttctatatagatgataat ggtatagttcaaattggtgtatttgatacttcagatggatataaatattttgcacctgctaatactgtaaatgataatattta cggacaagcagttgaatatagtggtttagttagagttggtgaagatgtatattatttttggagaaacatatacaattgaga ctggatggatatatgatatggaaaatgaaagtgataaatattatttcaatccagaaactaaaaaagcatgcaaaggta ttaatttaattgatgatataaaatattattttgatgagaagggcataatgagaacgggtcttatatcatttgaaaataataa ttattactttaatgagaatggtgaaatgcaatttggttatataaatatagaagataagatgttctattttggtgaagatggt gtcatgcagattggagtatttaatacaccagatggatttaaatactttgcacatcaaaatactttggatgagaattttgag ggagaatcaataaactatactggttggttagatttagatgaaaagagatattattttacagatgaatatattgcagcaac tggttcagttattattgatggtgaggagtattattttgatcctgatacagctcaattagtgattagtgaatag

(continued)

Name	SEQ ID NOs:	Sequences
Amino acid sequence of *trdB (strain 630)*	8	MSLVNRKQLEKMANVRFRTQEDEYVAILDALEEYHNMSENTVVEKYL KLKDINSLTDIYIDTYKKSGRNKALKKFKEYLVTEVLELKNNNLTPVEK NLHFVWIGGQINDTAINYINQWKDVNSDYNVNVFYDSNAFLINTLKKT VVESAINDTLESFRENLNDPRFDYNKFFRKRMEIIYDKQKNFINYYKAQR EENPELIIDDIVKTYLSNEYSKEIDELNTYIEESLNKITQNSGNDVRNFEEF KNGESFNLYEQELVERWNLAAASDILRISALKEIGGMYLDVDMLPGIQP DLFESIEKPSSVTVDFWEMTKLEAIMKYKEYIPEYTSEHFDMLDEEVQSS FESVLASKSDKSEIFSSLGDMEASPLEVKIAFNSKGIINQGLISVKDSYCSN LIVKQIENRYKILNNSLNPAISEDNDFNTTTNTFIDSIMAEANADNGRFM MELGKYLRVGFFPDVKTTINLSGPEAYAAAYQDLLMFKEGSMNIHLIEA
		DLRNFEISKTNISQSTEQEMASLWSFDDARAKAQFEEYKRNYFEGSLGE DDNLDFSQNIVVDKEYLLEKISSLARSSERGYIHYIVQLQGDKISYEAAC NLFAKTPYDSVLFQKNIEDSEIAYYYNPGDGEIQEIDKYKIPSIISDRPKIK LTFIGHGKDEFNTDIFAGFDVDSLSTEIEAAIDLAKEDISPKSIEINLLGCN MFSYSINVEETYPGKLLLKVKDKISELMPSISQDSIIVSANQYEVRINSEG RRELLDHSGEWINKEESIIKDISSKEYISFNPKENKITVKSKNLPELSTLLQ EIRNNSNSSDIELEEKVMLTECEINVISNIDTQIVEERIEEAKNLTSDSINYI KDEFKLIESISDALCDLKQQNELEDSHFISFEDISETDEGFSIRFINKETGES IFVETEKTIFSEYANHITEEISKIKGTIFDTVNGKLVKKVNLDTTHEVNTL NAAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQLFSTGLNTITDAAKVVE LVSTALDETIDLLPTLSEGLPIIATIIDGVSLGAAIKELSETSDPLLRQEIEA KIGIMAVNLTTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLRD KATKVVDYFKHVSLVETEGVFTLLDDKIMMPQDDLVISEIDFNNNSIVL GKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELD LSKDLMVLPNAPNRVFAWETGWTPGLRSLENDGTKLLDRIRDNYEGEF YWRYFAFIADALITTLKPRYEDTNIRINLDSNTRSFIVPIITTEYIREKLSYS FYGSGGTYALSLSQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKKG DLIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIIEVD LLSKSYKLLISGELKILMLNSNHIQQKIDYIGFNSELQKNIPYSFVDSEGK ENGFINGSTKEGLFVSELPDVVLISKVYMDDSKPSFGYYSNNLKDVKVIT KDNVNILTGYYLKDDIKISLSLTLQDEKTIKLNSVHLDESGVAEILKFMN RKGNTNTSDSLMSFLESMNIKSIFVNFLQSNIKFILDANFIISGTTSIGQFEF ICDENDNIQPYFIKFNTLETNYTLYVGNRQNMIVEPNYDLDDSGDISSTVI NFSQKYLYGIDSCVNKVVISPNIYTDEINITPVYETNNTYPEVIVLDANYI NEKINVNINDLSIRYVWSNDGNDFILMSTSEENKVSQVKIRFVNVFKDKT LANKLSFNFSDKQDVPVSEIILSFTPSYYEDGLIGYDLGLVSLYNEKFYIN NFGMMVSGLIYINDSLYYFKPPVNNLITGFVTVGDDKYYFNPINGGAASI GETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEGEAIDFT GKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLNQIGDY KYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAEN GEMQIGVFNTEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDS FTAVVGWKDLEDGSKYYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVG FVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYF APANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENES DKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENG EMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEG ESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYYFDPDTAQLVISE
Forward primer	9	caccACTAGTatgaacttagtaactggatggc
Reverse primer	10	CTCGAGttagccatatatcccaggggc
Forward primer	11	caccATGCATatgagtttagttaatagaaaacag
Reverse primer	12	ggcCTCGAGctattcactaatcactaattgagc

(continued)

Name	SEQ ID NOs:	Sequences
Forward primer	13	AGATCTATGCATGAGCTCctcgagcccaaaacgaaaggctcagc
Reverse primer	14	cggtccggggccatatatcccaggggcttttactcc
Forward primer	15	caccCCATTGatggtaacaggagtatttaaagga
Reverse primer	16	CTCGAGctattcactaatcactaattgagctg
C-TADCTB (nucleic acid sequence)	17	atggtaacaggagtatttaaaggacctaatggatttgagtattttgcacctgctaatactcacaataataacatagaag gtcaggctatagtttaccagaacaaattcttaactttgaatggcaaaaaatattattttgataatgactcaaaagcagtta ctggatggcaaaccattgatggtaaaaaatattactttaatcttaacactgctgaagcagctactggatggcaaactat tgatggtaaaaaatattacttttaatcttaacactgctgaagcagctactggatggcaaactattgatggtaaaaaatatt actttaatactaacactttcatagcctcaactggttatacaagtattaatggtaaacatttttatttttaatactgatggtatta tgcagataggagtgtttaaaggacctaatggatttgaatactttgcacctgctaatacggatgctaacaacatagaag gtcaagctatactttaccaaaataaattcttaactttgaatggtaaaaaatattactttggtagtgactcaaaagcagtta ccggactgcgaactattgatggtaaaaaatattactttaatactaacactgctgttgcagttactggatggcaaactatt aatggtaaaaaatactactttaatactaacacttctatagcttcaactggttatacaattattagtggtaaacatttttatttt aatactgatggtattatgcagataggagtgtttaaaggacctgatggatttgaatactttgcacctgctaatacagatg

(continued)

Name	SEQ ID NOs:	Sequences
		ctaacaatatagaaggtcaagctatacgttatcaaaatagattcctatatttacatgacaatatatattattttggtaataat tcaaaagcggctactggttgggtaactattgatggtaatagatattacttcgagcctaatacagctatgggtgcgaat ggttataaaactattgataataaaaattttactttagaaatggtttacctcagataggagtgtttaaagggtctaatggat ttgaatactttgcacctgctaatacggatgctaacaatatagaaggtcaagctatacgttatcaaaatagattcctacat ttacttggaaaaatatattactttggtaataattcaaaagcagttactggatggcaaactattaatggtaaagtatattact ttatgcctgatactgctatggctgcagctggtggacttttcgagattgatggtgttatatatttctttggtgttgatggagt aaaagcccctgggatatatggcAGATCTATGCATaatttgataactggatttgtgactgtaggcgatgata aatactactttaatccaattaatggtggagctgcttcaattggagagacaataattgatgacaaaaattattatttcaacc aaagtggagtgttacaaacaggtgtatttagtacagaagatggatttaaatattttgccccagctaatacacttgatga aaacctagaaggagaagcaattgattttactggaaaattaattattgacgaaaatatttattattttgatgataattataga ggagctgtagaatggaaagaattagatggtgaaatgcactattttagcccagaaacaggtaaagcttttaaaggtct aaatcaaataggtgattataaatactatttcaattctgatggagttatgcaaaaaggatttgttagtataaatgataataa acactattttgatgattctggtgttatgaaagtaggttacactgaaatagatggcaagcatttctactttgctgaaaacg gagaaatgcaaataggagtatttaatacagaagatggatttaaatattttgctcatcataatgaagatttaggaaatga agaaggtgaagaaatctcatattctggtatattaaatttcaataataaaatttactattttgatgattcatttacagctgtag ttggatggaaagatttagaggatggttcaaagtattatttttgatgaagatacagcagaagcatatataggtttgtcatta ataaatgatggtcaatattatttttaatgatgatggaattatgcaagttggatttgtcactataaatgataaagtcttctactt ctctgactctggaattatagaatctggagtacaaaacatagatgacaattatttctatatagatgataatggtatagttca aattggtgtatttgatacttcagatggatataaatattttgcacctgctaatactgtaaatgataatatttacggacaagca gttgaatatagtggtttagttagagttggggaagatgtatattatttttggagaaacatatacaattgagactggatggat atatgatatggaaaatgaaagtgataaatattatttcaatccagaaactaaaaaagcatgcaaaggtattaatttaattg atgatataaaatattattttgatgagaagggcataatgagaacgggtcttatatcatttgaaaataataattattactttaa tgagaatggtgaaatgcaatttggttatataaatatagaagataagatgttctattttggtgaagatggtgtcatgcaga ttggagtatttaatacaccagatggatttaaatactttgcacatcaaaatactttggatgagaattttgagggagaatca ataaactatactggttggttagatttagatgaaaagagatattattttacagatgaatatattgcagcaactggttcagtt attattgatggtgaggagtattattttgatcctgatacagctcaattagtgattagtgaaCTCGAGgattaatatat attaatgattcattatattattttaaaccaccagtaaataatttgataactggatttgtgactgtaggcgatgataaatacta ctttaatccaattaatggtggagctgcttcaattggagagacaataattgatgacaaaaattattatttcaaccaaagtg gagtgttacaaacaggtgtatttagtacagaagatggatttaaatattttgccccagctaatacacttgatgaaaaccta gaaggagaagcaattgattttactggaaaattaattattgacgaaaatatttattattttgatgataattatagaggagct gtagaatggaaagaattagatggtgaaatgcactattttagcccagaaacaggtaaagcttttaaaggtctaaatcaa ataggtgattataaatactatttcaattctgatggagttatgcaaaaaggatttgttagtataaatgataataaacactatt ttgatgattctggtgttatgaaagtaggttacactgaaatagatggcaagcatttctactttgctgaaaacggagaaat gcaaataggagtatttaatacagaagatggatttaaatattttgctcatcataatgaagatttaggaaatgaagaaggt gaagaaatctcatattctggtatattaaatttcaataataaaatttactattttgatgattcatttacagctgtagttggatgg aaagatttagaggatggttcaaagtattatttttgatgaagatacagcagaagcatatataggtttgtcattaataaatga tggtcaatattatttttaatgatgatggaattatgcaagttggatttgtcactataaatgataaagtcttctacttctctgact ctggaattatagaatctggagtacaaaacatagatgacaattatttctatatagatgataatggtatagttcaaatttggtg tatttgatacttcagatggatataaatattttgcacctgctaatactgtaaatgataatatttacggacaagcagttgaata tagtggtttagttagagttggggaagatgtatattatttttggagaaacatatacaattgagactggatggatatatgatat ggaaaatgaaagtgataaatattatttcaatccagaaactaaaaaagcatgcaaaggtattaatttaattgatgatataa aatattattttgatgagaagggcataatgagaacgggtcttatatcatttgaaaataataattattactttaatgagaatg gtgaaatgcaatttggttatataaatatagaagataagatgttctattttggtgaagatggtgtcatgcagattggagtat ttaatacaccagatggatttaaatactttgcacatcaaaatactttggatgagaattttgagggagaatcaataaactat actggttggttagatttagatgaaaagagatattattttacagatgaatatattgcagcaactggttcagttattattgatg gtgaggagtattattttgatcctgatacagctcaattagtgattagtgaatag

(continued)

Name	SEQ ID NOs:	Sequences
C-TADCTB (amino acid sequence)	18	MVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDN DSKAVTGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAA TGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPN GFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTI DGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYF NTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIY YFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLP QIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNS KAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGI YGRSMHNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGV
		LQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYR GAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFV SINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKY FAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSK YYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGII ESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAV EYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGI NLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYF GEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYYFDPDTAQLVISELEGLIYINDSLYYFKPPV NNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVF STEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEW KELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNK HYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNE DLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDED TAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNI DDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVR VGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGINLIDDIK YYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYFGEDGV MQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDE YIAATGSVIIDGEEYYFDPDTAQLVISE

(continued)

Name	SEQ ID NOs:	Sequences
C-TANCTB (nucleic acid sequence)	19	atggtaacaggagtatttaaaggacctaatggatttgagtattttgcacctgctaatactcacaataataacatagaag gtcaggctatagtttaccagaacaaattcttaactttgaatggcaaaaaatattattttgataatgactcaaaagcagtta ctggatggcaaaccattgatggtaaaaaatattactttaatcttaacactgctgaagcagctactggatggcaaactat tgatggtaaaaaatattacttttaatcttaacactgctgaagcagctactggatggcaaactattgatggtaaaaaatatt actttaatactaacactttcatagcctcaactggttatacaagtattaatggtaaacatttttatttttaatactgatggtatta tgcagataggagtgtttaaaggacctaatggatttgaatactttgcacctgctaatacggatgctaacaacatagaag gtcaagctatactttaccaaaataaattcttaactttgaatggtaaaaaatattactttggtagtgactcaaaagcagtta ccggactgcgaactattgatggtaaaaaatattactttaatactaacactgctgttgcagttactggatggcaaactatt aatggtaaaaaatactactttaatactaacacttctatagcttcaactggttatacaattattagtggtaaacatttttatttt aatactgatggtattatgcagataggagtgtttaaaggacctgatggatttgaatactttgcacctgctaatacagatg ctaacaatatagaaggtcaagctatacgttatcaaaatagattcctatatttacatgacaatatatattattttggtaataat tcaaaagcggctactggttgggtaactattgatggtaatagatattacttcgagcctaatacagctatgggtgcgaat ggttataaaactattgataataaaaattttactttagaaatggtttacctcagataggagtgtttaaagggtctaatggat ttgaatactttgcacctgctaatacggatgctaacaatatagaaggtcaagctatacgttatcaaaatagattcctacat ttacttggaaaaatatattactttggtaataattcaaaagcagttactggatggcaaactattaatggtaaagtatattact ttatgcctgatactgctatggctgcagctggtggacttttcgagattgatggtgttatatatttctttggtgttgatggagt aaaagcccctgggatatatggcAGATCTATGCATaatttgataactggatttgtgactgtaggcgatgata aatactactttaatccaattaatggtggagctgcttcaattggagagacaataattgatgacaaaaattattatttcaacc aaagtggagtgttacaaacaggtgtatttagtacagaagatggatttaaatattttgccccagctaatacacttgatga aaacctagaaggagaagcaattgattttactggaaaattaattattgacgaaaatatttattattttgatgataattataga ggagctgtagaatggaaagaattagatggtgaaatgcactattttagcccagaaacaggtaaagcttttaaaggtct aaatcaaataggtgattataaatactatttcaattctgatggagttatgcaaaaaggatttgttagtataaatgataataa acactattttgatgattctggtgttatgaaagtaggttacactgaaatagatggcaagcatttctactttgctgaaaacg gagaaatgcaaataggagtatttaatacagaagatggatttaaatattttgctcatcataatgaagatttaggaaatga agaaggtgaagaaatctcatattctggtatattaaatttcaataataaaatttactattttgatgattcatttacagctgtag ttggatggaaagatttagaggatggttcaaagtattattttgatgaagatacagcagaagcatatatagtttgtcatta ataaatgatggtcaatattattttaatgatgatggaattatgcaagttggatttgtcactataaatgataaagtcttctactt ctctgactctggaattatagaatctggagtacaaaacatagatgacaattatttctatatagatgataatggtatagttca aattggtgtatttgatacttcagatggatataaatattttgcacctgctaatactgtaaatgataatatttacggacaagca gttgaatatagtggtttagttagagttggggaagatgtatattatttttggagaaacatatacaattgagactggatggat atatgatatggaaaatgaaagtgataaatattatttcaatccagaaactaaaaaagcatgcaaaggtattaatttaattg atgatataaaatattattttgatgagaagggcataatgagaacgggtcttatatcatttgaaaataataattattactttaa tgagaatggtgaaatgcaatttggttatataaatatagaagataagatgttctatttttggtgaagatggtgtcatgcaga ttggagtatttaatacaccagatggatttaaatactttgcacatcaaaatactttggatgagaaattttgagggagaatca ataaactatactggttggttagatttagatgaaaagagatattattttacagatgaatatattgcagcaactggttcagtt
		attattgatggtgaggagtattattttgatcctgatacagctcaattagtgattagtgaaCTCGAGggattaatatat attaatgattcattatattattttaaaccaccagtaaataatttgataactggatttgtgactgtaggcgatgataaatacta ctttaatccaattaatggtggagctgcttcaattggagagacaataattgatgacaaaaattattatttcaaccaaagtg gagtgttacaaacaggtgtatttagtacagaagatggatttaaatattttgccccagctaatacacttgatgaaaaccta gaaggagaagcaattgattttactggaaaattaattattgacgaaaatatttattattttgatgataattatagaggagct gtagaatggaaagaattagatggtgaaatgcactattttagcccagaaacaggtaaagcttttaaaggtctaaatcaa ataggtgattataaatactatttcaattctgatggagttatgcaaaaaggatttgttagtataaatgataataaacactatt ttgatgattctggtgttatgaaagtaggttacactgaaatagatggcaagcatttctactttgctgaaaacggagaaat gcaaataggagtatttaatacagaagatggatttaaatattttgctcatcataatgaagatttaggaaatgaagaaggt gaagaaatctcatattct

(continued)

Name	SEQ ID NOs:	Sequences
C-TANCTB (amino acid sequence)	20	MVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDN DSKAVTGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAA TGWQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGPN GFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSKAVTGLRTI DGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYF NTDGIMQIGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIY YFGNNSKAATGWVTIDGNRYYFEPNTAMGANGYKTIDNKNFYFRNGLP QIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRFLHLLGKIYYFGNNS KAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGI YGRSMHNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGV LQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYR GAVEWKELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFV SINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKY FAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSK YYFDEDTAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGII ESGVQNIDDNYFYIDDNGIVQIGVFDTSDGYKYFAPANTVNDNIYGQAV EYSGLVRVGEDVYYFGETYTIETGWIYDMENESDKYYFNPETKKACKGI NLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFYF GEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRY YFTDEYIAATGSVIIDGEEYYFDPDTAQLVISELEGLIYINDSLYYFKPPV NNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVF STEDGFKYFAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEW KELDGEMHYFSPETGKAFKGLNQIGDYKYYFNSDGVMQKGFVSINDNK HYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFNTEDGFKYFAHHNE DLGNEEGEEISYS

SEQUENCE LISTING

[0242]

<110> Intercell AG
Intercell USA, Inc.

<120> Novel Isolated Polypeptide of the Toxin A and Toxin B Proteins of C. difficile and Uses Thereof

<130> ICP091/PCT

<150> US 61/379,892
<151> 2010-09-03

<160> 20

<170> PatentIn version 3.5

<210> 1
<211> 2877
<212> DNA
<213> artificial

<220>
<223> C-TAB.G5 fusion DNA

<400> 1

atggtaacag gagtatttaa aggacctaat ggatttgagt attttgcacc tgctaatact 60

cacaataata acatagaagg tcaggctata gtttaccaga acaaattctt aactttgaat 120

ggcaaaaaat attattttga taatgactca aaagcagtta ctggatggca aaccattgat 180

ggtaaaaaat attactttaa tcttaacact gctgaagcag ctactggatg gcaaactatt 240

gatggtaaaa aatattactt taatcttaac actgctgaag cagctactgg atggcaaact 300

attgatggta aaaaatatta ctttaatact aacactttca tagcctcaac tggttataca 360

agtattaatg gtaaacattt ttattttaat actgatggta ttatgcagat aggagtgttt 420

aaaggaccta atggatttga atactttgca cctgctaata ctcataataa caacatagaa 480

ggtcaagcta tactttacca aaataaattc ttaactttga atggtaaaaa atattacttt 540

ggtagtgact caaaagcagt taccggattg cgaactattg atggtaaaaa atattacttt 600

aatactaaca ctgctgttgc agttactgga tggcaaacta ttaatggtaa aaaatactac 660

tttaatacta acacttctat agcttcaact ggttatacaa ttattagtgg taaacatttt 720

tattttaata ctgatggtat tatgcagata ggagtgttta aaggacctga tggatttgaa 780

tactttgcac ctgctaatac agatgctaac aatatagaag gtcaagctat acgttatcaa 840

aatagattcc tatatttaca tgacaatata tattattttg gtaataattc aaaagcagct 900

actggttggg taactattga tggtaataga tattacttcg agcctaatac agctatgggt 960

gcgaatggtt ataaaactat tgataataaa aatttttact ttagaaatgg tttacctcag 1020

ataggagtgt ttaaagggtc taatggattt gaatactttg cacctgctaa tacggatgct 1080

aacaatatag aaggtcaagc tatacgttat caaaatagat tcctacattt acttggaaaa 1140

atatattact ttggtaataa ttcaaaagca gttactggat ggcaaactat taatggtaaa 1200

gtatattact ttatgcctga tactgctatg gctgcagctg gtggactttt cgagattgat 1260

ggtgttatat atttctttgg tgttgatgga gtaaaagccc ctgggatata tggcagatct 1320

atgcataatt tgataactgg atttgtgact gtaggcgatg ataaatacta ctttaatcca 1380

attaatggtg gagctgcttc aattggagag acaataattg atgacaaaaa ttattatttc 1440

aaccaaagtg gagtgttaca aacaggtgta tttagtacag aagatggatt taaatatttt 1500

gccccagcta atacacttga tgaaaaccta gaaggagaag caattgattt tactggaaaa 1560

ttaattattg acgaaaatat ttattatttt gatgataatt atagaggagc tgtagaatgg 1620

aaagaattag atggtgaaat gcactatttt agcccagaaa caggtaaagc ttttaaaggt 1680

ctaaatcaaa taggtgatta taaatactat ttcaattctg atggagttat gcaaaaagga 1740

tttgttagta taaatgataa taaacactat tttgatgatt ctggtgttat gaaagtaggt 1800

tacactgaaa tagatggcaa gcatttctac tttgctgaaa acggagaaat gcaaatagga 1860

gtatttaata cagaagatgg atttaaatat tttgctcatc ataatgaaga tttaggaaat 1920

gaagaaggtg aagaaatctc atattctggt atattaaatt tcaataataa aatttactat 1980

tttgatgatt catttacagc tgtagttgga tggaaagatt tagaggatgg ttcaaagtat 2040

tattttgatg aagatacagc agaagcatat ataggtttgt cattaataaa tgatggtcaa 2100

tattatttta atgatgatgg aattatgcaa gttggatttg tcactataaa tgataaagtc 2160

ttctacttct ctgactctgg aattatagaa tctggagtac aaaacataga tgacaattat 2220

ttctatatag atgataatgg tatagttcaa attggtgtat ttgatacttc agatggatat 2280

aaatattttg cacctgctaa tactgtaaat gataatattt acggacaagc agttgaatat 2340

agtggtttag ttagagttgg ggaagatgta tattattttg gagaaacata tacaattgag 2400

actggatgga tatatgatat ggaaaatgaa agtgataaat attatttcaa tccagaaact 2460

aaaaaagcat gcaaaggtat taatttaatt gatgatataa aatattattt tgatgagaag 2520

ggcataatga gaacgggtct tatatcattt gaaaataata attattactt taatgagaat 2580

ggtgaaatgc aatttggtta tataaatata gaagataaga tgttctattt tggtgaagat 2640

ggtgtcatgc agattggagt atttaataca ccagatggat ttaaatactt tgcacatcaa 2700

aatactttgg atgagaattt tgagggagaa tcaataaact atactggttg gttagattta 2760

gatgaaaaga gatattattt tacagatgaa tatattgcag caactggttc agttattatt 2820

gatggtgagg agtattattt tgatcctgat acagctcaat tagtgattag tgaatag 2877

<210> 2
<211> 958
<212> PRT

<213> artificial

<220>
<223> C-TAB.G5 fusion protein

<400> 2

```
Met Val Thr Gly Val Phe Lys Gly Pro Asn Gly Phe Glu Tyr Phe Ala
1               5                  10                  15

Pro Ala Asn Thr His Asn Asn Asn Ile Glu Gly Gln Ala Ile Val Tyr
            20                  25                  30

Gln Asn Lys Phe Leu Thr Leu Asn Gly Lys Lys Tyr Tyr Phe Asp Asn
        35                  40                  45

Asp Ser Lys Ala Val Thr Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr
    50                  55                  60

Tyr Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr Gly Trp Gln Thr Ile
65                  70                  75                  80

Asp Gly Lys Lys Tyr Tyr Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr
                85                  90                  95

Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr
            100                 105                 110

Phe Ile Ala Ser Thr Gly Tyr Thr Ser Ile Asn Gly Lys His Phe Tyr
        115                 120                 125

Phe Asn Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro Asn
    130                 135                 140

Gly Phe Glu Tyr Phe Ala Pro Ala Asn Thr His Asn Asn Asn Ile Glu
145                 150                 155                 160

Gly Gln Ala Ile Leu Tyr Gln Asn Lys Phe Leu Thr Leu Asn Gly Lys
                165                 170                 175

Lys Tyr Tyr Phe Gly Ser Asp Ser Lys Ala Val Thr Gly Leu Arg Thr
            180                 185                 190

Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Ala Val Ala Val
        195                 200                 205

Thr Gly Trp Gln Thr Ile Asn Gly Lys Lys Tyr Tyr Phe Asn Thr Asn
    210                 215                 220
```

```
Thr Ser Ile Ala Ser Thr Gly Tyr Thr Ile Ile Ser Gly Lys His Phe
225                 230                 235                 240

Tyr Phe Asn Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro
                245                 250                 255

Asp Gly Phe Glu Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile
            260                 265                 270

Glu Gly Gln Ala Ile Arg Tyr Gln Asn Arg Phe Leu Tyr Leu His Asp
        275                 280                 285

Asn Ile Tyr Tyr Phe Gly Asn Asn Ser Lys Ala Ala Thr Gly Trp Val
    290                 295                 300

Thr Ile Asp Gly Asn Arg Tyr Tyr Phe Glu Pro Asn Thr Ala Met Gly
305                 310                 315                 320

Ala Asn Gly Tyr Lys Thr Ile Asp Asn Lys Asn Phe Tyr Phe Arg Asn
                325                 330                 335

Gly Leu Pro Gln Ile Gly Val Phe Lys Gly Ser Asn Gly Phe Glu Tyr
            340                 345                 350

Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu Gly Gln Ala Ile
        355                 360                 365

Arg Tyr Gln Asn Arg Phe Leu His Leu Leu Gly Lys Ile Tyr Tyr Phe
    370                 375                 380

Gly Asn Asn Ser Lys Ala Val Thr Gly Trp Gln Thr Ile Asn Gly Lys
385                 390                 395                 400

Val Tyr Tyr Phe Met Pro Asp Thr Ala Met Ala Ala Ala Gly Gly Leu
                405                 410                 415

Phe Glu Ile Asp Gly Val Ile Tyr Phe Phe Gly Val Asp Gly Val Lys
            420                 425                 430

Ala Pro Gly Ile Tyr Gly Arg Ser Met His Asn Leu Ile Thr Gly Phe
        435                 440                 445

Val Thr Val Gly Asp Asp Lys Tyr Tyr Phe Asn Pro Ile Asn Gly Gly
    450                 455                 460

Ala Ala Ser Ile Gly Glu Thr Ile Ile Asp Asp Lys Asn Tyr Tyr Phe
465                 470                 475                 480
```

```
Asn Gln Ser Gly Val Leu Gln Thr Gly Val Phe Ser Thr Glu Asp Gly
                485                 490                 495

Phe Lys Tyr Phe Ala Pro Ala Asn Thr Leu Asp Glu Asn Leu Glu Gly
            500                 505                 510

Glu Ala Ile Asp Phe Thr Gly Lys Leu Ile Ile Asp Glu Asn Ile Tyr
        515                 520                 525

Tyr Phe Asp Asp Asn Tyr Arg Gly Ala Val Glu Trp Lys Glu Leu Asp
    530                 535                 540

Gly Glu Met His Tyr Phe Ser Pro Glu Thr Gly Lys Ala Phe Lys Gly
545                 550                 555                 560

Leu Asn Gln Ile Gly Asp Tyr Lys Tyr Tyr Phe Asn Ser Asp Gly Val
                565                 570                 575

Met Gln Lys Gly Phe Val Ser Ile Asn Asp Asn Lys His Tyr Phe Asp
            580                 585                 590

Asp Ser Gly Val Met Lys Val Gly Tyr Thr Glu Ile Asp Gly Lys His
        595                 600                 605

Phe Tyr Phe Ala Glu Asn Gly Glu Met Gln Ile Gly Val Phe Asn Thr
    610                 615                 620

Glu Asp Gly Phe Lys Tyr Phe Ala His His Asn Glu Asp Leu Gly Asn
625                 630                 635                 640

Glu Glu Gly Glu Glu Ile Ser Tyr Ser Gly Ile Leu Asn Phe Asn Asn
                645                 650                 655

Lys Ile Tyr Tyr Phe Asp Asp Ser Phe Thr Ala Val Val Gly Trp Lys
            660                 665                 670

Asp Leu Glu Asp Gly Ser Lys Tyr Tyr Phe Asp Glu Asp Thr Ala Glu
        675                 680                 685

Ala Tyr Ile Gly Leu Ser Leu Ile Asn Asp Gly Gln Tyr Tyr Phe Asn
    690                 695                 700

Asp Asp Gly Ile Met Gln Val Gly Phe Val Thr Ile Asn Asp Lys Val
705                 710                 715                 720

Phe Tyr Phe Ser Asp Ser Gly Ile Ile Glu Ser Gly Val Gln Asn Ile
                725                 730                 735
```

Asp Asp Asn Tyr Phe Tyr Ile Asp Asp Asn Gly Ile Val Gln Ile Gly
740 745 750

Val Phe Asp Thr Ser Asp Gly Tyr Lys Tyr Phe Ala Pro Ala Asn Thr
755 760 765

Val Asn Asp Asn Ile Tyr Gly Gln Ala Val Glu Tyr Ser Gly Leu Val
770 775 780

Arg Val Gly Glu Asp Val Tyr Tyr Phe Gly Glu Thr Tyr Thr Ile Glu
785 790 795 800

Thr Gly Trp Ile Tyr Asp Met Glu Asn Glu Ser Asp Lys Tyr Tyr Phe
805 810 815

Asn Pro Glu Thr Lys Lys Ala Cys Lys Gly Ile Asn Leu Ile Asp Asp
820 825 830

Ile Lys Tyr Tyr Phe Asp Glu Lys Gly Ile Met Arg Thr Gly Leu Ile
835 840 845

Ser Phe Glu Asn Asn Asn Tyr Tyr Phe Asn Glu Asn Gly Glu Met Gln
850 855 860

Phe Gly Tyr Ile Asn Ile Glu Asp Lys Met Phe Tyr Phe Gly Glu Asp
865 870 875 880

Gly Val Met Gln Ile Gly Val Phe Asn Thr Pro Asp Gly Phe Lys Tyr
885 890 895

Phe Ala His Gln Asn Thr Leu Asp Glu Asn Phe Glu Gly Glu Ser Ile
900 905 910

Asn Tyr Thr Gly Trp Leu Asp Leu Asp Glu Lys Arg Tyr Tyr Phe Thr
915 920 925

Asp Glu Tyr Ile Ala Ala Thr Gly Ser Val Ile Ile Asp Gly Glu Glu
930 935 940

Tyr Tyr Phe Asp Pro Asp Thr Ala Gln Leu Val Ile Ser Glu
945 950 955

<210> 3
<211> 2885
<212> DNA
<213> artificial

<220>

<223> C-TAB.G5.1 fusion DNA

<400> 3

ccatggttac aggtgttttc aaaggtccga acggctttga atattttgca ccggcaaata 60

cccacaataa taatattgaa ggccaggcca tcgtgtatca gaataaattt ctgaccctga 120

acggcaaaaa atactatttc gataacgata gcaaagcagt taccggttgg caaaccattg 180

atggcaaaaa atattacttc aacctgaata ccgcagaagc agcaaccggc tggcagacga 240

tcgacggtaa aaagtactat tttaacctga acacagccga agccgctaca ggctggcaga 300

caatagatgg gaagaagtat tattttaata ccaatacctt tattgccagc accggctata 360

ccagcattaa tggcaaacac ttctatttta acaccgatgg tattatgcag atcggtgtgt 420

ttaagggccc taatggtttt gagtacttcg ctccggctaa taccgatgca aataacatcg 480

aaggtcaggc aattctgtac cagaacaaat ttttaacgct gaacggtaag aaatattact 540

ttggtagcga ttcaaaagcc gttaccggtc tgcgtacgat cgacggcaag aaatattatt 600

tcaatacaaa caccgcagtt gccgtgacag gttggcagac gataaatggt aagaagtact 660

acttcaacac caataccagc attgcaagta ccggttatac cattatcagc ggcaaacact 720

tttacttcaa tacagacggc attatgcaga ttggcgtttt caaaggtccg gatggtttcg 780

agtactttgc ccctgcaaat acagatgcaa acaatattga gggacaggca attcgctatc 840

agaatcgttt tctgtatctg cacgataaca tctattactt cggcaataat tcaaaagcag 900

ccaccggttg ggttacaatt gatggtaatc gttattactt tgagccgaat accgcaatgg 960

gtgcaaatgg ttataaaacc atcgataaca aaaattttta tttccgcaac ggtctgccgc 1020

agattggtgt ttttaagggt agcaatggct tcgagtattt tgcgccagcc aacaccgatg 1080

ccaacaacat tgaaggccaa gcgattcgtt atcaaaaccg ctttctgcat ctgctgggca 1140

aaatttatta ctttggcaac aatagcaaag cggtgacggg ctggcaaacc attaacggta 1200

aagtttatta tttcatgccg gataccgcta tggcagcagc cggtggtctg tttgaaattg 1260

atggcgtgat ttatttttt ggcgtggatg gtgttaaagc accgggtatt tatggtcgta 1320

gcatgcataa tctgattacc ggttttgtta ccgtgggcga cgataaatac tactttaatc 1380

cgattaatgg tggtgcagca agcattggtg aaaccattat cgatgacaaa aactattatt 1440

ttaaccagag cggtgttctg cagacaggtg tttttagcac cgaagatggc ttcaaatatt 1500

ttgctcctgc gaatacactg gatgaaaatc tggaaggtga agcaattgat tttaccggca 1560

aactgatcat cgacgagaac atctactatt ttgatgataa ttatcgcggt gccgtggaat 1620

ggaaagaact ggatggtgaa atgcactatt ttagtccgga aaccggtaaa gcctttaaag 1680

gtctgaatca gatcggcgat tacaagtatt actttaattc agatggcgtg atgcagaaag 1740

gctttgtgag cattaacgac aacaaacact attttgacga cagcggtgtg atgaaagtgg 1800

gttataccga aatcgacggg aaacattttt attttgccga aaacggcgaa atgcagattg 1860

gagtatttaa taccgaggac ggctttaaat actttgccca tcataatgaa gatctgggta 1920

atgaagaagg cgaagaaatt agctatagcg gcattctgaa ttttaataac aagatctatt 1980

atttcgatga tagcttcacc gcagttgttg gttggaaaga tctggaagat ggcagcaaat 2040

attattttga tgaagatacc gcagaggcct atattggtct gagcctgatt aatgatggcc 2100

agtattattt caacgatgat ggtatcatgc aggttggttt tgtgaccatc aacgataaag 2160

tgttctattt cagcgatagc ggcattattg aaagcggtgt tcagaacatc gacgataact 2220

atttctacat cgatgataac ggtattgttc agattggcgt gtttgatacc tccgatggtt 2280

ataaatattt cgcaccagcc aataccgtga acgataatat ttatggtcag gcagttgaat 2340

attcaggtct ggttcgtgtt ggcgaagatg tttattattt tggcgaaacc tataccattg 2400

aaaccggctg gatctatgat atggaaaacg agagcgacaa gtactatttc aatccggaaa 2460

cgaaaaaagc ctgcaaaggc attaatctga tcgacgatat taagtactac tttgacgaaa 2520

aaggcattat gcgtaccggt ctgattagct ttgagaacaa caactattac ttcaatgaga 2580

acggtgagat gcagtttggc tatatcaaca tcgaggacaa aatgttttat tttggtgagg 2640

acggtgtgat gcagataggg gtttttaata caccggatgg gtttaagtat tttgcacatc 2700

agaacaccct ggatgaaaac tttgaaggcg aaagcattaa ttataccggt tggctggatc 2760

tggatgagaa acgttattat ttcaccgacg aatacattgc agcaaccggt agcgttatta 2820

ttgatggtga ggaatattac ttcgatccgg atacagcaca gctggttatt agcgaataac 2880

tcgag 2885

<210> 4
<211> 957
<212> PRT
<213> artificial

<220>
<223> C-TAB.G5.1 fusion protein

<400> 4

```
Val Thr Gly Val Phe Lys Gly Pro Asn Gly Phe Glu Tyr Phe Ala Pro
1               5                  10                  15

Ala Asn Thr His Asn Asn Asn Ile Glu Gly Gln Ala Ile Val Tyr Gln
            20                  25                  30

Asn Lys Phe Leu Thr Leu Asn Gly Lys Lys Tyr Tyr Phe Asp Asn Asp
        35                  40                  45

Ser Lys Ala Val Thr Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr Tyr
```

```
    50                  55                  60

Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr Gly Trp Gln Thr Ile Asp
65                  70                  75                  80

Gly Lys Lys Tyr Tyr Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr Gly
                85                  90                  95

Trp Gln Thr Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Phe
            100                 105                 110

Ile Ala Ser Thr Gly Tyr Thr Ser Ile Asn Gly Lys His Phe Tyr Phe
        115                 120                 125

Asn Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro Asn Gly
    130                 135                 140

Phe Glu Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu Gly
145                 150                 155                 160

Gln Ala Ile Leu Tyr Gln Asn Lys Phe Leu Thr Leu Asn Gly Lys Lys
                165                 170                 175

Tyr Tyr Phe Gly Ser Asp Ser Lys Ala Val Thr Gly Leu Arg Thr Ile
            180                 185                 190

Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Ala Val Ala Val Thr
        195                 200                 205

Gly Trp Gln Thr Ile Asn Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr
    210                 215                 220

Ser Ile Ala Ser Thr Gly Tyr Thr Ile Ile Ser Gly Lys His Phe Tyr
225                 230                 235                 240

Phe Asn Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro Asp
                245                 250                 255

Gly Phe Glu Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu
            260                 265                 270

Gly Gln Ala Ile Arg Tyr Gln Asn Arg Phe Leu Tyr Leu His Asp Asn
        275                 280                 285

Ile Tyr Tyr Phe Gly Asn Asn Ser Lys Ala Ala Thr Gly Trp Val Thr
    290                 295                 300
```

```
Ile Asp Gly Asn Arg Tyr Tyr Phe Glu Pro Asn Thr Ala Met Gly Ala
305                 310                 315                 320

Asn Gly Tyr Lys Thr Ile Asp Asn Lys Asn Phe Tyr Phe Arg Asn Gly
                325                 330                 335

Leu Pro Gln Ile Gly Val Phe Lys Gly Ser Asn Gly Phe Glu Tyr Phe
            340                 345                 350

Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu Gly Gln Ala Ile Arg
        355                 360                 365

Tyr Gln Asn Arg Phe Leu His Leu Leu Gly Lys Ile Tyr Tyr Phe Gly
    370                 375                 380

Asn Asn Ser Lys Ala Val Thr Gly Trp Gln Thr Ile Asn Gly Lys Val
385                 390                 395                 400

Tyr Tyr Phe Met Pro Asp Thr Ala Met Ala Ala Ala Gly Gly Leu Phe
                405                 410                 415

Glu Ile Asp Gly Val Ile Tyr Phe Phe Gly Val Asp Gly Val Lys Ala
            420                 425                 430

Pro Gly Ile Tyr Gly Arg Ser Met His Asn Leu Ile Thr Gly Phe Val
        435                 440                 445

Thr Val Gly Asp Asp Lys Tyr Tyr Phe Asn Pro Ile Asn Gly Gly Ala
    450                 455                 460

Ala Ser Ile Gly Glu Thr Ile Ile Asp Asp Lys Asn Tyr Tyr Phe Asn
465                 470                 475                 480

Gln Ser Gly Val Leu Gln Thr Gly Val Phe Ser Thr Glu Asp Gly Phe
                485                 490                 495

Lys Tyr Phe Ala Pro Ala Asn Thr Leu Asp Glu Asn Leu Glu Gly Glu
            500                 505                 510

Ala Ile Asp Phe Thr Gly Lys Leu Ile Ile Asp Glu Asn Ile Tyr Tyr
        515                 520                 525

Phe Asp Asp Asn Tyr Arg Gly Ala Val Glu Trp Lys Glu Leu Asp Gly
    530                 535                 540

Glu Met His Tyr Phe Ser Pro Glu Thr Gly Lys Ala Phe Lys Gly Leu
545                 550                 555                 560
```

```
Asn Gln Ile Gly Asp Tyr Lys Tyr Tyr Phe Asn Ser Asp Gly Val Met
                565                 570                 575

Gln Lys Gly Phe Val Ser Ile Asn Asp Asn Lys His Tyr Phe Asp Asp
            580                 585                 590

Ser Gly Val Met Lys Val Gly Tyr Thr Glu Ile Asp Gly Lys His Phe
        595                 600                 605

Tyr Phe Ala Glu Asn Gly Glu Met Gln Ile Gly Val Phe Asn Thr Glu
    610                 615                 620

Asp Gly Phe Lys Tyr Phe Ala His His Asn Glu Asp Leu Gly Asn Glu
625                 630                 635                 640

Glu Gly Glu Glu Ile Ser Tyr Ser Gly Ile Leu Asn Phe Asn Asn Lys
                645                 650                 655

Ile Tyr Tyr Phe Asp Asp Ser Phe Thr Ala Val Val Gly Trp Lys Asp
            660                 665                 670

Leu Glu Asp Gly Ser Lys Tyr Tyr Phe Asp Glu Asp Thr Ala Glu Ala
        675                 680                 685

Tyr Ile Gly Leu Ser Leu Ile Asn Asp Gly Gln Tyr Tyr Phe Asn Asp
    690                 695                 700

Asp Gly Ile Met Gln Val Gly Phe Val Thr Ile Asn Asp Lys Val Phe
705                 710                 715                 720

Tyr Phe Ser Asp Ser Gly Ile Ile Glu Ser Gly Val Gln Asn Ile Asp
                725                 730                 735

Asp Asn Tyr Phe Tyr Ile Asp Asp Asn Gly Ile Val Gln Ile Gly Val
            740                 745                 750

Phe Asp Thr Ser Asp Gly Tyr Lys Tyr Phe Ala Pro Ala Asn Thr Val
        755                 760                 765

Asn Asp Asn Ile Tyr Gly Gln Ala Val Glu Tyr Ser Gly Leu Val Arg
    770                 775                 780

Val Gly Glu Asp Val Tyr Tyr Phe Gly Glu Thr Tyr Thr Ile Glu Thr
785                 790                 795                 800

Gly Trp Ile Tyr Asp Met Glu Asn Glu Ser Asp Lys Tyr Tyr Phe Asn
                805                 810                 815
```

Pro Glu Thr Lys Lys Ala Cys Lys Gly Ile Asn Leu Ile Asp Asp Ile
820 825 830

Lys Tyr Tyr Phe Asp Glu Lys Gly Ile Met Arg Thr Gly Leu Ile Ser
835 840 845

Phe Glu Asn Asn Asn Tyr Tyr Phe Asn Glu Asn Gly Glu Met Gln Phe
850 855 860

Gly Tyr Ile Asn Ile Glu Asp Lys Met Phe Tyr Phe Gly Glu Asp Gly
865 870 875 880

Val Met Gln Ile Gly Val Phe Asn Thr Pro Asp Gly Phe Lys Tyr Phe
885 890 895

Ala His Gln Asn Thr Leu Asp Glu Asn Phe Glu Gly Glu Ser Ile Asn
900 905 910

Tyr Thr Gly Trp Leu Asp Leu Asp Glu Lys Arg Tyr Tyr Phe Thr Asp
915 920 925

Glu Tyr Ile Ala Ala Thr Gly Ser Val Ile Ile Asp Gly Glu Glu Tyr
930 935 940

Tyr Phe Asp Pro Asp Thr Ala Gln Leu Val Ile Ser Glu
945 950 955

<210> 5
<211> 8133
<212> DNA
<213> Clostridium difficile strain 630

<400> 5

atgtctttaa tatctaaaga agagttaata aaactcgcat atagcattag accaagagaa 60

aatgagtata aaactatact aactaattta gacgaatata ataagttaac tacaaacaat 120

aatgaaaata aatatttaca attaaaaaaa ctaaatgaat caattgatgt ttttatgaat 180

aaatataaaa cttcaagcag aaatagagca ctctctaatc taaaaaaaga tatattaaaa 240

gaagtaattc ttattaaaaa ttccaataca agccctgtag aaaaaaattt acattttgta 300

tggataggtg gagaagtcag tgatattgct cttgaataca taaaacaatg ggctgatatt 360

aatgcagaat ataatattaa actgtggtat gatagtgaag cattcttagt aaatacacta 420

aaaaaggcta tagttgaatc ttctaccact gaagcattac agctactaga ggaagagatt 480

caaaatcctc aatttgataa tatgaaattt tacaaaaaaa ggatggaatt tatatatgat 540

agacaaaaaa ggtttataaa ttattataaa tctcaaatca ataaacctac agtacctaca 600

atagatgata ttataaagtc tcatctagta tctgaatata atagagatga aactgtatta 660

gaatcatata gaacaaattc tttgagaaaa ataaatagta atcatgggat agatatcagg 720

gctaatagtt tgtttacaga acaagagtta ttaaatattt atagtcagga gttgttaaat 780

cgtggaaatt tagctgcagc atctgacata gtaagattat tagccctaaa aaattttggc 840

ggagtatatt tagatgttga tatgcttcca ggtattcact ctgatttatt taaaacaata 900

tctagaccta gctctattgg actagaccgt tgggaaatga taaaattaga ggctattatg 960

aagtataaaa aatatataaa taattataca tcagaaaact ttgataaact tgatcaacaa 1020

ttaaaagata attttaaact cattatagaa agtaaaagtg aaaaatctga gatattttct 1080

aaattagaaa atttaaatgt atctgatctt gaaattaaaa tagctttcgc tttaggcagt 1140

gttataaatc aagccttgat atcaaaacaa ggttcatatc ttactaacct agtaatagaa 1200

caagtaaaaa atagatatca atttttaaac caacacctta acccagccat agagtctgat 1260

aataacttca cagatactac taaaattttt catgattcat tatttaattc agctaccgca 1320

gaaaactcta tgtttttaac aaaaatagca ccatacttac aagtaggttt tatgccagaa 1380

gctcgctcca caataagttt aagtggtcca ggagcttatg cgtcagctta ctatgatttc 1440

ataaatttac aagaaaatac tatagaaaaa actttaaaag catcagattt aatagaattt 1500

aaattcccag aaaataatct atctcaattg acagaacaag aaataaatag tctatggagc 1560

tttgatcaag caagtgcaaa atatcaattt gagaaatatg taagagatta tactggtgga 1620

tctctttctg aagacaatgg ggtagacttt aataaaaata ctgccctcga caaaaactat 1680

ttattaaata ataaaattcc atcaaacaat gtagaagaag ctggaagtaa aaattatgtt 1740

cattatatca tacagttaca aggagatgat ataagttatg aagcaacatg caatttattt 1800

tctaaaaatc ctaaaaatag tattattata caacgaaata tgaatgaaag tgcaaaaagc 1860

tacttttaa gtgatgatgg agaatctatt ttagaattaa ataaatatag gatacctgaa 1920

agattaaaaa ataaggaaaa agtaaaagta acctttattg gacatggtaa agatgaattc 1980

aacacaagcg aatttgctag attaagtgta gattcacttt ccaatgagat aagttcattt 2040

ttagatacca taaaattaga tatatcacct aaaaatgtag aagtaaactt acttggatgt 2100

aatatgttta gttatgattt taatgttgaa gaaacttatc ctgggaagtt gctattaagt 2160

attatggaca aaattacttc cactttacct gatgtaaata aaaattctat tactatagga 2220

gcaaatcaat atgaagtaag aattaatagt gagggaagaa aagaacttct ggctcactca 2280

ggtaaatgga taaataaaga agaagctatt atgagcgatt tatctagtaa agaatacatt 2340

ttttttgatt ctatagataa taagctaaaa gcaaagtcca agaatattcc aggattagca 2400

tcaatatcag aagatataaa aacattatta cttgatgcaa gtgttagtcc tgatacaaaa 2460

tttattttaa ataatcttaa gcttaatatt gaatcttcta ttggtgatta catttattat 2520

gaaaaattag agcctgttaa aaatataatt cacaattcta tagatgattt aatagatgag 2580

ttcaatctac ttgaaaatgt atctgatgaa ttatatgaat taaaaaaatt aaataatcta 2640

gatgagaagt atttaatatc ttttgaagat atctcaaaaa ataattcaac ttactctgta 2700

agatttatta acaaaagtaa tggtgagtca gtttatgtag aaacagaaaa agaaattttt 2760

tcaaaatata gcgaacatat tacaaaagaa ataagtacta taaagaatag tataattaca 2820

gatgttaatg gtaatttatt ggataatata cagttagatc atacttctca agttaataca 2880

ttaaacgcag cattctttat tcaatcatta atagattata gtagcaataa agatgtactg 2940

aatgatttaa gtacctcagt taaggttcaa ctttatgctc aactatttag tacaggttta 3000

aatactatat atgactctat ccaattagta aatttaatat caaatgcagt aaatgatact 3060

ataaatgtac tacctacaat aacagagggg atacctattg tatctactat attagacgga 3120

ataaacttag gtgcagcaat taaggaatta ctagacgaac atgacccatt actaaaaaaa 3180

gaattagaag ctaaggtggg tgttttagca ataaatatgt cattatctat agctgcaact 3240

gtagcttcaa ttgttggaat aggtgctgaa gttactattt tcttattacc tatagctggt 3300

atatctgcag gaataccttc attagttaat aatgaattaa tattgcatga taaggcaact 3360

tcagtggtaa actattttaa tcatttgtct gaatctaaaa aatatggccc tcttaaaaca 3420

gaagatgata aaattttagt tcctattgat gatttagtaa tatcagaaat agattttaat 3480

aataattcga taaaactagg aacatgtaat atattagcaa tggaggggg atcaggacac 3540

acagtgactg gtaatataga tcacttttc tcatctccat ctataagttc tcatattcct 3600

tcattatcaa tttattctgc aataggtata gaaacagaaa atctagattt ttcaaaaaaa 3660

ataatgatgt tacctaatgc tccttcaaga gtgttttggt gggaaactgg agcagttcca 3720

ggtttaagat cattggaaaa tgacggaact agattacttg attcaataag agatttatac 3780

ccaggtaaat tttactggag attctatgct tttttcgatt atgcaataac tacattaaaa 3840

ccagtttatg aagacactaa tattaaaatt aaactagata aagatactag aaacttcata 3900

atgccaacta taactactaa cgaaattaga aacaaattat cttattcatt tgatggagca 3960

ggaggaactt actctttatt attatcttca tatccaatat caacgaatat aaatttatct 4020

aaagatgatt tatggatatt taatattgat aatgaagtaa gagaaatatc tatagaaaat 4080

ggtactatta aaaaaggaaa gttaataaaa gatgttttaa gtaaaattga tataaataaa 4140

aataaactta ttataggcaa tcaaacaata gatttttcag gcgatataga taataaagat 4200

agatatatat tcttgacttg tgagttagat gataaaatta gtttaataat agaaataaat 4260

cttgttgcaa aatcttatag tttgttattg tctggggata aaaattattt gatatccaat 4320

ttatctaata ttattgagaa aatcaatact ttaggcctag atagtaaaaa tatagcgtac 4380

aattacactg atgaatctaa taataaatat tttggagcta tatctaaaac aagtcaaaaa 4440

agcataatac attataaaaa agacagtaaa aatatattag aattttataa tgacagtaca 4500

ttagaattta acagtaaaga ttttattgct gaagatataa atgtatttat gaaagatgat 4560

attaatacta taacaggaaa atactatgtt gataataata ctgataaaag tatagatttc 4620

tctatttctt tagttagtaa aaatcaagta aaagtaaatg gattatattt aaatgaatcc 4680

gtatactcat cttaccttga ttttgtgaaa aattcagatg gacaccataa tacttctaat 4740

tttatgaatt tatttttgga caatataagt ttctggaaat tgtttgggtt tgaaaatata 4800

aattttgtaa tcgataaata ctttaccctt gttggtaaaa ctaatcttgg atatgtagaa 4860

tttatttgtg acaataataa aaatatagat atatattttg gtgaatggaa aacatcgtca 4920

tctaaaagca ctatatttag cggaaatggt agaaatgttg tagtagagcc tatatataat 4980

cctgatacgg gtgaagatat atctacttca ctagattttt cctatgaacc tctctatgga 5040

atagatagat atatcaataa agtattgata gcacctgatt tatatacaag tttaataaat 5100

attaatacca attattattc aaatgagtac taccctgaga ttatagttct taacccaaat 5160

acattccaca aaaaagtaaa tataaattta gatagttctt cttttgagta taaatggtct 5220

acagaaggaa gtgactttat tttagttaga tacttagaag aaagtaataa aaaaatatta 5280

caaaaaataa gaatcaaagg tatcttatct aatactcaat catttaataa aatgagtata 5340

gattttaaag atattaaaaa actatcatta ggatatataa tgagtaattt taaatcattt 5400

aattctgaaa atgaattaga tagagatcat ttaggattta aaataataga taataaaact 5460

tattactatg atgaagatag taaattagtt aaaggattaa tcaatataaa taattcatta 5520

ttctattttg atcctataga atttaactta gtaactggat ggcaaactat caatggtaaa 5580

aaatattatt ttgatataaa tactggagca gctttaatta gttataaaat tattaatggt 5640

aaacactttt attttaataa tgatggtgtg atgcagttgg gagtatttaa aggacctgat 5700

ggatttgaat attttgcacc tgccaatact caaaataata acatagaagg tcaggctata 5760

gtttatcaaa gtaaattctt aactttgaat ggcaaaaaat attattttga taatgactca 5820

aaagcagtca ctggatggag aattattaac aatgagaaat attactttaa tcctaataat 5880

gctattgctg cagtcggatt gcaagtaatt gacaataata agtattattt caatcctgac 5940

actgctatca tctcaaaagg ttggcagact gttaatggta gtagatacta ctttgatact 6000

gataccgcta ttgcctttaa tggttataaa actattgatg gtaaacactt ttattttgat 6060

agtgattgtg tagtgaaaat aggtgtgttt agtacctcta atggatttga atattttgca 6120

cctgctaata cttataataa taacatagaa ggtcaggcta tagtttatca aagtaaattc 6180

ttaactttga atggtaaaaa atattacttt gataataact caaaagcagt taccggatgg 6240

caaactattg atagtaaaaa atattacttt aatactaaca ctgctgaagc agctactgga 6300

tggcaaacta ttgatggtaa aaaatattac tttaatacta acactgctga agcagctact 6360

ggatggcaaa ctattgatgg taaaaaatat tactttaata ctaacactgc tatagcttca 6420

actggttata caattattaa tggtaaacat ttttatttta atactgatgg tattatgcag 6480

ataggagtgt ttaaaggacc taatggattt gaatattttg cacctgctaa tacggatgct 6540

aacaacatag aaggtcaagc tatactttac caaaatgaat tcttaacttt gaatggtaaa 6600

aaatattact ttggtagtga ctcaaaagca gttactggat ggagaattat taacaataag 6660

aaatattact ttaatcctaa taatgctatt gctgcaattc atctatgcac tataaataat 6720

gacaagtatt actttagtta tgatggaatt cttcaaaatg gatatattac tattgaaaga 6780

aataatttct attttgatgc taataatgaa tctaaaatgg taacaggagt atttaaagga 6840

cctaatggat ttgagtattt tgcacctgct aatactcaca ataataacat agaaggtcag 6900

gctatagttt accagaacaa attcttaact ttgaatggca aaaaatatta ttttgataat 6960

gactcaaaag cagttactgg atggcaaacc attgatggta aaaaatatta ctttaatctt 7020

aacactgctg aagcagctac tggatggcaa actattgatg gtaaaaaata ttactttaat 7080

cttaacactg ctgaagcagc tactggatgg caaactattg atggtaaaaa atattacttt 7140

aatactaaca ctttcatagc ctcaactggt tatacaagta ttaatggtaa acatttttat 7200

tttaatactg atggtattat gcagatagga gtgtttaaag gacctaatgg atttgaatac 7260

tttgcacctg ctaatactca taataataac atagaaggtc aagctatact ttaccaaaat 7320

aaattcttaa ctttgaatgg taaaaaatat tactttggta gtgactcaaa agcagttacc 7380

ggattgcgaa ctattgatgg taaaaaatat tactttaata ctaacactgc tgttgcagtt 7440

actggatggc aaactattaa tggtaaaaaa tactacttta atactaacac ttctatagct 7500

tcaactggtt atacaattat tagtggtaaa catttttatt ttaatactga tggtattatg 7560

cagataggag tgtttaaagg acctgatgga tttgaatact ttgcacctgc taatacagat 7620

gctaacaata tagaaggtca agctatacgt tatcaaaata gattcctata tttacatgac 7680

aatatatatt attttggtaa taattcaaaa gcagctactg gttgggtaac tattgatggt 7740

aatagatatt acttcgagcc taatacagct atgggtgcga atggttataa aactattgat 7800

aataaaaatt tttactttag aaatggttta cctcagatag gagtgtttaa agggtctaat 7860

ggatttgaat actttgcacc tgctaatacg gatgctaaca atatagaagg tcaagctata 7920

cgttatcaaa atagattcct acatttactt ggaaaaatat attactttgg taataattca 7980

aaagcagtta ctggatggca aactattaat ggtaaagtat attactttat gcctgatact 8040

gctatggctg cagctggtgg acttttcgag attgatggtg ttatatattt ctttggtgtt 8100

gatggagtaa aagcccctgg gatatatggc taa 8133

<210> 6
<211> 2710
<212> PRT
<213> Clostridium difficile strain 630

<400> 6

```
Met Ser Leu Ile Ser Lys Glu Glu Leu Ile Lys Leu Ala Tyr Ser Ile
1               5                  10                  15

Arg Pro Arg Glu Asn Glu Tyr Lys Thr Ile Leu Thr Asn Leu Asp Glu
            20                  25                  30

Tyr Asn Lys Leu Thr Thr Asn Asn Asn Glu Asn Lys Tyr Leu Gln Leu
        35                  40                  45

Lys Lys Leu Asn Glu Ser Ile Asp Val Phe Met Asn Lys Tyr Lys Thr
    50                  55                  60

Ser Ser Arg Asn Arg Ala Leu Ser Asn Leu Lys Lys Asp Ile Leu Lys
65                  70                  75                  80

Glu Val Ile Leu Ile Lys Asn Ser Asn Thr Ser Pro Val Glu Lys Asn
                85                  90                  95

Leu His Phe Val Trp Ile Gly Gly Glu Val Ser Asp Ile Ala Leu Glu
            100                 105                 110

Tyr Ile Lys Gln Trp Ala Asp Ile Asn Ala Glu Tyr Asn Ile Lys Leu
        115                 120                 125

Trp Tyr Asp Ser Glu Ala Phe Leu Val Asn Thr Leu Lys Lys Ala Ile
    130                 135                 140

Val Glu Ser Ser Thr Thr Glu Ala Leu Gln Leu Leu Glu Glu Glu Ile
145                 150                 155                 160

Gln Asn Pro Gln Phe Asp Asn Met Lys Phe Tyr Lys Lys Arg Met Glu
                165                 170                 175

Phe Ile Tyr Asp Arg Gln Lys Arg Phe Ile Asn Tyr Tyr Lys Ser Gln
            180                 185                 190

Ile Asn Lys Pro Thr Val Pro Thr Ile Asp Asp Ile Ile Lys Ser His
        195                 200                 205

Leu Val Ser Glu Tyr Asn Arg Asp Glu Thr Val Leu Glu Ser Tyr Arg
    210                 215                 220
```

Thr Asn Ser Leu Arg Lys Ile Asn Ser Asn His Gly Ile Asp Ile Arg
225 230 235 240

Ala Asn Ser Leu Phe Thr Glu Gln Glu Leu Leu Asn Ile Tyr Ser Gln
245 250 255

Glu Leu Leu Asn Arg Gly Asn Leu Ala Ala Ala Ser Asp Ile Val Arg
260 265 270

Leu Leu Ala Leu Lys Asn Phe Gly Gly Val Tyr Leu Asp Val Asp Met
275 280 285

Leu Pro Gly Ile His Ser Asp Leu Phe Lys Thr Ile Ser Arg Pro Ser
290 295 300

Ser Ile Gly Leu Asp Arg Trp Glu Met Ile Lys Leu Glu Ala Ile Met
305 310 315 320

Lys Tyr Lys Lys Tyr Ile Asn Asn Tyr Thr Ser Glu Asn Phe Asp Lys
325 330 335

Leu Asp Gln Gln Leu Lys Asp Asn Phe Lys Leu Ile Ile Glu Ser Lys
340 345 350

Ser Glu Lys Ser Glu Ile Phe Ser Lys Leu Glu Asn Leu Asn Val Ser
355 360 365

Asp Leu Glu Ile Lys Ile Ala Phe Ala Leu Gly Ser Val Ile Asn Gln
370 375 380

Ala Leu Ile Ser Lys Gln Gly Ser Tyr Leu Thr Asn Leu Val Ile Glu
385 390 395 400

Gln Val Lys Asn Arg Tyr Gln Phe Leu Asn Gln His Leu Asn Pro Ala
405 410 415

Ile Glu Ser Asp Asn Asn Phe Thr Asp Thr Thr Lys Ile Phe His Asp
420 425 430

Ser Leu Phe Asn Ser Ala Thr Ala Glu Asn Ser Met Phe Leu Thr Lys
435 440 445

Ile Ala Pro Tyr Leu Gln Val Gly Phe Met Pro Glu Ala Arg Ser Thr
450 455 460

Ile Ser Leu Ser Gly Pro Gly Ala Tyr Ala Ser Ala Tyr Tyr Asp Phe

```
465                 470                 475                 480

Ile Asn Leu Gln Glu Asn Thr Ile Glu Lys Thr Leu Lys Ala Ser Asp
                485                 490                 495

Leu Ile Glu Phe Lys Phe Pro Glu Asn Asn Leu Ser Gln Leu Thr Glu
            500                 505                 510

Gln Glu Ile Asn Ser Leu Trp Ser Phe Asp Gln Ala Ser Ala Lys Tyr
        515                 520                 525

Gln Phe Glu Lys Tyr Val Arg Asp Tyr Thr Gly Gly Ser Leu Ser Glu
    530                 535                 540

Asp Asn Gly Val Asp Phe Asn Lys Asn Thr Ala Leu Asp Lys Asn Tyr
545                 550                 555                 560

Leu Leu Asn Asn Lys Ile Pro Ser Asn Asn Val Glu Glu Ala Gly Ser
                565                 570                 575

Lys Asn Tyr Val His Tyr Ile Ile Gln Leu Gln Gly Asp Asp Ile Ser
            580                 585                 590

Tyr Glu Ala Thr Cys Asn Leu Phe Ser Lys Asn Pro Lys Asn Ser Ile
        595                 600                 605

Ile Ile Gln Arg Asn Met Asn Glu Ser Ala Lys Ser Tyr Phe Leu Ser
    610                 615                 620

Asp Asp Gly Glu Ser Ile Leu Glu Leu Asn Lys Tyr Arg Ile Pro Glu
625                 630                 635                 640

Arg Leu Lys Asn Lys Glu Lys Val Lys Val Thr Phe Ile Gly His Gly
                645                 650                 655

Lys Asp Glu Phe Asn Thr Ser Glu Phe Ala Arg Leu Ser Val Asp Ser
            660                 665                 670

Leu Ser Asn Glu Ile Ser Ser Phe Leu Asp Thr Ile Lys Leu Asp Ile
        675                 680                 685

Ser Pro Lys Asn Val Glu Val Asn Leu Leu Gly Cys Asn Met Phe Ser
    690                 695                 700

Tyr Asp Phe Asn Val Glu Glu Thr Tyr Pro Gly Lys Leu Leu Leu Ser
705                 710                 715                 720
```

```
Ile Met Asp Lys Ile Thr Ser Thr Leu Pro Asp Val Asn Lys Asn Ser
                725                 730                 735

Ile Thr Ile Gly Ala Asn Gln Tyr Glu Val Arg Ile Asn Ser Glu Gly
            740                 745                 750

Arg Lys Glu Leu Leu Ala His Ser Gly Lys Trp Ile Asn Lys Glu Glu
        755                 760                 765

Ala Ile Met Ser Asp Leu Ser Ser Lys Glu Tyr Ile Phe Phe Asp Ser
    770                 775                 780

Ile Asp Asn Lys Leu Lys Ala Lys Ser Lys Asn Ile Pro Gly Leu Ala
785                 790                 795                 800

Ser Ile Ser Glu Asp Ile Lys Thr Leu Leu Leu Asp Ala Ser Val Ser
                805                 810                 815

Pro Asp Thr Lys Phe Ile Leu Asn Asn Leu Lys Leu Asn Ile Glu Ser
            820                 825                 830

Ser Ile Gly Asp Tyr Ile Tyr Tyr Glu Lys Leu Glu Pro Val Lys Asn
        835                 840                 845

Ile Ile His Asn Ser Ile Asp Asp Leu Ile Asp Glu Phe Asn Leu Leu
    850                 855                 860

Glu Asn Val Ser Asp Glu Leu Tyr Glu Leu Lys Lys Leu Asn Asn Leu
865                 870                 875                 880

Asp Glu Lys Tyr Leu Ile Ser Phe Glu Asp Ile Ser Lys Asn Asn Ser
                885                 890                 895

Thr Tyr Ser Val Arg Phe Ile Asn Lys Ser Asn Gly Glu Ser Val Tyr
            900                 905                 910

Val Glu Thr Glu Lys Glu Ile Phe Ser Lys Tyr Ser Glu His Ile Thr
        915                 920                 925

Lys Glu Ile Ser Thr Ile Lys Asn Ser Ile Ile Thr Asp Val Asn Gly
    930                 935                 940

Asn Leu Leu Asp Asn Ile Gln Leu Asp His Thr Ser Gln Val Asn Thr
945                 950                 955                 960

Leu Asn Ala Ala Phe Phe Ile Gln Ser Leu Ile Asp Tyr Ser Ser Asn
                965                 970                 975
```

```
Lys Asp Val Leu Asn Asp Leu Ser Thr Ser Val Lys Val Gln Leu Tyr
            980                 985                 990

Ala Gln Leu Phe Ser Thr Gly Leu  Asn Thr Ile Tyr Asp  Ser Ile Gln
        995                 1000                1005

Leu Val  Asn Leu Ile Ser Asn  Ala Val Asn Asp Thr  Ile Asn Val
    1010                1015                1020

Leu Pro  Thr Ile Thr Glu Gly  Ile Pro Ile Val Ser  Thr Ile Leu
    1025                1030                1035

Asp Gly  Ile Asn Leu Gly Ala  Ala Ile Lys Glu Leu  Leu Asp Glu
    1040                1045                1050

His Asp  Pro Leu Leu Lys Lys  Glu Leu Glu Ala Lys  Val Gly Val
    1055                1060                1065

Leu Ala  Ile Asn Met Ser Leu  Ser Ile Ala Ala Thr  Val Ala Ser
    1070                1075                1080

Ile Val  Gly Ile Gly Ala Glu  Val Thr Ile Phe Leu  Leu Pro Ile
    1085                1090                1095

Ala Gly  Ile Ser Ala Gly Ile  Pro Ser Leu Val Asn  Asn Glu Leu
    1100                1105                1110

Ile Leu  His Asp Lys Ala Thr  Ser Val Val Asn Tyr  Phe Asn His
    1115                1120                1125

Leu Ser  Glu Ser Lys Lys Tyr  Gly Pro Leu Lys Thr  Glu Asp Asp
    1130                1135                1140

Lys Ile  Leu Val Pro Ile Asp  Asp Leu Val Ile Ser  Glu Ile Asp
    1145                1150                1155

Phe Asn  Asn Asn Ser Ile Lys  Leu Gly Thr Cys Asn  Ile Leu Ala
    1160                1165                1170

Met Glu  Gly Gly Ser Gly His  Thr Val Thr Gly Asn  Ile Asp His
    1175                1180                1185

Phe Phe  Ser Ser Pro Ser Ile  Ser Ser His Ile Pro  Ser Leu Ser
    1190                1195                1200

Ile Tyr  Ser Ala Ile Gly Ile  Glu Thr Glu Asn Leu  Asp Phe Ser
    1205                1210                1215
```

```
Lys Lys  Ile Met Met Leu Pro  Asn Ala Pro Ser Arg  Val Phe Trp
    1220                 1225                 1230

Trp Glu  Thr Gly Ala Val Pro  Gly Leu Arg Ser Leu  Glu Asn Asp
    1235                 1240                 1245

Gly Thr  Arg Leu Leu Asp Ser  Ile Arg Asp Leu Tyr  Pro Gly Lys
    1250                 1255                 1260

Phe Tyr  Trp Arg Phe Tyr Ala  Phe Phe Asp Tyr Ala  Ile Thr Thr
    1265                 1270                 1275

Leu Lys  Pro Val Tyr Glu Asp  Thr Asn Ile Lys Ile  Lys Leu Asp
    1280                 1285                 1290

Lys Asp  Thr Arg Asn Phe Ile  Met Pro Thr Ile Thr  Thr Asn Glu
    1295                 1300                 1305

Ile Arg  Asn Lys Leu Ser Tyr  Ser Phe Asp Gly Ala  Gly Gly Thr
    1310                 1315                 1320

Tyr Ser  Leu Leu Leu Ser Ser  Tyr Pro Ile Ser Thr  Asn Ile Asn
    1325                 1330                 1335

Leu Ser  Lys Asp Asp Leu Trp  Ile Phe Asn Ile Asp  Asn Glu Val
    1340                 1345                 1350

Arg Glu  Ile Ser Ile Glu Asn  Gly Thr Ile Lys Lys  Gly Lys Leu
    1355                 1360                 1365

Ile Lys  Asp Val Leu Ser Lys  Ile Asp Ile Asn Lys  Asn Lys Leu
    1370                 1375                 1380

Ile Ile  Gly Asn Gln Thr Ile  Asp Phe Ser Gly Asp  Ile Asp Asn
    1385                 1390                 1395

Lys Asp  Arg Tyr Ile Phe Leu  Thr Cys Glu Leu Asp  Asp Lys Ile
    1400                 1405                 1410

Ser Leu  Ile Ile Glu Ile Asn  Leu Val Ala Lys Ser  Tyr Ser Leu
    1415                 1420                 1425

Leu Leu  Ser Gly Asp Lys Asn  Tyr Leu Ile Ser Asn  Leu Ser Asn
    1430                 1435                 1440

Ile Ile  Glu Lys Ile Asn Thr  Leu Gly Leu Asp Ser  Lys Asn Ile
```

```
    1445                1450                1455

Ala Tyr  Asn Tyr Thr Asp Glu  Ser Asn Asn Lys Tyr  Phe Gly Ala
    1460                1465                1470

Ile Ser  Lys Thr Ser Gln Lys  Ser Ile Ile His Tyr  Lys Lys Asp
    1475                1480                1485

Ser Lys  Asn Ile Leu Glu Phe  Tyr Asn Asp Ser Thr  Leu Glu Phe
    1490                1495                1500

Asn Ser  Lys Asp Phe Ile Ala  Glu Asp Ile Asn Val  Phe Met Lys
    1505                1510                1515

Asp Asp  Ile Asn Thr Ile Thr  Gly Lys Tyr Tyr Val  Asp Asn Asn
    1520                1525                1530

Thr Asp  Lys Ser Ile Asp Phe  Ser Ile Ser Leu Val  Ser Lys Asn
    1535                1540                1545

Gln Val  Lys Val Asn Gly Leu  Tyr Leu Asn Glu Ser  Val Tyr Ser
    1550                1555                1560

Ser Tyr  Leu Asp Phe Val Lys  Asn Ser Asp Gly His  His Asn Thr
    1565                1570                1575

Ser Asn  Phe Met Asn Leu Phe  Leu Asp Asn Ile Ser  Phe Trp Lys
    1580                1585                1590

Leu Phe  Gly Phe Glu Asn Ile  Asn Phe Val Ile Asp  Lys Tyr Phe
    1595                1600                1605

Thr Leu  Val Gly Lys Thr Asn  Leu Gly Tyr Val Glu  Phe Ile Cys
    1610                1615                1620

Asp Asn  Asn Lys Asn Ile Asp  Ile Tyr Phe Gly Glu  Trp Lys Thr
    1625                1630                1635

Ser Ser  Ser Lys Ser Thr Ile  Phe Ser Gly Asn Gly  Arg Asn Val
    1640                1645                1650

Val Val  Glu Pro Ile Tyr Asn  Pro Asp Thr Gly Glu  Asp Ile Ser
    1655                1660                1665

Thr Ser  Leu Asp Phe Ser Tyr  Glu Pro Leu Tyr Gly  Ile Asp Arg
    1670                1675                1680
```

```
Tyr Ile  Asn Lys Val Leu Ile  Ala Pro Asp Leu Tyr  Thr Ser Leu
    1685                1690                1695

Ile Asn  Ile Asn Thr Asn Tyr  Tyr Ser Asn Glu Tyr  Tyr Pro Glu
    1700                1705                1710

Ile Ile  Val Leu Asn Pro Asn  Thr Phe His Lys Lys  Val Asn Ile
    1715                1720                1725

Asn Leu  Asp Ser Ser Ser Phe  Glu Tyr Lys Trp Ser  Thr Glu Gly
    1730                1735                1740

Ser Asp  Phe Ile Leu Val Arg  Tyr Leu Glu Glu Ser  Asn Lys Lys
    1745                1750                1755

Ile Leu  Gln Lys Ile Arg Ile  Lys Gly Ile Leu Ser  Asn Thr Gln
    1760                1765                1770

Ser Phe  Asn Lys Met Ser Ile  Asp Phe Lys Asp Ile  Lys Lys Leu
    1775                1780                1785

Ser Leu  Gly Tyr Ile Met Ser  Asn Phe Lys Ser Phe  Asn Ser Glu
    1790                1795                1800

Asn Glu  Leu Asp Arg Asp His  Leu Gly Phe Lys Ile  Ile Asp Asn
    1805                1810                1815

Lys Thr  Tyr Tyr Tyr Asp Glu  Asp Ser Lys Leu Val  Lys Gly Leu
    1820                1825                1830

Ile Asn  Ile Asn Asn Ser Leu  Phe Tyr Phe Asp Pro  Ile Glu Phe
    1835                1840                1845

Asn Leu  Val Thr Gly Trp Gln  Thr Ile Asn Gly Lys  Lys Tyr Tyr
    1850                1855                1860

Phe Asp  Ile Asn Thr Gly Ala  Ala Leu Ile Ser Tyr  Lys Ile Ile
    1865                1870                1875

Asn Gly  Lys His Phe Tyr Phe  Asn Asn Asp Gly Val  Met Gln Leu
    1880                1885                1890

Gly Val  Phe Lys Gly Pro Asp  Gly Phe Glu Tyr Phe  Ala Pro Ala
    1895                1900                1905

Asn Thr  Gln Asn Asn Asn Ile  Glu Gly Gln Ala Ile  Val Tyr Gln
    1910                1915                1920
```

```
Ser Lys  Phe Leu Thr Leu Asn  Gly Lys Lys Tyr Tyr  Phe Asp Asn
    1925                 1930                 1935

Asp Ser  Lys Ala Val Thr Gly  Trp Arg Ile Ile Asn  Asn Glu Lys
    1940                 1945                 1950

Tyr Tyr  Phe Asn Pro Asn Asn  Ala Ile Ala Ala Val  Gly Leu Gln
    1955                 1960                 1965

Val Ile  Asp Asn Asn Lys Tyr  Tyr Phe Asn Pro Asp  Thr Ala Ile
    1970                 1975                 1980

Ile Ser  Lys Gly Trp Gln Thr  Val Asn Gly Ser Arg  Tyr Tyr Phe
    1985                 1990                 1995

Asp Thr  Asp Thr Ala Ile Ala  Phe Asn Gly Tyr Lys  Thr Ile Asp
    2000                 2005                 2010

Gly Lys  His Phe Tyr Phe Asp  Ser Asp Cys Val Val  Lys Ile Gly
    2015                 2020                 2025

Val Phe  Ser Thr Ser Asn Gly  Phe Glu Tyr Phe Ala  Pro Ala Asn
    2030                 2035                 2040

Thr Tyr  Asn Asn Asn Ile Glu  Gly Gln Ala Ile Val  Tyr Gln Ser
    2045                 2050                 2055

Lys Phe  Leu Thr Leu Asn Gly  Lys Lys Tyr Tyr Phe  Asp Asn Asn
    2060                 2065                 2070

Ser Lys  Ala Val Thr Gly Trp  Gln Thr Ile Asp Ser  Lys Lys Tyr
    2075                 2080                 2085

Tyr Phe  Asn Thr Asn Thr Ala  Glu Ala Ala Thr Gly  Trp Gln Thr
    2090                 2095                 2100

Ile Asp  Gly Lys Lys Tyr Tyr  Phe Asn Thr Asn Thr  Ala Glu Ala
    2105                 2110                 2115

Ala Thr  Gly Trp Gln Thr Ile  Asp Gly Lys Lys Tyr  Tyr Phe Asn
    2120                 2125                 2130

Thr Asn  Thr Ala Ile Ala Ser  Thr Gly Tyr Thr Ile  Ile Asn Gly
    2135                 2140                 2145

Lys His  Phe Tyr Phe Asn Thr  Asp Gly Ile Met Gln  Ile Gly Val
    2150                 2155                 2160
```

```
Phe Lys  Gly Pro Asn Gly Phe  Glu Tyr Phe Ala Pro  Ala Asn Thr
    2165                 2170                 2175

Asp Ala  Asn Asn Ile Glu Gly  Gln Ala Ile Leu Tyr  Gln Asn Glu
    2180                 2185                 2190

Phe Leu  Thr Leu Asn Gly Lys  Lys Tyr Tyr Phe Gly  Ser Asp Ser
    2195                 2200                 2205

Lys Ala  Val Thr Gly Trp Arg  Ile Ile Asn Asn Lys  Lys Tyr Tyr
    2210                 2215                 2220

Phe Asn  Pro Asn Asn Ala Ile  Ala Ala Ile His Leu  Cys Thr Ile
    2225                 2230                 2235

Asn Asn  Asp Lys Tyr Tyr Phe  Ser Tyr Asp Gly Ile  Leu Gln Asn
    2240                 2245                 2250

Gly Tyr  Ile Thr Ile Glu Arg  Asn Asn Phe Tyr Phe  Asp Ala Asn
    2255                 2260                 2265

Asn Glu  Ser Lys Met Val Thr  Gly Val Phe Lys Gly  Pro Asn Gly
    2270                 2275                 2280

Phe Glu  Tyr Phe Ala Pro Ala  Asn Thr His Asn Asn  Asn Ile Glu
    2285                 2290                 2295

Gly Gln  Ala Ile Val Tyr Gln  Asn Lys Phe Leu Thr  Leu Asn Gly
    2300                 2305                 2310

Lys Lys  Tyr Tyr Phe Asp Asn  Asp Ser Lys Ala Val  Thr Gly Trp
    2315                 2320                 2325

Gln Thr  Ile Asp Gly Lys Lys  Tyr Tyr Phe Asn Leu  Asn Thr Ala
    2330                 2335                 2340

Glu Ala  Ala Thr Gly Trp Gln  Thr Ile Asp Gly Lys  Lys Tyr Tyr
    2345                 2350                 2355

Phe Asn  Leu Asn Thr Ala Glu  Ala Ala Thr Gly Trp  Gln Thr Ile
    2360                 2365                 2370

Asp Gly  Lys Lys Tyr Tyr Phe  Asn Thr Asn Thr Phe  Ile Ala Ser
    2375                 2380                 2385

Thr Gly  Tyr Thr Ser Ile Asn  Gly Lys His Phe Tyr  Phe Asn Thr
```

```
    2390                2395                2400

Asp Gly  Ile Met Gln Ile Gly  Val Phe Lys Gly Pro  Asn Gly Phe
    2405                2410                2415

Glu Tyr  Phe Ala Pro Ala Asn  Thr His Asn Asn Asn  Ile Glu Gly
    2420                2425                2430

Gln Ala  Ile Leu Tyr Gln Asn  Lys Phe Leu Thr Leu  Asn Gly Lys
    2435                2440                2445

Lys Tyr  Tyr Phe Gly Ser Asp  Ser Lys Ala Val Thr  Gly Leu Arg
    2450                2455                2460

Thr Ile  Asp Gly Lys Lys Tyr  Tyr Phe Asn Thr Asn  Thr Ala Val
    2465                2470                2475

Ala Val  Thr Gly Trp Gln Thr  Ile Asn Gly Lys Lys  Tyr Tyr Phe
    2480                2485                2490

Asn Thr  Asn Thr Ser Ile Ala  Ser Thr Gly Tyr Thr  Ile Ile Ser
    2495                2500                2505

Gly Lys  His Phe Tyr Phe Asn  Thr Asp Gly Ile Met  Gln Ile Gly
    2510                2515                2520

Val Phe  Lys Gly Pro Asp Gly  Phe Glu Tyr Phe Ala  Pro Ala Asn
    2525                2530                2535

Thr Asp  Ala Asn Asn Ile Glu  Gly Gln Ala Ile Arg  Tyr Gln Asn
    2540                2545                2550

Arg Phe  Leu Tyr Leu His Asp  Asn Ile Tyr Tyr Phe  Gly Asn Asn
    2555                2560                2565

Ser Lys  Ala Ala Thr Gly Trp  Val Thr Ile Asp Gly  Asn Arg Tyr
    2570                2575                2580

Tyr Phe  Glu Pro Asn Thr Ala  Met Gly Ala Asn Gly  Tyr Lys Thr
    2585                2590                2595

Ile Asp  Asn Lys Asn Phe Tyr  Phe Arg Asn Gly Leu  Pro Gln Ile
    2600                2605                2610

Gly Val  Phe Lys Gly Ser Asn  Gly Phe Glu Tyr Phe  Ala Pro Ala
    2615                2620                2625
```

```
Asn Thr  Asp Ala Asn Asn Ile  Glu Gly Gln Ala Ile  Arg Tyr Gln
    2630                 2635                 2640

Asn Arg  Phe Leu His Leu Leu  Gly Lys Ile Tyr Tyr  Phe Gly Asn
    2645                 2650                 2655

Asn Ser  Lys Ala Val Thr Gly  Trp Gln Thr Ile Asn  Gly Lys Val
    2660                 2665                 2670

Tyr Tyr  Phe Met Pro Asp Thr  Ala Met Ala Ala Ala  Gly Gly Leu
    2675                 2680                 2685

Phe Glu  Ile Asp Gly Val Ile  Tyr Phe Phe Gly Val  Asp Gly Val
    2690                 2695                 2700

Lys Ala  Pro Gly Ile Tyr Gly
    2705                 2710
```

<210> 7
<211> 7101
<212> DNA
<213> Clostridium difficile strain 630

<400> 7

atgagtttag ttaatagaaa acagttagaa aaaatggcaa atgtaagatt tcgtactcaa 60

gaagatgaat atgttgcaat attggatgct ttagaagaat atcataatat gtcagagaat 120

actgtagtcg aaaaatattt aaaattaaaa gatataaata gtttaacaga tatttatata 180

gatacatata aaaaatctgg tagaaataaa gccttaaaaa aatttaagga atatctagtt 240

acagaagtat tagagctaaa gaataataat ttaactccag ttgagaaaaa tttacatttt 300

gtttggattg gaggtcaaat aaatgacact gctattaatt atataaatca atggaaagat 360

gtaaatagtg attataatgt taatgttttt tatgatagta atgcattttt gataaacaca 420

ttgaaaaaaa ctgtagtaga atcagcaata aatgatacac ttgaatcatt tagagaaaac 480

ttaaatgacc ctagatttga ctataataaa ttcttcagaa aacgtatgga aataatttat 540

gataaacaga aaaatttcat aaactactat aaagctcaaa gagaagaaaa tcctgaactt 600

ataattgatg atattgtaaa gacatatctt tcaaatgagt attcaaagga gatagatgaa 660

cttaatacct atattgaaga atccttaaat aaaattacac agaatagtgg aaatgatgtt 720

agaaactttg aagaatttaa aaatggagag tcattcaact tatatgaaca agagttggta 780

gaaaggtgga atttagctgc tgcttctgac atattaagaa tatctgcatt aaaagaaatt 840

ggtggtatgt atttagatgt tgatatgtta ccaggaatac aaccagactt atttgagtct 900

atagagaaac ctagttcagt aacagtggat ttttgggaaa tgacaaagtt agaagctata 960

atgaaataca aagaatatat accagaatat acctcagaac attttgacat gttagacgaa 1020

gaagttcaaa gtagttttga atctgttcta gcttctaagt cagataaatc agaaatattc 1080

tcatcacttg gtgatatgga ggcatcacca ctagaagtta aaattgcatt taatagtaag 1140

ggtattataa atcaagggct aatttctgtg aaagactcat attgtagcaa tttaatagta 1200

aaacaaatcg agaatagata taaaatattg aataatagtt taaatccagc tattagcgag 1260

gataatgatt ttaatactac aacgaatacc tttattgata gtataatggc tgaagctaat 1320

gcagataatg gtagatttat gatggaacta ggaaagtatt taagagttgg tttcttccca 1380

gatgttaaaa ctactattaa cttaagtggc cctgaagcat atgcggcagc ttatcaagat 1440

ttattaatgt ttaaagaagg cagtatgaat atccatttga tagaagctga tttaagaaac 1500

tttgaaatct ctaaaactaa tatttctcaa tcaactgaac aagaaatggc tagcttatgg 1560

tcatttgacg atgcaagagc taaagctcaa tttgaagaat ataaaaggaa ttattttgaa 1620

ggttctcttg gtgaagatga taatcttgat ttttctcaaa atatagtagt tgacaaggag 1680

tatcttttag aaaaaatatc ttcattagca agaagttcag agagaggata tatacactat 1740

attgttcagt tacaaggaga taaaattagt tatgaagcag catgtaactt atttgcaaag 1800

actccttatg atagtgtact gtttcagaaa aatatagaag attcagaaat tgcatattat 1860

tataatcctg gagatggtga aatacaagaa atagacaagt ataaaattcc aagtataatt 1920

tctgatagac ctaagattaa attaacattt attggtcatg gtaaagatga atttaatact 1980

gatatatttg caggttttga tgtagattca ttatccacag aaatagaagc agcaatagat 2040

ttagctaaag aggatatttc tcctaagtca atagaaataa atttattagg atgtaatatg 2100

tttagctact ctatcaacgt agaggagact tatcctggaa aattattact taaagttaaa 2160

gataaaatat cagaattaat gccatctata agtcaagact ctattatagt aagtgcaaat 2220

caatatgaag ttagaataaa tagtgaagga agaagagaat tattggatca ttctggtgaa 2280

tggataaata aagaagaaag tattataaag gatatttcat caaaagaata tatatcattt 2340

aatcctaaag aaaataaaat tacagtaaaa tctaaaaatt tacctgagct atctacatta 2400

ttacaagaaa ttagaaataa ttctaattca agtgatattg aactagaaga aaaagtaatg 2460

ttaacagaat gtgagataaa tgttatttca aatatagata cgcaaattgt tgaggaaagg 2520

attgaagaag ctaagaattt aacttctgac tctattaatt atataaaaga tgaatttaaa 2580

ctaatagaat ctatttctga tgcactatgt gacttaaaac aacagaatga attagaagat 2640

tctcatttta tatcttttga ggacatatca gagactgatg agggatttag tataagattt 2700

attaataaag aaactggaga atctatattt gtagaaactg aaaaaacaat attctctgaa 2760

tatgctaatc atataactga agagatttct aagataaaag gtactatatt tgatactgta 2820

aatggtaagt tagtaaaaaa agtaaattta gatactacac acgaagtaaa tactttaaat 2880

gctgcatttt ttatacaatc attaatagaa tataatagtt ctaaagaatc tcttagtaat 2940

ttaagtgtag caatgaaagt ccaagtttac gctcaattat ttagtactgg tttaaatact 3000

attacagatg cagccaaagt tgttgaatta gtatcaactg cattagatga aactatagac 3060

ttacttccta cattatctga aggattacct ataattgcaa ctattataga tggtgtaagt 3120

ttaggtgcag caatcaaaga gctaagtgaa acgagtgacc cattattaag acaagaaata 3180

gaagctaaga taggtataat ggcagtaaat ttaacaacag ctacaactgc aatcattact 3240

tcatctttgg ggatagctag tggatttagt atacttttag ttcctttagc aggaatttca 3300

gcaggtatac caagcttagt aaacaatgaa cttgtacttc gagataaggc aacaaaggtt 3360

gtagattatt ttaaacatgt ttcattagtt gaaactgaag gagtatttac tttattagat 3420

gataaaataa tgatgccaca agatgattta gtgatatcag aaatagattt taataataat 3480

tcaatagttt taggtaaatg tgaaatctgg agaatggaag gtggttcagg tcatactgta 3540

actgatgata tagatcactt cttttcagca ccatcaataa catatagaga gccacactta 3600

tctatatatg acgtattgga agtacaaaaa gaagaacttg atttgtcaaa agatttaatg 3660

gtattaccta atgctccaaa tagagtattt gcttgggaaa caggatggac accaggttta 3720

agaagcttag aaaatgatgg cacaaaactg ttagaccgta taagagataa ctatgaaggt 3780

gagttttatt ggagatattt tgcttttata gctgatgctt taataacaac attaaaacca 3840

agatatgaag atactaatat aagaataaat ttagatagta atactagaag ttttatagtt 3900

ccaataataa ctacagaata tataagagaa aaattatcat attctttcta tggttcagga 3960

ggaacttatg cattgtctct ttctcaatat aatatgggta taaatataga attaagtgaa 4020

agtgatgttt ggattataga tgttgataat gttgtgagag atgtaactat agaatctgat 4080

aaaattaaaa aaggtgattt aatagaaggt attttatcta cactaagtat tgaagagaat 4140

aaaattatct taaatagcca tgagattaat ttttctggtg aggtaaatgg aagtaatgga 4200

tttgtttctt taacattttc aattttagaa ggaataaatg caattataga agttgattta 4260

ttatctaaat catataaatt acttatttct ggcgaattaa aaatattgat gttaaattca 4320

aatcatattc aacagaaaat agattatata ggattcaata gcgaattaca gaaaaatata 4380

ccatatagct ttgtagatag tgaaggaaaa gagaatggtt ttattaatgg ttcaacaaaa 4440

gaaggtttat ttgtatctga attacctgat gtagttctta taagtaaggt ttatatggat 4500

gatagtaagc cttcatttgg atattatagt aataatttga aagatgtcaa agttataact 4560

aaagataatg ttaatatatt aacaggttat tatcttaagg atgatataaa aatctctctt 4620

tctttgactc tacaagatga aaaaactata aagttaaata gtgtgcattt agatgaaagt 4680

ggagtagctg agattttgaa gttcatgaat agaaaaggta atacaaatac ttcagattct 4740

ttaatgagct ttttagaaag tatgaatata aaaagtattt tcgttaattt cttacaatct 4800

aatattaagt ttatattaga tgctaatttt ataataagtg gtactacttc tattggccaa 4860

tttgagttta tttgtgatga aaatgataat atacaaccat atttcattaa gtttaataca 4920

ctagaaacta attatacttt atatgtagga aatagacaaa atatgatagt ggaaccaaat 4980

tatgatttag atgattctgg agatatatct tcaactgtta tcaatttctc tcaaaagtat 5040

ctttatggaa tagacagttg tgttaataaa gttgtaattt caccaaatat ttatacagat 5100

gaaataaata taacgcctgt atatgaaaca aataatactt atccagaagt tattgtatta 5160

gatgcaaatt atataaatga aaaaataaat gttaatatca atgatctatc tatacgatat 5220

gtatggagta atgatggtaa tgattttatt cttatgtcaa ctagtgaaga aaataaggtg 5280

tcacaagtta aaataagatt cgttaatgtt tttaaagata agactttggc aaataagcta 5340

tcttttaact ttagtgataa acaagatgta cctgtaagtg aaataatctt atcatttaca 5400

ccttcatatt atgaggatgg attgattggc tatgatttgg gtctagtttc tttatataat 5460

gagaaatttt atattaataa ctttggaatg atggtatctg gattaatata tattaatgat 5520

tcattatatt attttaaacc accagtaaat aatttgataa ctggatttgt gactgtaggc 5580

gatgataaat actactttaa tccaattaat ggtggagctg cttcaattgg agagacaata 5640

attgatgaca aaaattatta tttcaaccaa agtggagtgt tacaaacagg tgtatttagt 5700

acagaagatg gatttaaata ttttgcccca gctaatacac ttgatgaaaa cctagaagga 5760

gaagcaattg attttactgg aaaattaatt attgacgaaa atatttatta ttttgatgat 5820

aattatagag gagctgtaga atggaaagaa ttagatggtg aaatgcacta ttttagccca 5880

gaaacaggta aagcttttaa aggtctaaat caaataggtg attataaata ctatttcaat 5940

tctgatggag ttatgcaaaa aggatttgtt agtataaatg ataataaaca ctattttgat 6000

gattctggtg ttatgaaagt aggttacact gaaatagatg gcaagcattt ctactttgct 6060

gaaaacggag aaatgcaaat aggagtattt aatacagaag atggatttaa atattttgct 6120

catcataatg aagatttagg aaatgaagaa ggtgaagaaa tctcatattc tggtatatta 6180

aatttcaata ataaaattta ctattttgat gattcattta cagctgtagt tggatggaaa 6240

gatttagagg atggttcaaa gtattatttt gatgaagata cagcagaagc atatataggt 6300

ttgtcattaa taaatgatgg tcaatattat tttaatgatg atggaattat gcaagttgga 6360

tttgtcacta taaatgataa agtcttctac ttctctgact ctggaattat agaatctgga 6420

gtacaaaaca tagatgacaa ttatttctat atagatgata atggtatagt tcaaattggt 6480

gtatttgata cttcagatgg atataaatat tttgcacctg ctaatactgt aaatgataat 6540

atttacggac aagcagttga atatagtggt ttagttagag ttggtgaaga tgtatattat 6600

tttggagaaa catatacaat tgagactgga tggatatatg atatggaaaa tgaaagtgat 6660

aaatattatt tcaatccaga aactaaaaaa gcatgcaaag gtattaattt aattgatgat 6720

ataaaatatt attttgatga gaagggcata atgagaacgg gtcttatatc atttgaaaat 6780

aataattatt actttaatga gaatggtgaa atgcaatttg gttatataaa tatagaagat 6840

aagatgttct attttggtga agatggtgtc atgcagattg gagtatttaa tacaccagat 6900

ggatttaaat actttgcaca tcaaaatact ttggatgaga attttgaggg agaatcaata 6960

aactatactg gttggttaga tttagatgaa aagagatatt attttacaga tgaatatatt 7020

gcagcaactg gttcagttat tattgatggt gaggagtatt attttgatcc tgatacagct 7080

caattagtga ttagtgaata g 7101

<210> 8
<211> 2366
<212> PRT
<213> Clostridium difficile strain 630

<400> 8

```
Met Ser Leu Val Asn Arg Lys Gln Leu Glu Lys Met Ala Asn Val Arg
1               5                  10                  15

Phe Arg Thr Gln Glu Asp Glu Tyr Val Ala Ile Leu Asp Ala Leu Glu
            20                  25                  30

Glu Tyr His Asn Met Ser Glu Asn Thr Val Val Glu Lys Tyr Leu Lys
        35                  40                  45

Leu Lys Asp Ile Asn Ser Leu Thr Asp Ile Tyr Ile Asp Thr Tyr Lys
    50                  55                  60

Lys Ser Gly Arg Asn Lys Ala Leu Lys Lys Phe Lys Glu Tyr Leu Val
65                  70                  75                  80

Thr Glu Val Leu Glu Leu Lys Asn Asn Asn Leu Thr Pro Val Glu Lys
                85                  90                  95

Asn Leu His Phe Val Trp Ile Gly Gly Gln Ile Asn Asp Thr Ala Ile
            100                 105                 110

Asn Tyr Ile Asn Gln Trp Lys Asp Val Asn Ser Asp Tyr Asn Val Asn
        115                 120                 125

Val Phe Tyr Asp Ser Asn Ala Phe Leu Ile Asn Thr Leu Lys Lys Thr
    130                 135                 140

Val Val Glu Ser Ala Ile Asn Asp Thr Leu Glu Ser Phe Arg Glu Asn
145                 150                 155                 160
```

```
Leu Asn Asp Pro Arg Phe Asp Tyr Asn Lys Phe Phe Arg Lys Arg Met
                165                 170                 175

Glu Ile Ile Tyr Asp Lys Gln Lys Asn Phe Ile Asn Tyr Tyr Lys Ala
            180                 185                 190

Gln Arg Glu Glu Asn Pro Glu Leu Ile Ile Asp Asp Ile Val Lys Thr
        195                 200                 205

Tyr Leu Ser Asn Glu Tyr Ser Lys Glu Ile Asp Glu Leu Asn Thr Tyr
    210                 215                 220

Ile Glu Glu Ser Leu Asn Lys Ile Thr Gln Asn Ser Gly Asn Asp Val
225                 230                 235                 240

Arg Asn Phe Glu Glu Phe Lys Asn Gly Glu Ser Phe Asn Leu Tyr Glu
                245                 250                 255

Gln Glu Leu Val Glu Arg Trp Asn Leu Ala Ala Ala Ser Asp Ile Leu
            260                 265                 270

Arg Ile Ser Ala Leu Lys Glu Ile Gly Gly Met Tyr Leu Asp Val Asp
        275                 280                 285

Met Leu Pro Gly Ile Gln Pro Asp Leu Phe Glu Ser Ile Glu Lys Pro
    290                 295                 300

Ser Ser Val Thr Val Asp Phe Trp Glu Met Thr Lys Leu Glu Ala Ile
305                 310                 315                 320

Met Lys Tyr Lys Glu Tyr Ile Pro Glu Tyr Thr Ser Glu His Phe Asp
                325                 330                 335

Met Leu Asp Glu Glu Val Gln Ser Ser Phe Glu Ser Val Leu Ala Ser
            340                 345                 350

Lys Ser Asp Lys Ser Glu Ile Phe Ser Ser Leu Gly Asp Met Glu Ala
        355                 360                 365

Ser Pro Leu Glu Val Lys Ile Ala Phe Asn Ser Lys Gly Ile Ile Asn
    370                 375                 380

Gln Gly Leu Ile Ser Val Lys Asp Ser Tyr Cys Ser Asn Leu Ile Val
385                 390                 395                 400

Lys Gln Ile Glu Asn Arg Tyr Lys Ile Leu Asn Asn Ser Leu Asn Pro
                405                 410                 415
```

```
Ala Ile Ser Glu Asp Asn Asp Phe Asn Thr Thr Thr Asn Thr Phe Ile
            420                 425                 430

Asp Ser Ile Met Ala Glu Ala Asn Ala Asp Asn Gly Arg Phe Met Met
        435                 440                 445

Glu Leu Gly Lys Tyr Leu Arg Val Gly Phe Phe Pro Asp Val Lys Thr
    450                 455                 460

Thr Ile Asn Leu Ser Gly Pro Glu Ala Tyr Ala Ala Ala Tyr Gln Asp
465                 470                 475                 480

Leu Leu Met Phe Lys Glu Gly Ser Met Asn Ile His Leu Ile Glu Ala
                485                 490                 495

Asp Leu Arg Asn Phe Glu Ile Ser Lys Thr Asn Ile Ser Gln Ser Thr
            500                 505                 510

Glu Gln Glu Met Ala Ser Leu Trp Ser Phe Asp Asp Ala Arg Ala Lys
        515                 520                 525

Ala Gln Phe Glu Glu Tyr Lys Arg Asn Tyr Phe Glu Gly Ser Leu Gly
    530                 535                 540

Glu Asp Asp Asn Leu Asp Phe Ser Gln Asn Ile Val Val Asp Lys Glu
545                 550                 555                 560

Tyr Leu Leu Glu Lys Ile Ser Ser Leu Ala Arg Ser Ser Glu Arg Gly
                565                 570                 575

Tyr Ile His Tyr Ile Val Gln Leu Gln Gly Asp Lys Ile Ser Tyr Glu
            580                 585                 590

Ala Ala Cys Asn Leu Phe Ala Lys Thr Pro Tyr Asp Ser Val Leu Phe
        595                 600                 605

Gln Lys Asn Ile Glu Asp Ser Glu Ile Ala Tyr Tyr Tyr Asn Pro Gly
    610                 615                 620

Asp Gly Glu Ile Gln Glu Ile Asp Lys Tyr Lys Ile Pro Ser Ile Ile
625                 630                 635                 640

Ser Asp Arg Pro Lys Ile Lys Leu Thr Phe Ile Gly His Gly Lys Asp
                645                 650                 655

Glu Phe Asn Thr Asp Ile Phe Ala Gly Phe Asp Val Asp Ser Leu Ser
```

660 665 670

Thr Glu Ile Glu Ala Ala Ile Asp Leu Ala Lys Glu Asp Ile Ser Pro
675 680 685

Lys Ser Ile Glu Ile Asn Leu Leu Gly Cys Asn Met Phe Ser Tyr Ser
690 695 700

Ile Asn Val Glu Glu Thr Tyr Pro Gly Lys Leu Leu Leu Lys Val Lys
705 710 715 720

Asp Lys Ile Ser Glu Leu Met Pro Ser Ile Ser Gln Asp Ser Ile Ile
725 730 735

Val Ser Ala Asn Gln Tyr Glu Val Arg Ile Asn Ser Glu Gly Arg Arg
740 745 750

Glu Leu Leu Asp His Ser Gly Glu Trp Ile Asn Lys Glu Glu Ser Ile
755 760 765

Ile Lys Asp Ile Ser Ser Lys Glu Tyr Ile Ser Phe Asn Pro Lys Glu
770 775 780

Asn Lys Ile Thr Val Lys Ser Lys Asn Leu Pro Glu Leu Ser Thr Leu
785 790 795 800

Leu Gln Glu Ile Arg Asn Asn Ser Asn Ser Ser Asp Ile Glu Leu Glu
805 810 815

Glu Lys Val Met Leu Thr Glu Cys Glu Ile Asn Val Ile Ser Asn Ile
820 825 830

Asp Thr Gln Ile Val Glu Glu Arg Ile Glu Glu Ala Lys Asn Leu Thr
835 840 845

Ser Asp Ser Ile Asn Tyr Ile Lys Asp Glu Phe Lys Leu Ile Glu Ser
850 855 860

Ile Ser Asp Ala Leu Cys Asp Leu Lys Gln Gln Asn Glu Leu Glu Asp
865 870 875 880

Ser His Phe Ile Ser Phe Glu Asp Ile Ser Glu Thr Asp Glu Gly Phe
885 890 895

Ser Ile Arg Phe Ile Asn Lys Glu Thr Gly Glu Ser Ile Phe Val Glu
900 905 910

```
Thr Glu Lys Thr Ile Phe Ser Glu Tyr Ala Asn His Ile Thr Glu Glu
        915                 920                 925

Ile Ser Lys Ile Lys Gly Thr Ile Phe Asp Thr Val Asn Gly Lys Leu
    930                 935                 940

Val Lys Lys Val Asn Leu Asp Thr Thr His Glu Val Asn Thr Leu Asn
945                 950                 955                 960

Ala Ala Phe Phe Ile Gln Ser Leu Ile Glu Tyr Asn Ser Ser Lys Glu
                965                 970                 975

Ser Leu Ser Asn Leu Ser Val Ala Met Lys Val Gln Val Tyr Ala Gln
            980                 985                 990

Leu Phe Ser Thr Gly Leu Asn Thr  Ile Thr Asp Ala Ala  Lys Val Val
        995                 1000                1005

Glu Leu  Val Ser Thr Ala Leu  Asp Glu Thr Ile Asp  Leu Leu Pro
    1010                1015                1020

Thr Leu  Ser Glu Gly Leu Pro  Ile Ile Ala Thr Ile  Ile Asp Gly
    1025                1030                1035

Val Ser  Leu Gly Ala Ala Ile  Lys Glu Leu Ser Glu  Thr Ser Asp
    1040                1045                1050

Pro Leu  Leu Arg Gln Glu Ile  Glu Ala Lys Ile Gly  Ile Met Ala
    1055                1060                1065

Val Asn  Leu Thr Thr Ala Thr  Thr Ala Ile Ile Thr  Ser Ser Leu
    1070                1075                1080

Gly Ile  Ala Ser Gly Phe Ser  Ile Leu Leu Val Pro  Leu Ala Gly
    1085                1090                1095

Ile Ser  Ala Gly Ile Pro Ser  Leu Val Asn Asn Glu  Leu Val Leu
    1100                1105                1110

Arg Asp  Lys Ala Thr Lys Val  Val Asp Tyr Phe Lys  His Val Ser
    1115                1120                1125

Leu Val  Glu Thr Glu Gly Val  Phe Thr Leu Leu Asp  Asp Lys Ile
    1130                1135                1140

Met Met  Pro Gln Asp Asp Leu  Val Ile Ser Glu Ile  Asp Phe Asn
    1145                1150                1155
```

```
Asn Asn  Ser Ile Val Leu Gly  Lys Cys Glu Ile Trp  Arg Met Glu
    1160                 1165                 1170

Gly Gly  Ser Gly His Thr Val  Thr Asp Asp Ile Asp  His Phe Phe
    1175                 1180                 1185

Ser Ala  Pro Ser Ile Thr Tyr  Arg Glu Pro His Leu  Ser Ile Tyr
    1190                 1195                 1200

Asp Val  Leu Glu Val Gln Lys  Glu Glu Leu Asp Leu  Ser Lys Asp
    1205                 1210                 1215

Leu Met  Val Leu Pro Asn Ala  Pro Asn Arg Val Phe  Ala Trp Glu
    1220                 1225                 1230

Thr Gly  Trp Thr Pro Gly Leu  Arg Ser Leu Glu Asn  Asp Gly Thr
    1235                 1240                 1245

Lys Leu  Leu Asp Arg Ile Arg  Asp Asn Tyr Glu Gly  Glu Phe Tyr
    1250                 1255                 1260

Trp Arg  Tyr Phe Ala Phe Ile  Ala Asp Ala Leu Ile  Thr Thr Leu
    1265                 1270                 1275

Lys Pro  Arg Tyr Glu Asp Thr  Asn Ile Arg Ile Asn  Leu Asp Ser
    1280                 1285                 1290

Asn Thr  Arg Ser Phe Ile Val  Pro Ile Ile Thr Thr  Glu Tyr Ile
    1295                 1300                 1305

Arg Glu  Lys Leu Ser Tyr Ser  Phe Tyr Gly Ser Gly  Gly Thr Tyr
    1310                 1315                 1320

Ala Leu  Ser Leu Ser Gln Tyr  Asn Met Gly Ile Asn  Ile Glu Leu
    1325                 1330                 1335

Ser Glu  Ser Asp Val Trp Ile  Ile Asp Val Asp Asn  Val Val Arg
    1340                 1345                 1350

Asp Val  Thr Ile Glu Ser Asp  Lys Ile Lys Lys Gly  Asp Leu Ile
    1355                 1360                 1365

Glu Gly  Ile Leu Ser Thr Leu  Ser Ile Glu Glu Asn  Lys Ile Ile
    1370                 1375                 1380

Leu Asn  Ser His Glu Ile Asn  Phe Ser Gly Glu Val  Asn Gly Ser
    1385                 1390                 1395
```

```
Asn Gly  Phe Val Ser Leu Thr  Phe Ser Ile Leu Glu  Gly Ile Asn
    1400                 1405                 1410

Ala Ile  Ile Glu Val Asp Leu  Leu Ser Lys Ser Tyr  Lys Leu Leu
    1415                 1420                 1425

Ile Ser  Gly Glu Leu Lys Ile  Leu Met Leu Asn Ser  Asn His Ile
    1430                 1435                 1440

Gln Gln  Lys Ile Asp Tyr Ile  Gly Phe Asn Ser Glu  Leu Gln Lys
    1445                 1450                 1455

Asn Ile  Pro Tyr Ser Phe Val  Asp Ser Glu Gly Lys  Glu Asn Gly
    1460                 1465                 1470

Phe Ile  Asn Gly Ser Thr Lys  Glu Gly Leu Phe Val  Ser Glu Leu
    1475                 1480                 1485

Pro Asp  Val Val Leu Ile Ser  Lys Val Tyr Met Asp  Asp Ser Lys
    1490                 1495                 1500

Pro Ser  Phe Gly Tyr Tyr Ser  Asn Asn Leu Lys Asp  Val Lys Val
    1505                 1510                 1515

Ile Thr  Lys Asp Asn Val Asn  Ile Leu Thr Gly Tyr  Tyr Leu Lys
    1520                 1525                 1530

Asp Asp  Ile Lys Ile Ser Leu  Ser Leu Thr Leu Gln  Asp Glu Lys
    1535                 1540                 1545

Thr Ile  Lys Leu Asn Ser Val  His Leu Asp Glu Ser  Gly Val Ala
    1550                 1555                 1560

Glu Ile  Leu Lys Phe Met Asn  Arg Lys Gly Asn Thr  Asn Thr Ser
    1565                 1570                 1575

Asp Ser  Leu Met Ser Phe Leu  Glu Ser Met Asn Ile  Lys Ser Ile
    1580                 1585                 1590

Phe Val  Asn Phe Leu Gln Ser  Asn Ile Lys Phe Ile  Leu Asp Ala
    1595                 1600                 1605

Asn Phe  Ile Ile Ser Gly Thr  Thr Ser Ile Gly Gln  Phe Glu Phe
    1610                 1615                 1620

Ile Cys  Asp Glu Asn Asp Asn  Ile Gln Pro Tyr Phe  Ile Lys Phe
```

```
    1625                1630                1635

Asn Thr  Leu Glu Thr Asn Tyr  Thr Leu Tyr Val Gly  Asn Arg Gln
    1640                1645                1650

Asn Met  Ile Val Glu Pro Asn  Tyr Asp Leu Asp Asp  Ser Gly Asp
    1655                1660                1665

Ile Ser  Ser Thr Val Ile Asn  Phe Ser Gln Lys Tyr  Leu Tyr Gly
    1670                1675                1680

Ile Asp  Ser Cys Val Asn Lys  Val Val Ile Ser Pro  Asn Ile Tyr
    1685                1690                1695

Thr Asp  Glu Ile Asn Ile Thr  Pro Val Tyr Glu Thr  Asn Asn Thr
    1700                1705                1710

Tyr Pro  Glu Val Ile Val Leu  Asp Ala Asn Tyr Ile  Asn Glu Lys
    1715                1720                1725

Ile Asn  Val Asn Ile Asn Asp  Leu Ser Ile Arg Tyr  Val Trp Ser
    1730                1735                1740

Asn Asp  Gly Asn Asp Phe Ile  Leu Met Ser Thr Ser  Glu Glu Asn
    1745                1750                1755

Lys Val  Ser Gln Val Lys Ile  Arg Phe Val Asn Val  Phe Lys Asp
    1760                1765                1770

Lys Thr  Leu Ala Asn Lys Leu  Ser Phe Asn Phe Ser  Asp Lys Gln
    1775                1780                1785

Asp Val  Pro Val Ser Glu Ile  Ile Leu Ser Phe Thr  Pro Ser Tyr
    1790                1795                1800

Tyr Glu  Asp Gly Leu Ile Gly  Tyr Asp Leu Gly Leu  Val Ser Leu
    1805                1810                1815

Tyr Asn  Glu Lys Phe Tyr Ile  Asn Asn Phe Gly Met  Met Val Ser
    1820                1825                1830

Gly Leu  Ile Tyr Ile Asn Asp  Ser Leu Tyr Tyr Phe  Lys Pro Pro
    1835                1840                1845

Val Asn  Asn Leu Ile Thr Gly  Phe Val Thr Val Gly  Asp Asp Lys
    1850                1855                1860
```

```
Tyr Tyr  Phe Asn Pro Ile Asn  Gly Gly Ala Ala Ser  Ile Gly Glu
    1865                 1870                 1875

Thr Ile  Ile Asp Asp Lys Asn  Tyr Tyr Phe Asn Gln  Ser Gly Val
    1880                 1885                 1890

Leu Gln  Thr Gly Val Phe Ser  Thr Glu Asp Gly Phe  Lys Tyr Phe
    1895                 1900                 1905

Ala Pro  Ala Asn Thr Leu Asp  Glu Asn Leu Glu Gly  Glu Ala Ile
    1910                 1915                 1920

Asp Phe  Thr Gly Lys Leu Ile  Ile Asp Glu Asn Ile  Tyr Tyr Phe
    1925                 1930                 1935

Asp Asp  Asn Tyr Arg Gly Ala  Val Glu Trp Lys Glu  Leu Asp Gly
    1940                 1945                 1950

Glu Met  His Tyr Phe Ser Pro  Glu Thr Gly Lys Ala  Phe Lys Gly
    1955                 1960                 1965

Leu Asn  Gln Ile Gly Asp Tyr  Lys Tyr Tyr Phe Asn  Ser Asp Gly
    1970                 1975                 1980

Val Met  Gln Lys Gly Phe Val  Ser Ile Asn Asp Asn  Lys His Tyr
    1985                 1990                 1995

Phe Asp  Asp Ser Gly Val Met  Lys Val Gly Tyr Thr  Glu Ile Asp
    2000                 2005                 2010

Gly Lys  His Phe Tyr Phe Ala  Glu Asn Gly Glu Met  Gln Ile Gly
    2015                 2020                 2025

Val Phe  Asn Thr Glu Asp Gly  Phe Lys Tyr Phe Ala  His His Asn
    2030                 2035                 2040

Glu Asp  Leu Gly Asn Glu Glu  Gly Glu Glu Ile Ser  Tyr Ser Gly
    2045                 2050                 2055

Ile Leu  Asn Phe Asn Asn Lys  Ile Tyr Tyr Phe Asp  Asp Ser Phe
    2060                 2065                 2070

Thr Ala  Val Val Gly Trp Lys  Asp Leu Glu Asp Gly  Ser Lys Tyr
    2075                 2080                 2085

Tyr Phe  Asp Glu Asp Thr Ala  Glu Ala Tyr Ile Gly  Leu Ser Leu
    2090                 2095                 2100
```

```
Ile Asn  Asp Gly Gln Tyr Tyr  Phe Asn Asp Asp Gly  Ile Met Gln
    2105                 2110                 2115

Val Gly  Phe Val Thr Ile Asn  Asp Lys Val Phe Tyr  Phe Ser Asp
    2120                 2125                 2130

Ser Gly  Ile Ile Glu Ser Gly  Val Gln Asn Ile Asp  Asp Asn Tyr
    2135                 2140                 2145

Phe Tyr  Ile Asp Asp Asn Gly  Ile Val Gln Ile Gly  Val Phe Asp
    2150                 2155                 2160

Thr Ser  Asp Gly Tyr Lys Tyr  Phe Ala Pro Ala Asn  Thr Val Asn
    2165                 2170                 2175

Asp Asn  Ile Tyr Gly Gln Ala  Val Glu Tyr Ser Gly  Leu Val Arg
    2180                 2185                 2190

Val Gly  Glu Asp Val Tyr Tyr  Phe Gly Glu Thr Tyr  Thr Ile Glu
    2195                 2200                 2205

Thr Gly  Trp Ile Tyr Asp Met  Glu Asn Glu Ser Asp  Lys Tyr Tyr
    2210                 2215                 2220

Phe Asn  Pro Glu Thr Lys Lys  Ala Cys Lys Gly Ile  Asn Leu Ile
    2225                 2230                 2235

Asp Asp  Ile Lys Tyr Tyr Phe  Asp Glu Lys Gly Ile  Met Arg Thr
    2240                 2245                 2250

Gly Leu  Ile Ser Phe Glu Asn  Asn Asn Tyr Tyr Phe  Asn Glu Asn
    2255                 2260                 2265

Gly Glu  Met Gln Phe Gly Tyr  Ile Asn Ile Glu Asp  Lys Met Phe
    2270                 2275                 2280

Tyr Phe  Gly Glu Asp Gly Val  Met Gln Ile Gly Val  Phe Asn Thr
    2285                 2290                 2295

Pro Asp  Gly Phe Lys Tyr Phe  Ala His Gln Asn Thr  Leu Asp Glu
    2300                 2305                 2310

Asn Phe  Glu Gly Glu Ser Ile  Asn Tyr Thr Gly Trp  Leu Asp Leu
    2315                 2320                 2325

Asp Glu  Lys Arg Tyr Tyr Phe  Thr Asp Glu Tyr Ile  Ala Ala Thr
    2330                 2335                 2340
```

Gly Ser Val Ile Ile Asp Gly Glu Glu Tyr Tyr Phe Asp Pro Asp
2345 2350 2355

Thr Ala Gln Leu Val Ile Ser Glu
2360 2365

<210> 9
<211> 32
<212> DNA
<213> artificial

<220>
<223> forward primer

<400> 9
caccactagt atgaacttag taactggatg gc 32

<210> 10
<211> 27
<212> DNA
<213> artificial

<220>
<223> reverse primer

<400> 10
ctcgagttag ccatatatcc caggggc 27

<210> 11
<211> 34
<212> DNA
<213> artificial

<220>
<223> forward primer

<400> 11
caccatgcat atgagtttag ttaatagaaa acag 34

<210> 12
<211> 33
<212> DNA
<213> artificial

<220>
<223> reverse primer

<400> 12
ggcctcgagc tattcactaa tcactaattg agc 33

<210> 13
<211> 44
<212> DNA
<213> artificial

<220>
<223> forward primer

<400> 13
agatctatgc atgagctcct cgagcccaaa acgaaaggct cagc 44

<210> 14
<211> 36
<212> DNA
<213> artificial

<220>
<223> reverse primer

<400> 14
cggtccgggg ccatatatcc caggggcttt tactcc 36

<210> 15
<211> 34
<212> DNA
<213> artificial

<220>
<223> forward primer

<400> 15
caccccattg atggtaacag gagtatttaa agga 34

<210> 16
<211> 32
<212> DNA
<213> artificial

<220>
<223> reverse primer

<400> 16
ctcgagctat tcactaatca ctaattgagc tg 32

<210> 17
<211> 4482
<212> DNA
<213> artificial

<220>
<223> C-TADCTB fusion DNA

<400> 17

atggtaacag gagtatttaa aggacctaat ggatttgagt attttgcacc tgctaatact 60

cacaataata acatagaagg tcaggctata gtttaccaga acaaattctt aactttgaat 120

ggcaaaaaat attattttga taatgactca aaagcagtta ctggatggca aaccattgat 180

ggtaaaaaat attactttaa tcttaacact gctgaagcag ctactggatg gcaaactatt 240

gatggtaaaa aatattactt taatcttaac actgctgaag cagctactgg atggcaaact 300

attgatggta aaaaatatta ctttaatact aacactttca tagcctcaac tggttataca 360

agtattaatg gtaaacattt ttattttaat actgatggta ttatgcagat aggagtgttt 420

aaaggaccta atggatttga atactttgca cctgctaata cggatgctaa caacatagaa 480

ggtcaagcta tactttacca aaataaattc ttaactttga atggtaaaaa atattacttt 540

ggtagtgact caaaagcagt taccggactg cgaactattg atggtaaaaa atattacttt 600

aatactaaca ctgctgttgc agttactgga tggcaaacta ttaatggtaa aaaatactac 660

tttaatacta acacttctat agcttcaact ggttatacaa ttattagtgg taaacatttt 720

tattttaata ctgatggtat tatgcagata ggagtgttta aaggacctga tggatttgaa 780

tactttgcac ctgctaatac agatgctaac aatatagaag gtcaagctat acgttatcaa 840

aatagattcc tatatttaca tgacaatata tattattttg gtaataattc aaaagcggct 900

actggttggg taactattga tggtaataga tattacttcg agcctaatac agctatgggt 960

gcgaatggtt ataaaactat tgataataaa aatttttact ttagaaatgg tttacctcag 1020

ataggagtgt ttaaagggtc taatggattt gaatactttg cacctgctaa tacggatgct 1080

aacaatatag aaggtcaagc tatacgttat caaaatagat tcctacattt acttggaaaa 1140

atatattact ttggtaataa ttcaaaagca gttactggat ggcaaactat taatggtaaa 1200

gtatattact ttatgcctga tactgctatg gctgcagctg gtggactttt cgagattgat 1260

ggtgttatat atttctttgg tgttgatgga gtaaaagccc ctgggatata tggcagatct 1320

atgcataatt tgataactgg atttgtgact gtaggcgatg ataaatacta ctttaatcca 1380

attaatggtg gagctgcttc aattggagag acaataattg atgacaaaaa ttattatttc 1440

aaccaaagtg gagtgttaca aacaggtgta tttagtacag aagatggatt taaatatttt 1500

gccccagcta atacacttga tgaaaaccta gaaggagaag caattgattt tactggaaaa 1560

ttaattattg acgaaaatat ttattatttt gatgataatt atagaggagc tgtagaatgg 1620

aaagaattag atggtgaaat gcactatttt agcccagaaa caggtaaagc ttttaaaggt 1680

ctaaatcaaa taggtgatta taaatactat ttcaattctg atggagttat gcaaaaagga 1740

tttgttagta taaatgataa taaacactat tttgatgatt ctggtgttat gaaagtaggt 1800

tacactgaaa tagatggcaa gcatttctac tttgctgaaa acggagaaat gcaaatagga 1860

gtatttaata cagaagatgg atttaaatat tttgctcatc ataatgaaga tttaggaaat 1920

gaagaaggtg aagaaatctc atattctggt atattaaatt tcaataataa aatttactat 1980

tttgatgatt catttacagc tgtagttgga tggaaagatt tagaggatgg ttcaaagtat 2040

tattttgatg aagatacagc agaagcatat ataggtttgt cattaataaa tgatggtcaa 2100

tattatttta atgatgatgg aattatgcaa gttggatttg tcactataaa tgataaagtc 2160

ttctacttct ctgactctgg aattatagaa tctggagtac aaaacataga tgacaattat 2220

ttctatatag atgataatgg tatagttcaa attggtgtat ttgatacttc agatggatat 2280

aaatattttg cacctgctaa tactgtaaat gataatattt acggacaagc agttgaatat 2340

agtggtttag ttagagttgg ggaagatgta tattattttg gagaaacata tacaattgag 2400

actggatgga tatatgatat ggaaaatgaa agtgataaat attatttcaa tccagaaact 2460

aaaaaagcat gcaaaggtat taatttaatt gatgatataa aatattattt tgatgagaag 2520

ggcataatga gaacgggtct tatatcattt gaaaataata attattactt taatgagaat 2580

ggtgaaatgc aatttggtta tataaatata gaagataaga tgttctattt tggtgaagat 2640

ggtgtcatgc agattggagt atttaataca ccagatggat ttaaatactt tgcacatcaa 2700

aatactttgg atgagaattt tgagggagaa tcaataaact atactggttg gttagattta 2760

gatgaaaaga gatattattt tacagatgaa tatattgcag caactggttc agttattatt 2820

gatggtgagg agtattattt tgatcctgat acagctcaat tagtgattag tgaactcgag 2880

ggattaatat atattaatga ttcattatat tattttaaac caccagtaaa taatttgata 2940

actggatttg tgactgtagg cgatgataaa tactacttta atccaattaa tggtggagct 3000

gcttcaattg gagagacaat aattgatgac aaaaattatt atttcaacca aagtggagtg 3060

ttacaaacag gtgtatttag tacagaagat ggatttaaat attttgcccc agctaataca 3120

cttgatgaaa acctagaagg agaagcaatt gattttactg gaaaattaat tattgacgaa 3180

aatatttatt attttgatga taattataga ggagctgtag aatggaaaga attagatggt 3240

gaaatgcact attttagccc agaaacaggt aaagctttta aaggtctaaa tcaaataggt 3300

gattataaat actatttcaa ttctgatgga gttatgcaaa aaggatttgt tagtataaat 3360

gataataaac actattttga tgattctggt gttatgaaag taggttacac tgaaatagat 3420

ggcaagcatt tctactttgc tgaaaacgga gaaatgcaaa taggagtatt taatacagaa 3480

gatggattta aatattttgc tcatcataat gaagatttag gaaatgaaga aggtgaagaa 3540

atctcatatt ctggtatatt aaatttcaat aataaaattt actattttga tgattcattt 3600

acagctgtag ttggatggaa agatttagag gatggttcaa agtattattt tgatgaagat 3660

acagcagaag catatatagg tttgtcatta ataaatgatg gtcaatatta ttttaatgat 3720

gatggaatta tgcaagttgg atttgtcact ataaatgata aagtcttcta cttctctgac 3780

tctggaatta tagaatctgg agtacaaaac atagatgaca attatttcta tatagatgat 3840

aatggtatag ttcaaattgg tgtatttgat acttcagatg gatataaata ttttgcacct 3900

gctaatactg taaatgataa tatttacgga caagcagttg aatatagtgg tttagttaga 3960

gttggggaag atgtatatta ttttggagaa acatatacaa ttgagactgg atggatatat 4020

gatatggaaa atgaaagtga taaatattat ttcaatccag aaactaaaaa agcatgcaaa 4080

ggtattaatt taattgatga tataaaatat tattttgatg agaagggcat aatgagaacg 4140

ggtcttatat catttgaaaa taataattat tactttaatg agaatggtga aatgcaattt 4200

ggttatataa atatagaaga taagatgttc tattttggtg aagatggtgt catgcagatt 4260

ggagtattta atacaccaga tggatttaaa tactttgcac atcaaaatac tttggatgag 4320

aattttgagg gagaatcaat aaactatact ggttggttag atttagatga aaagagatat 4380

tattttacag atgaatatat tgcagcaact ggttcagtta ttattgatgg tgaggagtat 4440

tattttgatc ctgatacagc tcaattagtg attagtgaat ag 4482

<210> 18
<211> 1493
<212> PRT
<213> artificial

<220>
<223> C-TADCTB fusion protein

<400> 18

Met Val Thr Gly Val Phe Lys Gly Pro Asn Gly Phe Glu Tyr Phe Ala
1 5 10 15

Pro Ala Asn Thr His Asn Asn Asn Ile Glu Gly Gln Ala Ile Val Tyr
20 25 30

Gln Asn Lys Phe Leu Thr Leu Asn Gly Lys Lys Tyr Tyr Phe Asp Asn
35 40 45

Asp Ser Lys Ala Val Thr Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr
50 55 60

Tyr Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr Gly Trp Gln Thr Ile
65 70 75 80

Asp Gly Lys Lys Tyr Tyr Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr
85 90 95

Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr
100 105 110

Phe Ile Ala Ser Thr Gly Tyr Thr Ser Ile Asn Gly Lys His Phe Tyr
115 120 125

Phe Asn Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro Asn
130 135 140

Gly Phe Glu Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu
145 150 155 160

```
Gly Gln Ala Ile Leu Tyr Gln Asn Lys Phe Leu Thr Leu Asn Gly Lys
                165                 170                 175

Lys Tyr Tyr Phe Gly Ser Asp Ser Lys Ala Val Thr Gly Leu Arg Thr
            180                 185                 190

Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Ala Val Ala Val
        195                 200                 205

Thr Gly Trp Gln Thr Ile Asn Gly Lys Lys Tyr Tyr Phe Asn Thr Asn
    210                 215                 220

Thr Ser Ile Ala Ser Thr Gly Tyr Thr Ile Ile Ser Gly Lys His Phe
225                 230                 235                 240

Tyr Phe Asn Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro
                245                 250                 255

Asp Gly Phe Glu Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile
            260                 265                 270

Glu Gly Gln Ala Ile Arg Tyr Gln Asn Arg Phe Leu Tyr Leu His Asp
        275                 280                 285

Asn Ile Tyr Tyr Phe Gly Asn Asn Ser Lys Ala Ala Thr Gly Trp Val
    290                 295                 300

Thr Ile Asp Gly Asn Arg Tyr Tyr Phe Glu Pro Asn Thr Ala Met Gly
305                 310                 315                 320

Ala Asn Gly Tyr Lys Thr Ile Asp Asn Lys Asn Phe Tyr Phe Arg Asn
                325                 330                 335

Gly Leu Pro Gln Ile Gly Val Phe Lys Gly Ser Asn Gly Phe Glu Tyr
            340                 345                 350

Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu Gly Gln Ala Ile
        355                 360                 365

Arg Tyr Gln Asn Arg Phe Leu His Leu Leu Gly Lys Ile Tyr Tyr Phe
    370                 375                 380

Gly Asn Asn Ser Lys Ala Val Thr Gly Trp Gln Thr Ile Asn Gly Lys
385                 390                 395                 400

Val Tyr Tyr Phe Met Pro Asp Thr Ala Met Ala Ala Ala Gly Gly Leu
                405                 410                 415
```

```
Phe Glu Ile Asp Gly Val Ile Tyr Phe Phe Gly Val Asp Gly Val Lys
            420                 425                 430

Ala Pro Gly Ile Tyr Gly Arg Ser Met His Asn Leu Ile Thr Gly Phe
        435                 440                 445

Val Thr Val Gly Asp Asp Lys Tyr Tyr Phe Asn Pro Ile Asn Gly Gly
    450                 455                 460

Ala Ala Ser Ile Gly Glu Thr Ile Ile Asp Asp Lys Asn Tyr Tyr Phe
465                 470                 475                 480

Asn Gln Ser Gly Val Leu Gln Thr Gly Val Phe Ser Thr Glu Asp Gly
                485                 490                 495

Phe Lys Tyr Phe Ala Pro Ala Asn Thr Leu Asp Glu Asn Leu Glu Gly
            500                 505                 510

Glu Ala Ile Asp Phe Thr Gly Lys Leu Ile Ile Asp Glu Asn Ile Tyr
        515                 520                 525

Tyr Phe Asp Asp Asn Tyr Arg Gly Ala Val Glu Trp Lys Glu Leu Asp
    530                 535                 540

Gly Glu Met His Tyr Phe Ser Pro Glu Thr Gly Lys Ala Phe Lys Gly
545                 550                 555                 560

Leu Asn Gln Ile Gly Asp Tyr Lys Tyr Tyr Phe Asn Ser Asp Gly Val
                565                 570                 575

Met Gln Lys Gly Phe Val Ser Ile Asn Asp Asn Lys His Tyr Phe Asp
            580                 585                 590

Asp Ser Gly Val Met Lys Val Gly Tyr Thr Glu Ile Asp Gly Lys His
        595                 600                 605

Phe Tyr Phe Ala Glu Asn Gly Glu Met Gln Ile Gly Val Phe Asn Thr
    610                 615                 620

Glu Asp Gly Phe Lys Tyr Phe Ala His His Asn Glu Asp Leu Gly Asn
625                 630                 635                 640

Glu Glu Gly Glu Glu Ile Ser Tyr Ser Gly Ile Leu Asn Phe Asn Asn
                645                 650                 655

Lys Ile Tyr Tyr Phe Asp Asp Ser Phe Thr Ala Val Val Gly Trp Lys
```

660 665 670

Asp Leu Glu Asp Gly Ser Lys Tyr Tyr Phe Asp Glu Asp Thr Ala Glu
675 680 685

Ala Tyr Ile Gly Leu Ser Leu Ile Asn Asp Gly Gln Tyr Tyr Phe Asn
690 695 700

Asp Asp Gly Ile Met Gln Val Gly Phe Val Thr Ile Asn Asp Lys Val
705 710 715 720

Phe Tyr Phe Ser Asp Ser Gly Ile Ile Glu Ser Gly Val Gln Asn Ile
725 730 735

Asp Asp Asn Tyr Phe Tyr Ile Asp Asp Asn Gly Ile Val Gln Ile Gly
740 745 750

Val Phe Asp Thr Ser Asp Gly Tyr Lys Tyr Phe Ala Pro Ala Asn Thr
755 760 765

Val Asn Asp Asn Ile Tyr Gly Gln Ala Val Glu Tyr Ser Gly Leu Val
770 775 780

Arg Val Gly Glu Asp Val Tyr Tyr Phe Gly Glu Thr Tyr Thr Ile Glu
785 790 795 800

Thr Gly Trp Ile Tyr Asp Met Glu Asn Glu Ser Asp Lys Tyr Tyr Phe
805 810 815

Asn Pro Glu Thr Lys Lys Ala Cys Lys Gly Ile Asn Leu Ile Asp Asp
820 825 830

Ile Lys Tyr Tyr Phe Asp Glu Lys Gly Ile Met Arg Thr Gly Leu Ile
835 840 845

Ser Phe Glu Asn Asn Asn Tyr Tyr Phe Asn Glu Asn Gly Glu Met Gln
850 855 860

Phe Gly Tyr Ile Asn Ile Glu Asp Lys Met Phe Tyr Phe Gly Glu Asp
865 870 875 880

Gly Val Met Gln Ile Gly Val Phe Asn Thr Pro Asp Gly Phe Lys Tyr
885 890 895

Phe Ala His Gln Asn Thr Leu Asp Glu Asn Phe Glu Gly Glu Ser Ile
900 905 910

```
Asn Tyr Thr Gly Trp Leu Asp Leu Asp Glu Lys Arg Tyr Tyr Phe Thr
        915                 920                 925

Asp Glu Tyr Ile Ala Ala Thr Gly Ser Val Ile Ile Asp Gly Glu Glu
    930                 935                 940

Tyr Tyr Phe Asp Pro Asp Thr Ala Gln Leu Val Ile Ser Glu Leu Glu
945                 950                 955                 960

Gly Leu Ile Tyr Ile Asn Asp Ser Leu Tyr Tyr Phe Lys Pro Pro Val
                965                 970                 975

Asn Asn Leu Ile Thr Gly Phe Val Thr Val Gly Asp Asp Lys Tyr Tyr
            980                 985                 990

Phe Asn Pro Ile Asn Gly Gly Ala  Ala Ser Ile Gly Glu  Thr Ile Ile
        995                 1000                1005

Asp Asp  Lys Asn Tyr Tyr Phe  Asn Gln Ser Gly Val  Leu Gln Thr
    1010                1015                1020

Gly Val  Phe Ser Thr Glu Asp  Gly Phe Lys Tyr Phe  Ala Pro Ala
    1025                1030                1035

Asn Thr  Leu Asp Glu Asn Leu  Glu Gly Glu Ala Ile  Asp Phe Thr
    1040                1045                1050

Gly Lys  Leu Ile Ile Asp Glu  Asn Ile Tyr Tyr Phe  Asp Asp Asn
    1055                1060                1065

Tyr Arg  Gly Ala Val Glu Trp  Lys Glu Leu Asp Gly  Glu Met His
    1070                1075                1080

Tyr Phe  Ser Pro Glu Thr Gly  Lys Ala Phe Lys Gly  Leu Asn Gln
    1085                1090                1095

Ile Gly  Asp Tyr Lys Tyr Tyr  Phe Asn Ser Asp Gly  Val Met Gln
    1100                1105                1110

Lys Gly  Phe Val Ser Ile Asn  Asp Asn Lys His Tyr  Phe Asp Asp
    1115                1120                1125

Ser Gly  Val Met Lys Val Gly  Tyr Thr Glu Ile Asp  Gly Lys His
    1130                1135                1140

Phe Tyr  Phe Ala Glu Asn Gly  Glu Met Gln Ile Gly  Val Phe Asn
    1145                1150                1155
```

```
Thr Glu  Asp Gly Phe Lys Tyr  Phe Ala His His Asn  Glu Asp Leu
    1160                1165                1170

Gly Asn  Glu Glu Gly Glu Glu  Ile Ser Tyr Ser Gly  Ile Leu Asn
    1175                1180                1185

Phe Asn  Asn Lys Ile Tyr Tyr  Phe Asp Asp Ser Phe  Thr Ala Val
    1190                1195                1200

Val Gly  Trp Lys Asp Leu Glu  Asp Gly Ser Lys Tyr  Tyr Phe Asp
    1205                1210                1215

Glu Asp  Thr Ala Glu Ala Tyr  Ile Gly Leu Ser Leu  Ile Asn Asp
    1220                1225                1230

Gly Gln  Tyr Tyr Phe Asn Asp  Asp Gly Ile Met Gln  Val Gly Phe
    1235                1240                1245

Val Thr  Ile Asn Asp Lys Val  Phe Tyr Phe Ser Asp  Ser Gly Ile
    1250                1255                1260

Ile Glu  Ser Gly Val Gln Asn  Ile Asp Asp Asn Tyr  Phe Tyr Ile
    1265                1270                1275

Asp Asp  Asn Gly Ile Val Gln  Ile Gly Val Phe Asp  Thr Ser Asp
    1280                1285                1290

Gly Tyr  Lys Tyr Phe Ala Pro  Ala Asn Thr Val Asn  Asp Asn Ile
    1295                1300                1305

Tyr Gly  Gln Ala Val Glu Tyr  Ser Gly Leu Val Arg  Val Gly Glu
    1310                1315                1320

Asp Val  Tyr Tyr Phe Gly Glu  Thr Tyr Thr Ile Glu  Thr Gly Trp
    1325                1330                1335

Ile Tyr  Asp Met Glu Asn Glu  Ser Asp Lys Tyr Tyr  Phe Asn Pro
    1340                1345                1350

Glu Thr  Lys Lys Ala Cys Lys  Gly Ile Asn Leu Ile  Asp Asp Ile
    1355                1360                1365

Lys Tyr  Tyr Phe Asp Glu Lys  Gly Ile Met Arg Thr  Gly Leu Ile
    1370                1375                1380

Ser Phe  Glu Asn Asn Asn Tyr  Tyr Phe Asn Glu Asn  Gly Glu Met
    1385                1390                1395
```

```
Gln Phe  Gly Tyr Ile Asn Ile  Glu Asp Lys Met Phe  Tyr Phe Gly
    1400                1405                1410

Glu Asp  Gly Val Met Gln Ile  Gly Val Phe Asn Thr  Pro Asp Gly
    1415                1420                1425

Phe Lys  Tyr Phe Ala His Gln  Asn Thr Leu Asp Glu  Asn Phe Glu
    1430                1435                1440

Gly Glu  Ser Ile Asn Tyr Thr  Gly Trp Leu Asp Leu  Asp Glu Lys
    1445                1450                1455

Arg Tyr  Tyr Phe Thr Asp Glu  Tyr Ile Ala Ala Thr  Gly Ser Val
    1460                1465                1470

Ile Ile  Asp Gly Glu Glu Tyr  Tyr Phe Asp Pro Asp  Thr Ala Gln
    1475                1480                1485

Leu Val  Ile Ser Glu
    1490
```

<210> 19
<211> 3552
<212> DNA
<213> artificial

<220>
<223> C-TANCTB fusion DNA

<400> 19

```
atggtaacag gagtatttaa aggacctaat ggatttgagt attttgcacc tgctaatact   60
cacaataata acatagaagg tcaggctata gtttaccaga acaaattctt aactttgaat  120
ggcaaaaaat attattttga taatgactca aaagcagtta ctggatggca aaccattgat  180
ggtaaaaaat attactttaa tcttaacact gctgaagcag ctactggatg gcaaactatt  240
gatggtaaaa aatattactt taatcttaac actgctgaag cagctactgg atggcaaact  300
attgatggta aaaaatatta ctttaatact aacactttca tagcctcaac tggttataca  360
agtattaatg gtaaacattt ttattttaat actgatggta ttatgcagat aggagtgttt  420
aaaggaccta atggatttga atactttgca cctgctaata cggatgctaa caacatagaa  480
ggtcaagcta tactttacca aaataaattc ttaactttga atggtaaaaa atattacttt  540
ggtagtgact caaaagcagt taccggactg cgaactattg atggtaaaaa atattacttt  600
aatactaaca ctgctgttgc agttactgga tggcaaacta ttaatggtaa aaaatactac  660
tttaatacta acacttctat agcttcaact ggttatacaa ttattagtgg taaacatttt  720
```

tattttaata ctgatggtat tatgcagata ggagtgttta aaggacctga tggatttgaa 780

tactttgcac ctgctaatac agatgctaac aatatagaag gtcaagctat acgttatcaa 840

aatagattcc tatatttaca tgacaatata tattattttg gtaataattc aaaagcggct 900

actggttggg taactattga tggtaataga tattacttcg agcctaatac agctatgggt 960

gcgaatggtt ataaaactat tgataataaa aatttttact ttagaaatgg tttacctcag 1020

ataggagtgt ttaaagggtc taatggattt gaatactttg cacctgctaa tacggatgct 1080

aacaatatag aaggtcaagc tatacgttat caaaatagat tcctacattt acttggaaaa 1140

atatattact ttggtaataa ttcaaaagca gttactggat ggcaaactat taatggtaaa 1200

gtatattact ttatgcctga tactgctatg gctgcagctg gtggactttt cgagattgat 1260

ggtgttatat atttctttgg tgttgatgga gtaaaagccc ctgggatata tggcagatct 1320

atgcataatt tgataactgg atttgtgact gtaggcgatg ataaatacta ctttaatcca 1380

attaatggtg gagctgcttc aattggagag acaataattg atgacaaaaa ttattatttc 1440

aaccaaagtg gagtgttaca aacaggtgta tttagtacag aagatggatt taaatatttt 1500

gccccagcta atacacttga tgaaaaccta gaaggagaag caattgattt tactggaaaa 1560

ttaattattg acgaaaatat ttattatttt gatgataatt atagaggagc tgtagaatgg 1620

aaagaattag atggtgaaat gcactatttt agcccagaaa caggtaaagc ttttaaaggt 1680

ctaaatcaaa taggtgatta taaatactat ttcaattctg atggagttat gcaaaaagga 1740

tttgttagta taaatgataa taaacactat tttgatgatt ctggtgttat gaaagtaggt 1800

tacactgaaa tagatggcaa gcatttctac tttgctgaaa acggagaaat gcaaatagga 1860

gtatttaata cagaagatgg atttaaatat tttgctcatc ataatgaaga tttaggaaat 1920

gaagaaggtg aagaaatctc atattctggt atattaaatt tcaataataa aatttactat 1980

tttgatgatt catttacagc tgtagttgga tggaaagatt tagaggatgg ttcaaagtat 2040

tattttgatg aagatacagc agaagcatat ataggtttgt cattaataaa tgatggtcaa 2100

tattatttta atgatgatgg aattatgcaa gttggatttg tcactataaa tgataaagtc 2160

ttctacttct ctgactctgg aattatagaa tctggagtac aaaacataga tgacaattat 2220

ttctatatag atgataatgg tatagttcaa attggtgtat ttgatacttc agatggatat 2280

aaatattttg cacctgctaa tactgtaaat gataatattt acggacaagc agttgaatat 2340

agtggtttag ttagagttgg ggaagatgta tattattttg gagaaacata tacaattgag 2400

actggatgga tatatgatat ggaaaatgaa agtgataaat attatttcaa tccagaaact 2460

aaaaaagcat gcaaaggtat taatttaatt gatgatataa aatattattt tgatgagaag 2520

ggcataatga gaacgggtct tatatcattt gaaaataata attattactt taatgagaat 2580

ggtgaaatgc aatttggtta tataaatata gaagataaga tgttctattt tggtgaagat 2640

ggtgtcatgc agattggagt atttaataca ccagatggat ttaaatactt tgcacatcaa 2700

aatactttgg atgagaattt tgagggagaa tcaataaact atactggttg gttagattta 2760

gatgaaaaga gatattattt tacagatgaa tatattgcag caactggttc agttattatt 2820

gatggtgagg agtattattt tgatcctgat acagctcaat tagtgattag tgaactcgag 2880

ggattaatat atattaatga ttcattatat tattttaaac caccagtaaa taatttgata 2940

actggatttg tgactgtagg cgatgataaa tactacttta atccaattaa tggtggagct 3000

gcttcaattg gagagacaat aattgatgac aaaaattatt atttcaacca aagtggagtg 3060

ttacaaacag gtgtatttag tacagaagat ggatttaaat attttgcccc agctaataca 3120

cttgatgaaa acctagaagg agaagcaatt gattttactg gaaaattaat tattgacgaa 3180

aatatttatt attttgatga taattataga ggagctgtag aatggaaaga attagatggt 3240

gaaatgcact attttagccc agaaacaggt aaagctttta aaggtctaaa tcaaataggt 3300

gattataaat actatttcaa ttctgatgga gttatgcaaa aaggatttgt tagtataaat 3360

gataataaac actattttga tgattctggt gttatgaaag taggttacac tgaaatagat 3420

ggcaagcatt tctactttgc tgaaaacgga gaaatgcaaa taggagtatt taatacagaa 3480

gatggattta aatattttgc tcatcataat gaagatttag gaaatgaaga aggtgaagaa 3540

atctcatatt ct 3552

<210> 20
<211> 1184
<212> PRT
<213> artificial

<220>
<223> C-TANCTB fusion protein

<400> 20

Met Val Thr Gly Val Phe Lys Gly Pro Asn Gly Phe Glu Tyr Phe Ala
1 5 10 15

Pro Ala Asn Thr His Asn Asn Asn Ile Glu Gly Gln Ala Ile Val Tyr
20 25 30

Gln Asn Lys Phe Leu Thr Leu Asn Gly Lys Lys Tyr Tyr Phe Asp Asn
35 40 45

Asp Ser Lys Ala Val Thr Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr
50 55 60

Tyr Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr Gly Trp Gln Thr Ile
65 70 75 80

```
Asp Gly Lys Lys Tyr Tyr Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr
                85                  90                  95

Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr
            100                 105                 110

Phe Ile Ala Ser Thr Gly Tyr Thr Ser Ile Asn Gly Lys His Phe Tyr
        115                 120                 125

Phe Asn Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro Asn
    130                 135                 140

Gly Phe Glu Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu
145                 150                 155                 160

Gly Gln Ala Ile Leu Tyr Gln Asn Lys Phe Leu Thr Leu Asn Gly Lys
                165                 170                 175

Lys Tyr Tyr Phe Gly Ser Asp Ser Lys Ala Val Thr Gly Leu Arg Thr
            180                 185                 190

Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Ala Val Ala Val
        195                 200                 205

Thr Gly Trp Gln Thr Ile Asn Gly Lys Lys Tyr Tyr Phe Asn Thr Asn
    210                 215                 220

Thr Ser Ile Ala Ser Thr Gly Tyr Thr Ile Ile Ser Gly Lys His Phe
225                 230                 235                 240

Tyr Phe Asn Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro
                245                 250                 255

Asp Gly Phe Glu Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile
            260                 265                 270

Glu Gly Gln Ala Ile Arg Tyr Gln Asn Arg Phe Leu Tyr Leu His Asp
        275                 280                 285

Asn Ile Tyr Tyr Phe Gly Asn Asn Ser Lys Ala Ala Thr Gly Trp Val
    290                 295                 300

Thr Ile Asp Gly Asn Arg Tyr Tyr Phe Glu Pro Asn Thr Ala Met Gly
305                 310                 315                 320

Ala Asn Gly Tyr Lys Thr Ile Asp Asn Lys Asn Phe Tyr Phe Arg Asn
                325                 330                 335
```

```
Gly Leu Pro Gln Ile Gly Val Phe Lys Gly Ser Asn Gly Phe Glu Tyr
            340                 345                 350

Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu Gly Gln Ala Ile
        355                 360                 365

Arg Tyr Gln Asn Arg Phe Leu His Leu Leu Gly Lys Ile Tyr Tyr Phe
    370                 375                 380

Gly Asn Asn Ser Lys Ala Val Thr Gly Trp Gln Thr Ile Asn Gly Lys
385                 390                 395                 400

Val Tyr Tyr Phe Met Pro Asp Thr Ala Met Ala Ala Ala Gly Gly Leu
                405                 410                 415

Phe Glu Ile Asp Gly Val Ile Tyr Phe Phe Gly Val Asp Gly Val Lys
            420                 425                 430

Ala Pro Gly Ile Tyr Gly Arg Ser Met His Asn Leu Ile Thr Gly Phe
        435                 440                 445

Val Thr Val Gly Asp Asp Lys Tyr Tyr Phe Asn Pro Ile Asn Gly Gly
    450                 455                 460

Ala Ala Ser Ile Gly Glu Thr Ile Ile Asp Asp Lys Asn Tyr Tyr Phe
465                 470                 475                 480

Asn Gln Ser Gly Val Leu Gln Thr Gly Val Phe Ser Thr Glu Asp Gly
                485                 490                 495

Phe Lys Tyr Phe Ala Pro Ala Asn Thr Leu Asp Glu Asn Leu Glu Gly
            500                 505                 510

Glu Ala Ile Asp Phe Thr Gly Lys Leu Ile Ile Asp Glu Asn Ile Tyr
        515                 520                 525

Tyr Phe Asp Asp Asn Tyr Arg Gly Ala Val Glu Trp Lys Glu Leu Asp
    530                 535                 540

Gly Glu Met His Tyr Phe Ser Pro Glu Thr Gly Lys Ala Phe Lys Gly
545                 550                 555                 560

Leu Asn Gln Ile Gly Asp Tyr Lys Tyr Tyr Phe Asn Ser Asp Gly Val
                565                 570                 575

Met Gln Lys Gly Phe Val Ser Ile Asn Asp Asn Lys His Tyr Phe Asp
```

580 585 590

Asp Ser Gly Val Met Lys Val Gly Tyr Thr Glu Ile Asp Gly Lys His
595 600 605

Phe Tyr Phe Ala Glu Asn Gly Glu Met Gln Ile Gly Val Phe Asn Thr
610 615 620

Glu Asp Gly Phe Lys Tyr Phe Ala His His Asn Glu Asp Leu Gly Asn
625 630 635 640

Glu Glu Gly Glu Glu Ile Ser Tyr Ser Gly Ile Leu Asn Phe Asn Asn
645 650 655

Lys Ile Tyr Tyr Phe Asp Asp Ser Phe Thr Ala Val Val Gly Trp Lys
660 665 670

Asp Leu Glu Asp Gly Ser Lys Tyr Tyr Phe Asp Glu Asp Thr Ala Glu
675 680 685

Ala Tyr Ile Gly Leu Ser Leu Ile Asn Asp Gly Gln Tyr Tyr Phe Asn
690 695 700

Asp Asp Gly Ile Met Gln Val Gly Phe Val Thr Ile Asn Asp Lys Val
705 710 715 720

Phe Tyr Phe Ser Asp Ser Gly Ile Ile Glu Ser Gly Val Gln Asn Ile
725 730 735

Asp Asp Asn Tyr Phe Tyr Ile Asp Asp Asn Gly Ile Val Gln Ile Gly
740 745 750

Val Phe Asp Thr Ser Asp Gly Tyr Lys Tyr Phe Ala Pro Ala Asn Thr
755 760 765

Val Asn Asp Asn Ile Tyr Gly Gln Ala Val Glu Tyr Ser Gly Leu Val
770 775 780

Arg Val Gly Glu Asp Val Tyr Tyr Phe Gly Glu Thr Tyr Thr Ile Glu
785 790 795 800

Thr Gly Trp Ile Tyr Asp Met Glu Asn Glu Ser Asp Lys Tyr Tyr Phe
805 810 815

Asn Pro Glu Thr Lys Lys Ala Cys Lys Gly Ile Asn Leu Ile Asp Asp
820 825 830

```
Ile Lys Tyr Tyr Phe Asp Glu Lys Gly Ile Met Arg Thr Gly Leu Ile
        835                 840                 845

Ser Phe Glu Asn Asn Asn Tyr Tyr Phe Asn Glu Asn Gly Glu Met Gln
    850                 855                 860

Phe Gly Tyr Ile Asn Ile Glu Asp Lys Met Phe Tyr Phe Gly Glu Asp
865                 870                 875                 880

Gly Val Met Gln Ile Gly Val Phe Asn Thr Pro Asp Gly Phe Lys Tyr
                885                 890                 895

Phe Ala His Gln Asn Thr Leu Asp Glu Asn Phe Glu Gly Glu Ser Ile
            900                 905                 910

Asn Tyr Thr Gly Trp Leu Asp Leu Asp Glu Lys Arg Tyr Tyr Phe Thr
        915                 920                 925

Asp Glu Tyr Ile Ala Ala Thr Gly Ser Val Ile Ile Asp Gly Glu Glu
    930                 935                 940

Tyr Tyr Phe Asp Pro Asp Thr Ala Gln Leu Val Ile Ser Glu Leu Glu
945                 950                 955                 960

Gly Leu Ile Tyr Ile Asn Asp Ser Leu Tyr Tyr Phe Lys Pro Pro Val
                965                 970                 975

Asn Asn Leu Ile Thr Gly Phe Val Thr Val Gly Asp Asp Lys Tyr Tyr
            980                 985                 990

Phe Asn Pro Ile Asn Gly Gly Ala  Ala Ser Ile Gly Glu  Thr Ile Ile
        995                 1000                1005

Asp Asp  Lys Asn Tyr Tyr Phe  Asn Gln Ser Gly Val  Leu Gln Thr
    1010                1015                1020

Gly Val  Phe Ser Thr Glu Asp  Gly Phe Lys Tyr Phe  Ala Pro Ala
    1025                1030                1035

Asn Thr  Leu Asp Glu Asn Leu  Glu Gly Glu Ala Ile  Asp Phe Thr
    1040                1045                1050

Gly Lys  Leu Ile Ile Asp Glu  Asn Ile Tyr Tyr Phe  Asp Asp Asn
    1055                1060                1065

Tyr Arg  Gly Ala Val Glu Trp  Lys Glu Leu Asp Gly  Glu Met His
    1070                1075                1080
```

```
Tyr Phe  Ser Pro Glu Thr Gly  Lys Ala Phe Lys Gly  Leu Asn Gln
    1085                 1090                 1095

Ile Gly  Asp Tyr Lys Tyr Tyr  Phe Asn Ser Asp Gly  Val Met Gln
    1100                 1105                 1110

Lys Gly  Phe Val Ser Ile Asn  Asp Asn Lys His Tyr  Phe Asp Asp
    1115                 1120                 1125

Ser Gly  Val Met Lys Val Gly  Tyr Thr Glu Ile Asp  Gly Lys His
    1130                 1135                 1140

Phe Tyr  Phe Ala Glu Asn Gly  Glu Met Gln Ile Gly  Val Phe Asn
    1145                 1150                 1155

Thr Glu  Asp Gly Phe Lys Tyr  Phe Ala His His Asn  Glu Asp Leu
    1160                 1165                 1170

Gly Asn  Glu Glu Gly Glu Glu  Ile Ser Tyr Ser
    1175                 1180
```

Claims

1. An immunogenic or vaccine composition comprising an isolated polypeptide having at least 85% sequence identity to the amino acid sequence as set forth in SEQ ID NO: 4 and a pharmaceutically acceptable carrier or excipient.

2. The composition of claim 1, wherein said polypeptide provides 100 % survival of hamsters vaccinated with said polypeptide after intragastric administration of a dose of 10^2, 10^3 or 10^4 *Clostridium difficile* spores.

3. The composition of claim 1 or 2, wherein the polypeptide comprises 19 repeating units derived from the C-terminal domain of toxin A of *Clostridium difficile.*

4. The composition of any one of claims 1 to 3, wherein the polypeptide comprises 23 repeating units derived from the C-terminal domain of toxin B of *Clostridium difficile.*

5. The composition of any one of claims 1 to 4, wherein the polypeptide has at least 90%, more preferably 95%, most preferably 99% sequence identity to the amino acid sequence as set forth in SEQ ID NO: 4.

6. The composition of any one of claims 1 to 5, wherein the polypeptide has the sequence as set forth in SEQ ID NO: 2 or SEQ ID NO: 4.

7. The composition of any one of claims 1 to 6, further comprising an adjuvant.

8. The composition of claim 7, wherein the adjuvant comprises alum.

9. An isolated polypeptide having at least 85% sequence identity to the amino acid sequence as set forth in SEQ ID NO: 4, or the composition of any one of claims 1 to 8, for use in the prevention or treatment of a *C. difficile* associated disease (CDAD).

10. The polypeptide or composition for use of claim 9, for use in the prevention of CDAD in a subject at risk of a CDAD.

11. The polypeptide or composition for use of claim 10, wherein said subject at risk of CDAD is: i) an elderly subject; ii)

a subject with AIDS; iii) a subject taking or planning to take immunosuppressing drugs; iv) a subject with planned hospitalization or a subject that is in hospital; v) a subject in or expected to go to an intensive care unit; vi) a subject that is undergoing or is planning to undergo gastrointestinal surgery; vii) a subject that is in or planning to go to a long-term care such as a nursing home; viii) a subject with co-morbidities requiring frequent and/or prolonged antibiotic use; ix) a subject with recurrent CDAD; x) a subject below 18 years of age; xi) a subject from 18 to 65 years of age; or xii) a subject over 65 years of age.

12. The polypeptide or composition for use of claim 10, wherein said subject at risk of CDAD is over 65 years.

13. The polypeptide or composition for use of claim 9, wherein the treatment of CDAD includes reducing the severity and/or duration of disease, ameliorating symptoms of disease, or combinations thereof.

14. The composition of any one of claims 1 to 13, further comprising an additional antigen.

15. The composition of any one of claims 1 to 14, further comprising a therapeutic agent selected from the group consisting of anesthetics, analgesics, anti-inflammatories, steroids, antibiotics, antiarthritics, anorectics, antihistamines and antineoplastics.

Patentansprüche

1. Immunisierende oder Impfstoff-Zusammensetzung, umfassend ein isoliertes Polypeptid mit mindestens 85% Sequenzidentität zu der Aminosäure-Sequenz, wie in SEQ ID NR.: 4 angeführt, und einen pharmazeutisch verträglichen Träger oder Exzipienten.

2. Zusammensetzung nach Anspruch 1, wobei das Polypeptid 100 % Überleben von Hamstern, beimpft mit dem Polypeptid nach intragastrischer Verabreichung einer Dosis von 10^2, 10^3 oder 10^4 *Clostridium difficile*-Sporen, bereitstellt.

3. Zusammensetzung nach Anspruch 1 oder 2, wobei das Polypeptid 19 wiederkehrende Einheiten, abgeleitet von der C-terminalen Domäne von Toxin A von *Clostridium difficile,* umfasst.

4. Zusammensetzung nach einem der Ansprüche 1 bis 3, wobei das Polypeptid 23 wiederkehrende Einheiten, abgeleitet von der C-terminalen Domäne von Toxin B von *Clostridium difficile,* umfasst.

5. Zusammensetzung nach einem der Ansprüche 1 bis 4, wobei das Polypeptid mindestens 90%, bevorzugter 95%, besonders bevorzugt 99% Sequenzidentität zu der Aminosäure-Sequenz, wie in SEQ ID NR.: 4 angeführt, aufweist.

6. Zusammensetzung nach einem der Ansprüche 1 bis 5, wobei das Polypeptid die Sequenz, wie in SEQ ID NR.: 2 oder SEQ ID NR.: 4 angeführt, aufweist.

7. Zusammensetzung nach einem der Ansprüche 1 bis 6, weiterhin umfassend ein Adjuvans.

8. Zusammensetzung nach Anspruch 7, wobei das Adjuvans Alaun umfasst.

9. Isoliertes Polypeptid mit mindestens 85% Sequenzidentität zu der Aminosäure-Sequenz, wie in SEQ ID NR.: 4 angeführt, oder die Zusammensetzung nach einem der Ansprüche 1 bis 8, zur Verwendung bei der Prävention oder Behandlung von einer mit C. *difficile* verbundenen Krankheit (CDAD).

10. Polypeptid oder Zusammensetzung zur Verwendung nach Anspruch 9, zur Verwendung bei der Prävention von CDAD bei einer Person mit Risiko für eine CDAD.

11. Polypeptid oder Zusammensetzung zur Verwendung nach Anspruch 10, wobei die Person mit Risiko für CDAD ist: i) eine ältere Person; ii) eine Person mit AIDS; iii) eine Person, die immunosuppressive Arzneimittel einnimmt oder plant, einzunehmen; iv) eine Person mit geplantem Krankenhausaufenthalt oder eine Person, die sich in einem Krankenhaus befindet; v) eine Person in der Intensivstation oder von der erwartet wird, dort eingeliefert zu werden; vi) eine Person, an der eine Gastrointestinal-Operation vorgenommen wird oder die plant, an sich eine vornehmen zu lassen; vii) eine Person, die sich in Langzeitpflege, wie ein Pflegeheim, befindet oder plant, in eine zu gehen;

viii) eine Person mit Komorbiditäten, die häufige und/oder längere Anwendungen von Antibiotika erfordern; ix) eine Person mit wiederkehrender CDAD; x) eine Person im Alter von unter 18 Jahren; xi) eine Person im Alter von 18 bis 65 Jahren; oder xii) eine Person im Alter von über 65 Jahren.

12. Polypeptid oder Zusammensetzung zur Verwendung nach Anspruch 10, wobei die Person mit Risiko für CDAD über 65 Jahre ist.

13. Polypeptid oder Zusammensetzung zur Verwendung nach Anspruch 9, wobei die Behandlung von CDAD Vermindern der Schwere und/oder Dauer der Krankheit, Lindern der Symptome der Krankheit oder Kombinationen davon einschließt.

14. Zusammensetzung nach einem der Ansprüche 1 bis 13, weiterhin umfassend ein zusätzliches Antigen.

15. Zusammensetzung nach einem der Ansprüche 1 bis 14, weiterhin umfassend ein therapeutisches Mittel, ausgewählt aus der Gruppe, bestehend aus Anästhetika, Analgetika, Entzündungshemmern, Steroiden, Antibiotika, Antiarthritika, Anorektika, Antihistaminika und Antineoplastika.

Revendications

1. Composition immunogène ou vaccinale comprenant un polypeptide isolé ayant au moins 85% d'identité de séquence avec la séquence d'acides aminés telle que représentée par SEQ ID NO : 4 et un véhicule ou excipient pharmaceutiquement acceptable.

2. Composition selon la revendication 1, dans laquelle ledit polypeptide donne 100% de survie chez des hamsters vaccinés avec ledit polypeptide après une administration intragastrique d'une dose de 10^2, 10^3 ou 10^4 spores de *Clostridium difficile.*

3. Composition selon la revendication 1 ou 2, dans laquelle le polypeptide comprend 19 unités répétées dérivées du domaine C-terminal de la toxine A de *Clostridium difficile.*

4. Composition selon l'une quelconque des revendications 1 à 3, dans laquelle le polypeptide comprend 23 unités répétées dérivées du domaine C-terminal de la toxine B de *Clostridium difficile.*

5. Composition selon l'une quelconque des revendications 1 à 4, dans laquelle le polypeptide a au moins 90%, plus préférablement 95%, le plus préférablement 99% d'identité de séquence avec la séquence d'acide aminé telle que représentée par SEQ ID NO : 4.

6. Composition selon l'une quelconque des revendications 1 à 5, dans laquelle le polypeptide a la séquence telle que représentée par SEQ ID NO : 2 ou SEQ ID NO : 4.

7. Composition selon l'une quelconque des revendications 1 à 6, comprenant en outre un adjuvant.

8. Composition selon la revendication 7, dans laquelle l'adjuvant comprend de l'alun.

9. Peptide isolé ayant au moins 85% d'identité de séquence avec la séquence d'acide aminé telle que représentée par SEQ ID NO : 4, ou la composition selon l'une quelconque des revendications 1 à 8, pour une utilisation dans la prévention ou le traitement d'une maladie associée à *Clostridium difficile* (MACD).

10. Polypeptide ou composition pour l'utilisation selon la revendication 9, pour une utilisation dans la prévention d'une MACD chez un sujet à risque d'une MACD.

11. Polypeptide ou composition pour l'utilisation selon la revendication 10, dans lequel ledit sujet à risque d'une MACD est : i) un sujet âgé, ii) un sujet avec le SIDA; iii) un sujet prenant ou envisageant de prendre des médicaments immunosuppresseurs ; iv) un sujet avec une hospitalisation planifiée ou un sujet qui est à l'hôpital ; v) un sujet dans où devant aller dans une unité de soins intensifs; vi) un sujet qui est en train de subir ou qui envisage de subir une chirurgie gastro-intestinale ; vii) un sujet qui est dans ou qui prévoit d'aller dans une unité de soins de longue durée telle qu'une maison de repos; viii) un sujet avec des comorbidités nécessitant une utilisation prolongée ou fréquente

d'antibiotiques ; ix) un sujet avec des MACD récurrentes ; x) un sujet de moins de 18 ans ; xi) un sujet de 18 à 65 ans ; ou xii) un sujet de plus de 65 ans.

12. Polypeptide ou composition pour l'utilisation selon la revendication 10, dans lequel ledit sujet à risque d'une MACD a plus de 65 ans.

13. Polypeptide ou composition pour l'utilisation selon la revendication 9, dans lequel le traitement de la MACD inclut réduire la sévérité et/ou la durée de la maladie, améliorer les symptômes de la maladie, ou les combinaisons de ceux-ci.

14. Composition selon l'une quelconque des revendications 1 à 13, comprenant en outre un antigène additionnel.

15. Composition selon l'une quelconque des revendications 1 à 14, comprenant en outre un agent thérapeutique sélectionné dans le groupe constitué d'anesthésiques, d'analgésiques, d'anti-inflammatoires, de stéroïdes, d'antibiotiques, d'antiarthritiques, d'anorexigènes, d'antihistaminiques et d'antinéoplasiques.

Figure 1

A. C-TAB.G5 nucleic acid

ATGGTAACAGGAGTATTTAAAGGACCTAATGGATTTGAGTATTTTGCACCTGCTAATACTCACAATAATAACATAGAAGGTCAG
GCTATAGTTTACCAGAACAAATTCTTAACTTTGAATGGCAAAAAATATTATTTTGATAATGACTCAAAAGCAGTTACTGGATGG
CAAACCATTGATGGTAAAAAATATTACTTTAATCTTAACACTGCTGAAGCAGCTACTGGATGGCAAACTATTGATGGTAAAAAA
TATTACTTTAATCTTAACACTGCTGAAGCAGCTACTGGATGGCAAACTATTGATGGTAAAAAATATTACTTTAATACTAACACTT
TCATAGCCTCAACTGGTTATACAAGTATTAATGGTAAACATTTTTATTTTAATACTGATGGTATTATGCAGATAGGAGTGTTTAA
AGGACCTAATGGATTTGAATACTTTGCACCTGCTAATACTCATAATAACAACATAGAAGGTCAAGCTATACTTTACCAAAATAA
ATTCTTAACTTTGAATGGTAAAAAATATTACTTTGGTAGTGACTCAAAAGCAGTTACCGGATTGCGAACTATTGATGGTAAAAA
ATATTACTTTAATACTAACACTGCTGTTGCAGTTACTGGATGGCAAACTATTAATGGTAAAAAATACTACTTTAATACTAACACT
TCTATAGCTTCAACTGGTTATACAATTATTAGTGGTAAACATTTTTATTTTAATACTGATGGTATTATGCAGATAGGAGTGTTTAA
AGGACCTGATGGATTTGAATACTTTGCACCTGCTAATACAGATGCTAACAATATAGAAGGTCAAGCTATACGTTATCAAAATAG
ATTCCTATATTTACATGACAATATATATTATTTTGGTAATAATTCAAAAGCAGCTACTGGTTGGGTAACTATTGATGGTAATAGA
TATTACTTCGAGCCTAATACAGCTATGGGTGCGAATGGTTATAAAACTATTGATAATAAAAATTTTTACTTTAGAAATGGTTTAC
CTCAGATAGGAGTGTTTAAAGGGTCTAATGGATTTGAATACTTTGCACCTGCTAATACGGATGCTAACAATATAGAAGGTCAA
GCTATACGTTATCAAAATAGATTCCTACATTTACTTGGAAAAATATATTACTTTGGTAATAATTCAAAAGCAGTTACTGGATGGC
AAACTATTAATGGTAAAGTATATTACTTTATGCCTGATACTGCTATGGCTGCAGCTGGTGGACTTTTCGAGATTGATGGTGTTAT
ATATTTCTTTGGTGTTGATGGAGTAAAAGCCCCTGGGATATATGGCAGATCTATGCATAATTTGATAACTGGATTTGTGACTGT
AGGCGATGATAAATACTACTTTAATCCAATTAATGGTGGAGCTGCTTCAATTGGAGAGACAATAATTGATGACAAAAATTATTA
TTTCAACCAAAGTGGAGTGTTACAAACAGGTGTATTTAGTACAGAAGATGGATTTAAATATTTTGCCCCAGCTAATACACTTGA
TGAAAACCTAGAAGGAGAAGCAATTGATTTTACTGGAAAATTAATTATTGACGAAAATATTTATTATTTTGATGATAATTATAGA
GGAGCTGTAGAATGGAAAGAATTAGATGGTGAAATGCACTATTTTAGCCCAGAAACAGGTAAAGCTTTTAAAGGTCTAAATCA
AATAGGTGATTATAAATACTATTTCAATTCTGATGGAGTTATGCAAAAAGGATTTGTTAGTATAAATGATAATAAACACTATTTT
GATGATTCTGGTGTTATGAAAGTAGGTTACACTGAAATAGATGGCAAGCATTTCTACTTTGCTGAAAACGGAGAAATGCAAAT
AGGAGTATTTAATACAGAAGATGGATTTAAATATTTTGCTCATCATAATGAAGATTTAGGAAATGAAGAAGGTGAAGAAATCTC
ATATTCTGGTATATTAAATTTCAATAATAAAATTTACTATTTTGATGATTCATTTACAGCTGTAGTTGGATGGAAAGATTTAGAG
GATGGTTCAAAGTATTATTTTGATGAAGATACAGCAGAAGCATATATAGGTTTGTCATTAATAAATGATGGTCAATATTATTTTA
ATGATGATGGAATTATGCAAGTTGGATTTGTCACTATAAATGATAAAGTCTTCTACTTCTCTGACTCTGGAATTATAGAATCTGG
AGTACAAAACATAGATGACAATTATTTCTATATAGATGATAATGGTATAGTTCAAATTGGTGTATTTGATACTTCAGATGGATAT
AAATATTTTGCACCTGCTAATACTGTAAATGATAATATTTACGGACAAGCAGTTGAATATAGTGGTTTAGTTAGAGTTGGGGAA
GATGTATATTATTTTGGAGAAACATATACAATTGAGACTGGATGGATATATGATATGGAAAATGAAAGTGATAAATATTATTTC
AATCCAGAAACTAAAAAAGCATGCAAAGGTATTAATTTAATTGATGATATAAAATATTATTTTGATGAGAAGGGCATAATGAGA
ACGGGTCTTATATCATTTGAAAATAATAATTATTACTTTAATGAGAATGGTGAAATGCAATTTGGTTATATAAATATAGAAGATA
AGATGTTCTATTTTGGTGAAGATGGTGTCATGCAGATTGGAGTATTTAATACACCAGATGGATTTAAATACTTTGCACATCAAA
ATACTTTGGATGAGAATTTTGAGGGAGAATCAATAAACTATACTGGTTGGTTAGATTTAGATGAAAAGAGATATTATTTTACAG
ATGAATATATTGCAGCAACTGGTTCAGTTATTATTGATGGTGAGGAGTATTATTTTGATCCTGATACAGCTCAATTAGTGATTAG
TGAATAG

EP 2 753 352 B1

Figure 1 (cont.)

B. C-TAB.G5 amino acid

MVTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTIDG
KKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSING
KHFYFNTDGIMQIGVFKGPNGFEYFAPANTHNNNIEGQAILYQNKFLTLNGKKYYFGSDSKA
VTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQI
GVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRYY
FEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNRF
LHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGIY
GRSMHNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDGFKYF
APANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLN
QIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFN
TEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDED
TAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIVQ
IGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENESD
KYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMFY
FGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGSV
IIDGEEYYFDPDTAQLVISE*

Figure 2

A. C-TAB.G5 .1 nucleic acid

ATGGTTACAGGTGTTTTCAAAGGTCCGAACGGCTTTGAATATTTTGCACCGGCAAATACCCACAATAATAATATTGAAGGCCAGGCC
ATCGTGTATCAGAATAAATTTCTGACCCTGAACGGCAAAAAATACTATTTCGATAACGATAGCAAAGCAGTTACCGGTTGGCAAACC
ATTGATGGCAAAAAATATTACTTCAACCTGAATACCGCAGAAGCAGCAACCGGCTGGCAGACGATCGACGGTAAAAAGTACTATTTT
AACCTGAACACAGCCGAAGCCGCTACAGGCTGGCAGACAATAGATGGGAAGAAGTATTATTTTAATACCAATACCTTTATTGCCAGC
ACCGGCTATACCAGCATTAATGGCAAACACTTCTATTTTAACACCGATGGTATTATGCAGATCGGTGTGTTTAAGGGCCCTAATGGT
TTTGAGTACTTCGCTCCGGCTAATACCGATGCAAATAACATCGAAGGTCAGGCAATTCTGTACCAGAACAAATTTTTAACGCTGAAC
GGTAAGAAATATTACTTTGGTAGCGATTCAAAAGCCGTTACCGGTCTGCGTACGATCGACGGCAAGAAATATTATTTCAATACAAAC
ACCGCAGTTGCCGTGACAGGTTGGCAGACGATAAATGGTAAGAAGTACTACTTCAACACCAATACCAGCATTGCAAGTACCGGTTA
TACCATTATCAGCGGCAAACACTTTTACTTCAATACAGACGGCATTATGCAGATTGGCGTTTTCAAAGGTCCGGATGGTTTCGAGTA
CTTTGCCCCTGCAAATACAGATGCAAACAATATTGAGGGACAGGCAATTCGCTATCAGAATCGTTTTCTGTATCTGCACGATAACAT
CTATTACTTCGGCAATAATTCAAAAGCAGCCACCGGTTGGGTTACAATTGATGGTAATCGTTATTACTTTGAGCCGAATACCGCAAT
GGGTGCAAATGGTTATAAAACCATCGATAACAAAAATTTTTATTTCCGCAACGGTCTGCCGCAGATTGGTGTTTTTAAGGGTAGCAAT
GGCTTCGAGTATTTTGCGCCAGCCAACACCGATGCCAACAACATTGAAGGCCAAGCGATTCGTTATCAAAACCGCTTTCTGCATCT
GCTGGGCAAAATTTATTACTTTGGCAACAATAGCAAAGCGGTGACGGGCTGGCAAACCATTAACGGTAAAGTTTATTATTTCATGCC
GGATACCGCTATGGCAGCAGCCGGTGGTCTGTTTGAAATTGATGGCGTGATTTATTTTTTTGGCGTGGATGGTGTTAAAGCACCGG
GTATTTATGGTCGTAGCATGCATAATCTGATTACCGGTTTTGTTACCGTGGGCGACGATAAATACTACTTTAATCCGATTAATGGTGG
TGCAGCAAGCATTGGTGAAACCATTATCGATGACAAAAACTATTATTTTAACCAGAGCGGTGTTCTGCAGACAGGTGTTTTTAGCAC
CGAAGATGGCTTCAAATATTTTGCTCCTGCGAATACACTGGATGAAAATCTGGAAGGTGAAGCAATTGATTTTACCGGCAAACTGAT
CATCGACGAGAACATCTACTATTTTGATGATAATTATCGCGGTGCCGTGGAATGGAAAGAACTGGATGGTGAAATGCACTATTTTAG
TCCGGAAACCGGTAAAGCCTTTAAAGGTCTGAATCAGATCGGCGATTACAAGTATTACTTTAATTCAGATGGCGTGATGCAGAAAGG
CTTTGTGAGCATTAACGACAACAAACACTATTTTGACGACAGCGGTGTGATGAAAGTGGGTTATACCGAAATCGACGGGAAACATTT
TTATTTTGCCGAAAACGGCGAAATGCAGATTGGAGTATTTAATACCGAGGACGGCTTTAAATACTTTGCCCATCATAATGAAGATCTG
GGTAATGAAGAAGGCGAAGAAATTAGCTATAGCGGCATTCTGAATTTTAATAACAAGATCTATTATTTCGATGATAGCTTCACCGCAG
TTGTTGGTTGGAAAGATCTGGAAGATGGCAGCAAATATTATTTTGATGAAGATACCGCAGAGGCCTATATTGGTCTGAGCCTGATTA
ATGATGGCCAGTATTATTTCAACGATGATGGTATCATGCAGGTTGGTTTTGTGACCATCAACGATAAAGTGTTCTATTTCAGCGATAG
CGGCATTATTGAAAGCGGTGTTCAGAACATCGACGATAACTATTTCTACATCGATGATAACGGTATTGTTCAGATTGGCGTGTTTGAT
ACCTCCGATGGTTATAAATATTTCGCACCAGCCAATACCGTGAACGATAATATTTATGGTCAGGCAGTTGAATATTCAGGTCTGGTTC
GTGTTGGCGAAGATGTTTATTATTTTGGCGAAACCTATACCATTGAAACCGGCTGGATCTATGATATGGAAAACGAGAGCGACAAGT
ACTATTTCAATCCGGAAACGAAAAAAGCCTGCAAAGGCATTAATCTGATCGACGATATTAAGTACTACTTTGACGAAAAAGGCATTAT
GCGTACCGGTCTGATTAGCTTTGAGAACAACAACTATTACTTCAATGAGAACGGTGAGATGCAGTTTGGCTATATCAACATCGAGGA
CAAAATGTTTTATTTTGGTGAGGACGGTGTGATGCAGATAGGGGTTTTTAATACACCGGATGGGTTTAAGTATTTTGCACATCAGAAC
ACCCTGGATGAAAACTTTGAAGGCGAAAGCATTAATTATACCGGTTGGCTGGATCTGGATGAGAAACGTTATTATTTCACCGACGAA
TACATTGCAGCAACCGGTAGCGTTATTATTGATGGTGAGGAATATTACTTCGATCCGGATACAGCACAGCTGGTTATTAGCGAATAA

Figure 2 (cont.)

B. C-TAB.G5.1 amino acid

(M*)VTGVFKGPNGFEYFAPANTHNNNIEGQAIVYQNKFLTLNGKKYYFDNDSKAVTGWQTID
GKKYYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATGWQTIDGKKYYFNTNTFIASTGYTSIN
GKHFYFNTDGIMQIGVFKGPNGFEYFAPANTDANNIEGQAILYQNKFLTLNGKKYYFGSDSK
AVTGLRTIDGKKYYFNTNTAVAVTGWQTINGKKYYFNTNTSIASTGYTIISGKHFYFNTDGIMQ
IGVFKGPDGFEYFAPANTDANNIEGQAIRYQNRFLYLHDNIYYFGNNSKAATGWVTIDGNRY
YFEPNTAMGANGYKTIDNKNFYFRNGLPQIGVFKGSNGFEYFAPANTDANNIEGQAIRYQNR
FLHLLGKIYYFGNNSKAVTGWQTINGKVYYFMPDTAMAAAGGLFEIDGVIYFFGVDGVKAPGI
YGRSMHNLITGFVTVGDDKYYFNPINGGAASIGETIIDDKNYYFNQSGVLQTGVFSTEDGFKY
FAPANTLDENLEGEAIDFTGKLIIDENIYYFDDNYRGAVEWKELDGEMHYFSPETGKAFKGLN
QIGDYKYYFNSDGVMQKGFVSINDNKHYFDDSGVMKVGYTEIDGKHFYFAENGEMQIGVFN
TEDGFKYFAHHNEDLGNEEGEEISYSGILNFNNKIYYFDDSFTAVVGWKDLEDGSKYYFDED
TAEAYIGLSLINDGQYYFNDDGIMQVGFVTINDKVFYFSDSGIIESGVQNIDDNYFYIDDNGIV
QIGVFDTSDGYKYFAPANTVNDNIYGQAVEYSGLVRVGEDVYYFGETYTIETGWIYDMENES
DKYYFNPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENNNYYFNENGEMQFGYINIEDKMF
YFGEDGVMQIGVFNTPDGFKYFAHQNTLDENFEGESINYTGWLDLDEKRYYFTDEYIAATGS
VIIDGEEYYFDPDTAQLVISE

*the N-terminal Met is cleaved off during expression

Figure 3

Antibody titers on day 14 (1st immunization) and day 28 (2nd immunizations) in mice

	Anti-C-TAB			Anti-Toxin A			Anti-Toxin B		
C-TAB immunization	**Pre***	**Day 14**	**Day 28**	**Pre**	**Day 14**	**Day 28**	**Pre**	**Day 14**	**Day 28**
0	2	1	1	1	2	1	1	1	1
100 (154) ng	5	3	349	4	8	122	1	1	4
300 (462) ng	4	2	1,172	3	4	530	1	1	5
1,000 (1,540) ng	5	144	5,199	9	100	2,382	1	1	4
3.000 (4,620) ng	3	1,105	24,047	2	655	11,563	1	1	284
10,000 (15,400) ng	5	11,195	84,519	4	6,980	48,595	2	46	2,724
PBS + 50 µg alum	9	3	3	2	1	1	1	2	2
10 (15.4) ng + 50 µg alum	6	45,375	123,373	4	21,157	87,492	1	1	436
30 (46.2) ng + 50 µg alum	5	42,410	152,074	4	18,691	111,424	1	1	3,835
100 (154) ng + 50 µg alum	2	44,777	250,705	2	34,660	150,931	1	4	13,568
300 (462) ng + 50 µg alum	2	97,667	389,670	2	23,167	234,546	1	154	20,955
1.000 (1,540) ng + 50 µg alum	8	110,184	268,865	8	26,864	188,545	1	1,558	31,088

* pre-bleed

Figure 4

Comparison of antibody titers in mice

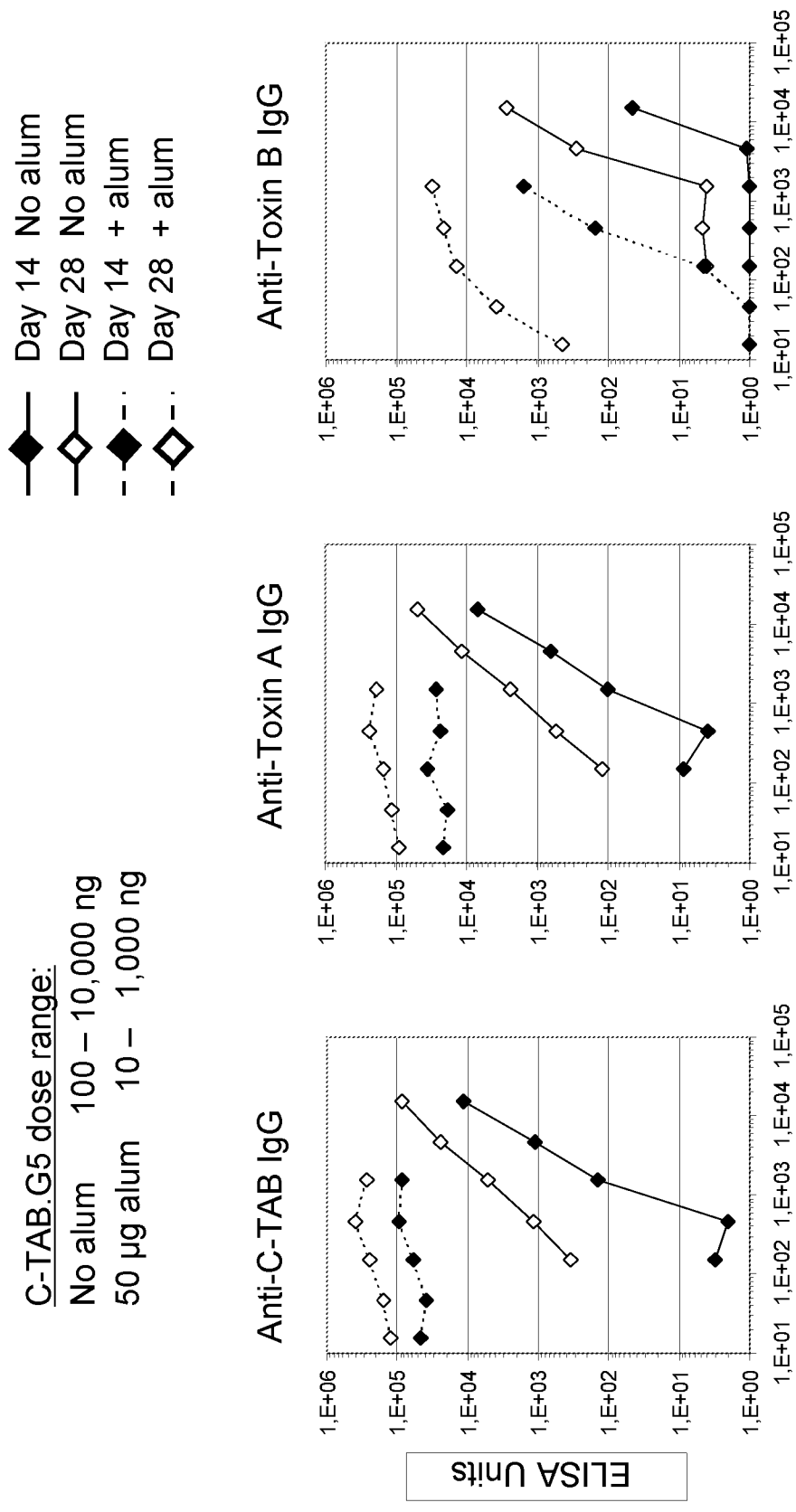

Figure 5

Antibody titers on day 28 (post 2nd immunization) in mice

GP	Vaccine		Antibody titers		
	C-TAB.G5	Adjuvant	Anti-C-TAB	Anti-toxin A	Anti-toxin B
1	PBS	---	0	0	0
2	3 μg	No adjuvant	27,930	19,001	422
3	10 μg	No adjuvant	51,518	34,058	1,464
4	30 μg	No adjuvant	69,836	50,907	1,801
5	3 μg	50 μg alum	584,447	241,362	45,181
6	10 μg	50 μg alum	789,145	311,657	67,462
7	30 μg	50 μg alum	1,081,219	404,399	53,849

Figure 6

TNA and protection against challenge with toxin A or toxin B in mice

Vaccine	Toxin A		Toxin B	
	TNA	Protection*, %	TNA	Protection*, %
PBS	0	16.6	0	0
3 µg C-TAB.G5	103	100	0	0
10 µg C-TAB.G5	171	100	0	12.5
30 µg C-TAB.G5	150	100	0	50
3 µg C-TAB.G5 + alum	683	100	339	100
10 µg C-TAB.G5 + alum	778	100	300	100
30 µg C-TAB.G5 + alum	1010	100	669	87.5

*Challenge dose:
toxin A: 25 ng/mouse
toxin B: 50 ng/mouse

Figure 7

Comparison of C-TAB.G5 immunogenicity in young vs. old mice

Age	C-TAB.5 immunization	IgG titers			TNA		Protection*, %	
		C-TAB	Toxin A	Toxin B	Toxin A	Toxin B	Toxin A	Toxin B
Y	PBS	0	1	1	0	0	25	0
O	PBS	25	4	1	0	0	0	25
Y	10 µg	48,911	34,519	2,338	184	0	0	0
Y	30 µg	91,110	47,559	3,722	285	0	100	0
O	10 µg	3,186	2,970	6	6	0	100	71
O	30 µg	19,836	8,924	151	58	0	62.5	25
Y	10 µg + alum	580,987	302,148	41,603	779	1,428	100	85
Y	30 µg + alum	747,839	451,641	55,705	682	1,042	100	100
O	10 µg + alum	99,981	61,105	2,455	283	0	100	75
O	30 µg + alum	351,373	134,631	11,211	682	0	100	100

*Challenge dose:
toxin A: 25 ng/mouse
toxin B: 50 ng/mouse

Figure 8

Kinetics of anti-C-TAB antibody development in young vs old mice

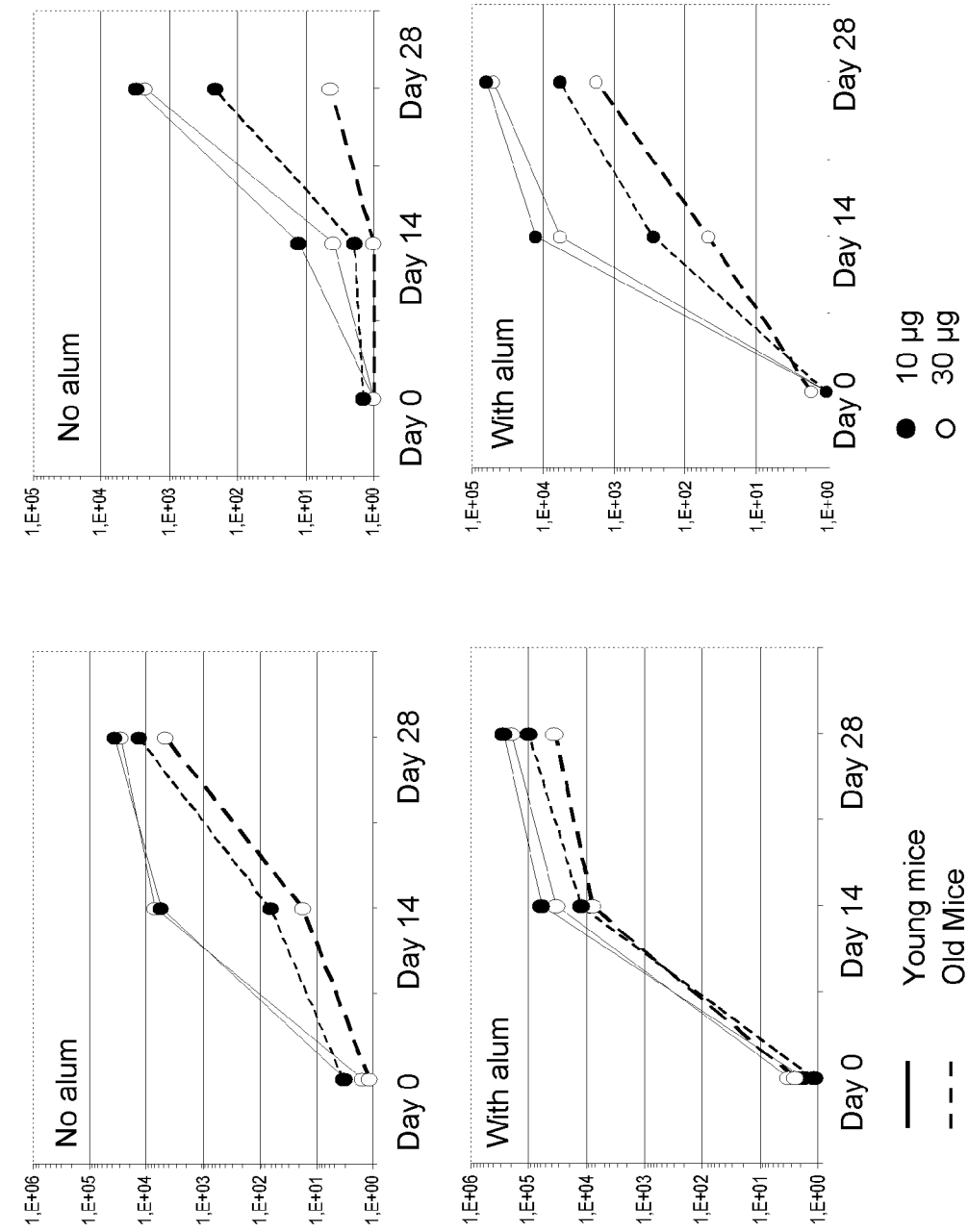

Figure 9

Comparison of the immunogenicity of C-TAB.G5.1 and *C. difficile* toxoid A and B

Antigen	Dose, µg	Study Day	Anti-C-TAB		Anti-Toxin A*		Anti-Toxin B*	
			No alum	Alum	No alum	Alum	No alum	Alum
C-TAB	10	14	4	30,271	1	21,238	1	10
		28	987	574,975	1,249	1,134,277	2	15,081
	30	14	1,274	64,228	1,391	31,391	2	3,836
		28	23,301	898,654	44,069	1,055,061	379	53,063
Toxoid	10	14	6,802	48,496	9,275	49,081	19	2,659
		28	20,140	978,277	72,954	1,128,721	380	40,295
	30	14	17,810	86,595	22,319	64,214	70	5,833
		28	65,536	731,416	348,019	1,292,560	4,967	43,098

* Immunization with toxoid induces antibody to the N-terminal portion of the toxin molecule which is read out in the anti-toxin ELISA, while immunization with C-TAB induces antibody to the C-terminal portion of the toxin molecule.

Figure 10

TNA and protection data in comparative C-TAB.G5.1 and toxoid A/B study

Antigen	Dose, µg	Alum	Toxin A challenge		Toxin B challenge	
			TNA	Protection, %	TNA	Protection, %
PBS	-	-	0	12.5	0	7.1
C-TAB	10	-	0	100	0	12.5
		+	532	100	0	100
	30	-	98	100	0	25
		+	735	100	487	100
Toxoid	10	-	362	100	0	0
		+	769	100	261	100
	30	-	1291	100	0	37.5
		+	1330	100	692	100

*Challenge dose:
toxin A: 28 ng/mouse
toxin B: 50 ng/mouse

Figure 11

A. **Anti-C-TAB IgG antibody response in hamsters**

Group	Vaccine	Pre*	Day 14	Day 28	Day 35
1	PBS	12	18	8	6
2	10 µg C-TAB.G5.1	44	320	16,105	15,100
3	10 µg C-TAB.G5.1 + alum	30	25,958	227,632	257,474
4	30 µg G5.1	35	745	20,377	56,476
5	30 µg C-TAB.G5.1 + alum	54	101,331	602,697	411,660
6	100 µg C-TAB. G5.1	19	2,508	31,978	42,131
7	100 µg C-TAB.G5.1 + alum	9	156,909	1,021,300	789,069

* pre-bleed

Figure 11

B. **Anti-toxin A IgG antibody response in hamsters**

Group	Vaccine	Pre*	Day 14	Day 28	Day 35
1	PBS	5	15	6	3
2	10 µg C-TAB.G5.1	5	117	2,988	3,233
3	10 µg C-TAB.G5.1 + alum	10	32,375	294,613	226,237
4	30 µg C-TAB.G5.1	6	486	6,872	16,336
5	30 µg C-TAB.G5.1 + alum	4	89,773	517,244	276,755
6	100 µg C-TAB.G5.1	10	224	2,053	2,675
7	100 µg C-TAB.G5.1 +alum	3	88,889	412,052	277,145

C. **Anti-toxin B IgG antibody response in hamsters**

Group	Vaccine	Pre*	Day 14	Day 28	Day 35
1	PBS	10	9	5	6
2	10 µg C-TAB.G5.1	25	271	9,627	16,104
3	10 µg C-TAB.G5.1 +alum	6	5,046	69,084	75,377
4	30 µg C-TAB.G5.1	18	375	14,621	47,642
5	30 µg C-TAB.G5.1 + alum	17	19,204	270,035	191,430
6	100 µg C-TAB.G5.1	8	263	5,029	32,928
7	100 µg C-TAB.G5.1 + alum	22	48,726	1,912,238	757,756

* pre-bleed

Figure 12

Kinetics of anti-C-TAB antibody development in hamsters

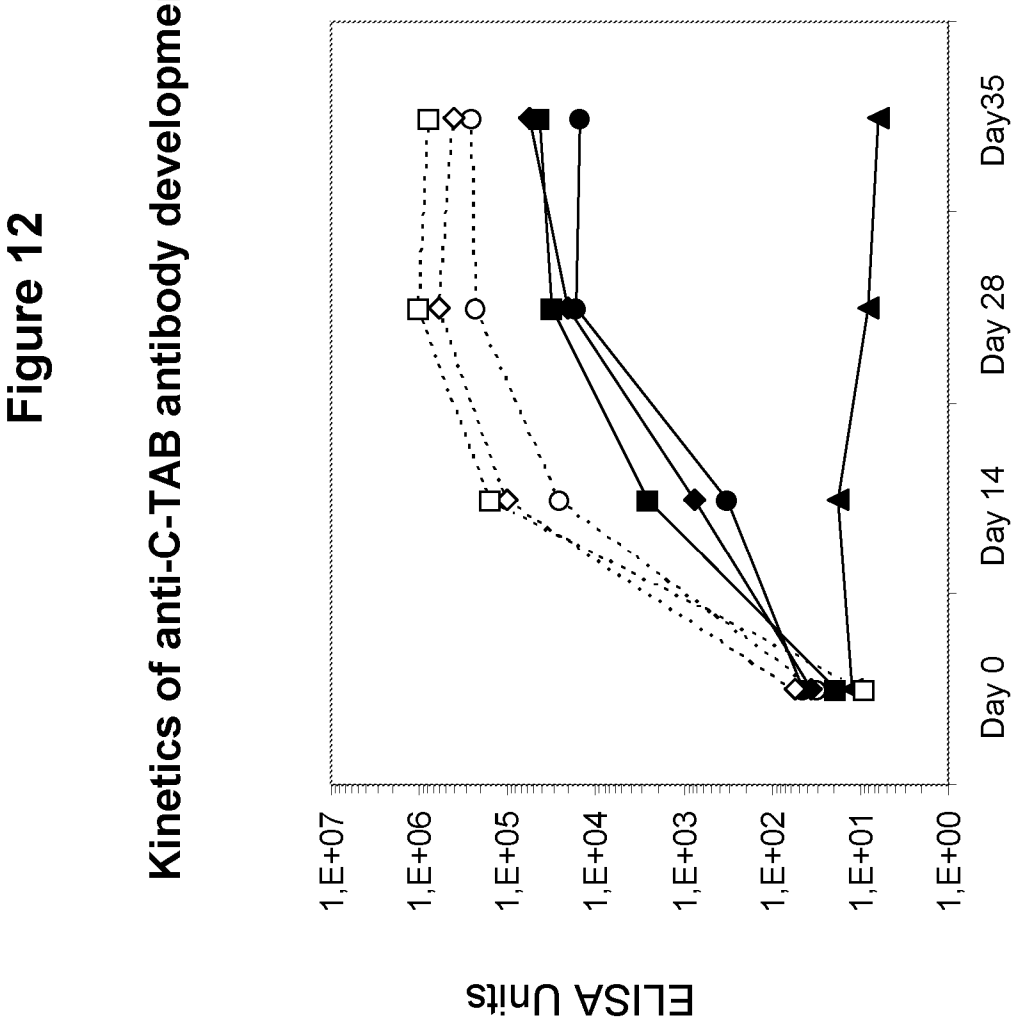

Figure 13

TNA and protection against *in vivo* toxin challenge in hamster

Vaccine	Toxin A challenge*		Toxin B challenge*	
	TNA	Protection, %	TNA	Protection, %
PBS	0	16.6	0	50
10 µg C-TAB.G5.1	267	100	0	83
30 µg C-TAB.G5.1	509	100	0	100
100 µg C-TAB.G5.1	179	83	0	50
10 µg C-TAB.G5.1 + alum	5,025	100	2,061	100
30 µg C-TAB.G5.1 + alum	2,824	100	4598	100
100 µg C-TAB.G5.1 + alum	3,262	100	16214	100

*Challenge dose:
toxin A: 75 ng/hamster
toxin B: 125 ng/hamster

Figure 14

Survival curves of *C. difficile* spore challenge study in hamsters

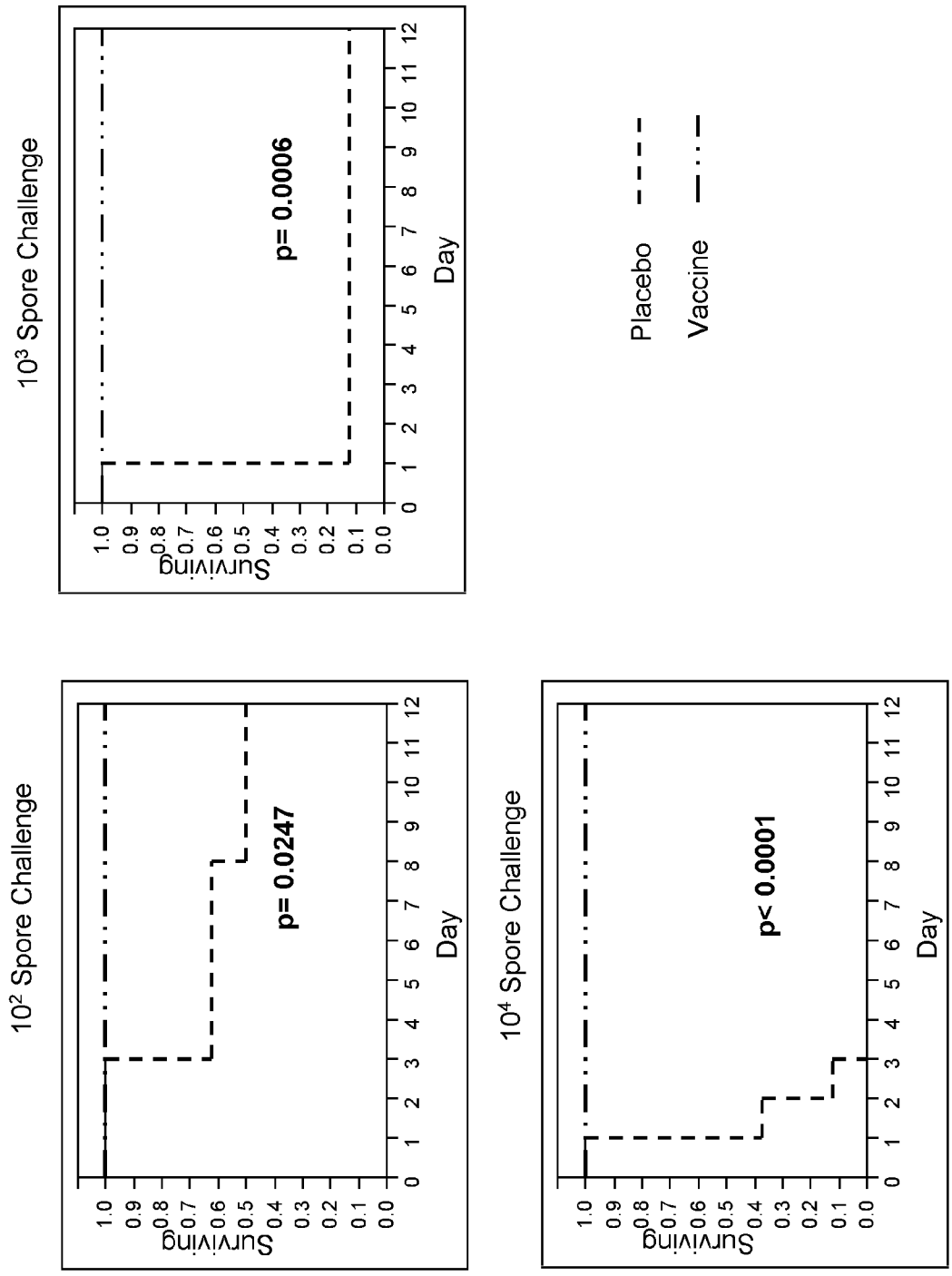

Figure 15

Anti-C-TAB antibody responses in monkeys

		+Alum		+Alum		+Alum
Study Day	**Anti-C-TAB**		**Anti-Toxin A**		**Anti-Toxin B**	
Pre-Bleed	330	486	1,222	1,751	566	744
Day 14	1,179	463,164	2,670	278,162	872	17,892
Day 28	31,518	2,688,109	23,726	1,128,621	10,104	179,050
Day 42	344,025	1,875,802	137,221	2,331,556	166,415	517,139

Figure 16

Comparison of C-TAB.G5 and C-TAB.G5.1 immunogenicity

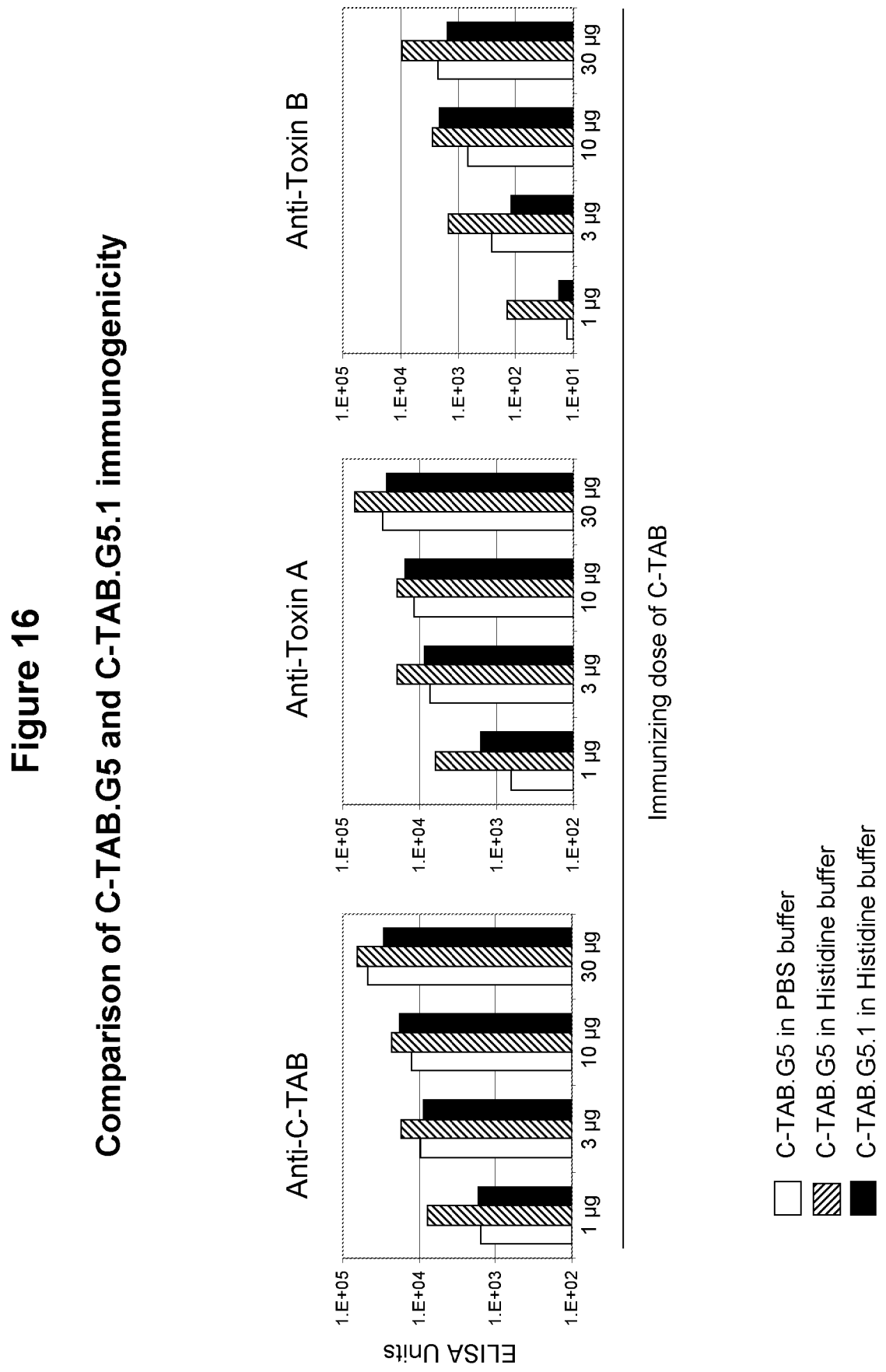

Figure 17

Immunogenicity of C-TAB.G5, C-TABNCTB and C-TADCTB in mice

Anti-C-TAB

Dose	C-TAB.G5	C-TANCTB	C-TADCTB
0.3 µg	2,397	2,196	2,062
1 µg	5,030	15,933	2,396
3 µg	37,542	20,699	27,717
10 µg	58,132	87,783	92,496
30 µg	95,179	278,713	278,534

Anti-Toxin A

Dose	C-TAB.G5	C-TANCTB	C-TADCTB
0.3 µg	2,127	1,235	1,464
1 µg	3,960	10,644	1,257
3 µg	28,701	13,178	13,590
10 µg	40,992	55,013	32,597
30 µg	64,063	152,425	163,932

Anti-Toxin B

Dose	C-TAB.G5	C-TANCTB	C-TADCTB
0.3 µg	31	23	266
1 µg	122	175	269
3 µg	2,170	1,284	4,241
10 µg	1,409	5,004	17,926
30 µg	4,616	8,327	138,098

Figure 18

Protection against challenge with toxin B* in mice

Vaccine§	C-TAB.G5		C-TABNCTB		C-TADCTB	
	Survival/total	Protection, %	Survival/total	Protection, %	Survival/total	Protection, %
0.33 µg	2/6	33	1/6	17	4/6	67
1 µg	4/6	67	2/6	33	2/6	33
3.3 µg	2/6	33	3/6	50	4/6	67
10 µg	3/6	50	2/6	33	5/6	83
33 µg	2/5	33	2/6	33	6/6	100

*Challenge dose: 50 ng/mouse

§ Negative control - vaccination with PBS: 2/6, 33 % protection

Figure 19

Comparison of TNA and protective efficacy of C-TAB.G5.1 and C-TADCTB in hamsters

Vaccine	Toxin A challenge*		Toxin B challenge*	
	TNA	Protection, %	TNA	Protection,%
PBS	0	16.6	0	50
10 µg C-TAB.G5.1	267	100	0	83
10 µg C-TAB.G5.1 + alum	5,025	100	2,061	100
30 µg C-TAB.G5.1	509	100	0	100
30 µg C-TAB.G5.1 + alum	2,824	100	4,598	100
100 µg C-TAB.G5.1	3,262	83	0	30
100 µg C-TAB.G5.1 + alum	3,262	100	16,214	100
30 µg C-TADCTB	159	100	2,886	100
30 µg C-TADCTB + alum	1,101	100	5,614	100

*Challenge dose:
toxin A - 75 ng/hamster
toxin B – 125 ng /hamster

Figure 20

TNA and protection against challenge with toxin A or B in mice immunized in different regimens

A.

C-TAB.5.1 immunization, days	Toxin A challenge§		Toxin B challenge§	
	TNA*	Protection, %	TNA**	Protection, %
PBS	0	25	0	0
0/3/14	51	75	0	12.5
0/3/14 + alum	742	100	399	28.6
0/7/21	71	100	9	12.5
0/7/21 + alum	1316	100	1318	37.5
0/14/28	107	100	24	37.5
0/14/28 + alum	1798	100	1275	87.5

§ Challenge dose:
toxin A: 28 ng/mouse
toxin B: 50 ng/mouse

* Geometric mean of one experiment, n=2
** Geometric mean of thee experiments, n=6

Figure 20

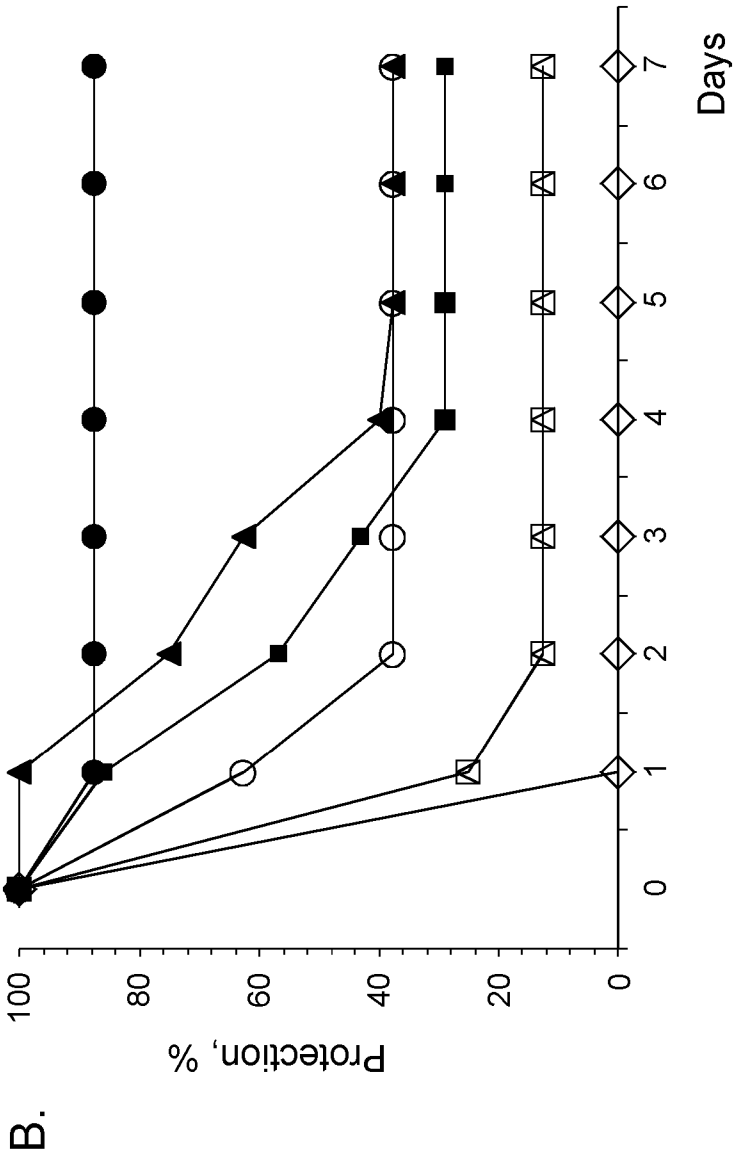
B.
Protection, %
100
80
60
40
20
0
0
1
2
3
4
5
6
7
Days
PBS
0/3/14
0/3/14+Alum
0/7/21
0/7/21/Alum
0/14/28
0/14/28+Alum

Figure 21

Protection against challenge with toxin A in mice received one dose of the vaccine

Immunization group	Toxin A challenge	
	Survival/total	Protection, %
Naive	3/12	25
Day 21	6/8	75
Day 35	6/8	75
Day 49	8/8	100

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 5736139 A **[0010]**
- WO 9702836 A, Thomas **[0013]**
- US 5919463 A **[0013]**
- WO 0061762 A, Wilkins **[0016]**
- US 6733760 B **[0016]**
- WO 2010017383 A, Telfer **[0018]**
- WO 9602555 A **[0050] [0106] [0141]**
- WO 0193903 A **[0050] [0106] [0141]**
- WO 0193905 A **[0050] [0106] [0141]**
- WO 02095027 A **[0050] [0106] [0141]**
- WO 0124822 A **[0050] [0106] [0141]**
- US 6291212 B **[0074]**
- US 5547871 A **[0074] [0081]**
- US 4683195 A, Mullis **[0092]**
- US 4683202 A, Mullis **[0092]**
- WO 0232451 A **[0106]**
- WO 0154720 A **[0106]**
- WO 0213857 A **[0106]**
- WO 03047602 A **[0106]**
- US 5730969 A **[0130]**
- EP 0399843 A **[0141]**
- US 61379892 B **[0242]**

Non-patent literature cited in the description

- **REVILL, P. ; SERRADELL, N. ; BOLOS, J.** Tiacumicin B: macrolide antibiotic treatment of C. difficile-associated diarrhea. *Drugs of the Future,* 2006, vol. 31 (6), 494-497 **[0006]**
- **JOHNSON ; STUART.** Recurrent Clostridium difficile infection: a review of risk factors, treatments, and outcomes. *Journal of Infection,* June 2009, vol. 58 (6), 403-410 **[0006]**
- **VOTH DE ; BALLARD JD.** *Clinical Microbiology Reviews,* 2005, vol. 18, 247-263 **[0007]**
- **HO et al.** *PNAS,* 2005, vol. 102, 18373-18378 **[0007]**
- **DOVE et al.** *Infect. Immun.,* 1990, vol. 58, 480-488 **[0007]**
- **BARROSO et al.** *Nucleic Acids Res.,* 1990, vol. 18, 4004 **[0007]**
- **EICHEL-STREIBER et al.** *Gene,* 1992, vol. 96, 107-113 **[0007]**
- **GIANNASCA PJ ; WARNY M.** *Vaccine,* 2004, vol. 22, 848-856 **[0008]**
- **KYNE L et al.** *The Lancet,* 2001, vol. 357, 189-193 **[0008]**
- **LYERLY et al.** *Current Microbiology,* 1990, vol. 21, 29-32 **[0009]**
- **RYAN et al.** *Infect. Immun.,* 1997, vol. 65, 2941-49 **[0009]**
- **KINK ; WILLIAMS.** *Infect. Immun.,* 1998, vol. 66, 2018-25 **[0010]**
- **BABCOCK et al.** *Infect. Immun.,* 2006, vol. 74, 6339-6347 **[0010]**
- **WARD et al.** *Infect. Immun.,* 1999, vol. 67, 5124-32 **[0011]**
- **DEMAREST SJ et al.** *J. Mol. Bio.,* 2005, vol. 346, 1197-1206 **[0012]**
- **DINGLE T.** *Glycobiology,* 2008, vol. 18, 698-706 **[0012]**
- **KOTLOFF KL et al.** *Infect. Immun.,* 2001, vol. 69, 988-995 **[0014]**
- **ABOUDOLA S et al.** *Infect. Immun.,* 2003, vol. 71, 1608-1610 **[0014]**
- **VARFOLOMEEVA et al.** *Mol. Genetics, Microb. and Virol.,* 2003, vol. 3, 6-10 **[0015]**
- **BELYI ; VARFOLOMEEVA.** *FEMS Letters,* 2003, vol. 225, 325-9 **[0015]**
- **DOVE et al.** *Infect. Immun.,* 1990, vol. 58, 480-8 **[0039]**
- **BARROSO et al.** *Nucleic Acids Research,* 1990, vol. 18, 4004 **[0039]**
- **CRYZ et al.** *Vaccine,* 1994, vol. 13, 67-71 **[0043]**
- **LIANG et al.** *J. Immunology,* 1988, vol. 141, 1495-501 **[0043]**
- **CZERKINSKY et al.** *Infect. Immun.,* 1989, vol. 57, 1072-77 **[0043]**
- **SKEIKY et al.** *J. Exp. Med.,* 1995, vol. 181, 1527-1537 **[0050]**
- **DICKENSON ; CLEMENTS.** *Infection and Immunity,* 1995, vol. 63 (5), 1617-1623 **[0050]**
- Computational Molecular Biology. Oxford University Press, 1988 **[0067]**
- Biocomputing: Informatics and Genome Projects. Academic Press, 1993 **[0067]**
- Computer Analysis of Sequence Data. Humana Press, 1994 **[0067]**
- **VON HEINJE, G.** Sequence Analysis in Molecular Biology. Academic Press, 1987 **[0067]**
- Sequence Analysis Primer. M Stockton Press, 1991 **[0067]**

- Sequence Analysis in Molecular Biology. Academic Press, 1987 **[0067]**
- **CARILLO, H. ; LIPMAN, D.** *SIAM J. Applied Math.,* 1988, vol. 48, 1073 **[0067]**
- Guide to Huge Computers. Academic Press, 1994 **[0067]**
- **MERRIFIELD.** *J. Am Chem. Soc.,* vol. 85, 2149-2154 **[0071]**
- **KENT et al.** Synthetic Peptides in Biology and Medicine. Elsevier Science Publishers, 1985, 295-358 **[0071]**
- **GRANTHAM, R. et al.** *Nucl. Acid Res.,* 1980, vol. 8, 49-62 **[0081]**
- **GRANTHAM, R. et al.** *Nucl. Acid Res.,* 1981, vol. 9, 43-74 **[0081]**
- **MAROYAMA, T. et al.** *Nucl. Acid Res.,* 1986, vol. 14, 151-197 **[0081]**
- **AOTA, S. et al.** *Nucl. Acid Res.,* 1988, vol. 16, 315-402 **[0081]**
- **WADA, K. et al.** *Nucl. Acid Res.,* 1991, vol. 19, 1981-1985 **[0081]**
- **KURLAND, C. G.** *FEBS Lett.,* 1991, vol. 285, 165-169 **[0081]**
- **PEDERSEN, S.** *EMBO J.,* 1984, vol. 3, 2895-2898 **[0081]**
- **SORENSEN, M. A.** *J. Mol. Biol.,* 1989, vol. 207, 365-377 **[0081]**
- **RANDALL, L. L. et al.** *Eur. J. Biochem.,* 1980, vol. 107, 375-379 **[0081]**
- **CURRAN, J. F. ; YARUS, M.** *J. Mol. Biol.,* 1989, vol. 209, 65-77 **[0081]**
- **VARENNE, S. et al.** *J. Mol. Biol.,* 1984, vol. 180, 549-576 **[0081]**
- **VARENNE, S. et al.** *J. Mol, Biol.,* 1984, vol. 180, 549-576 **[0081]**
- **GAREL, J.-P.** *J. Theor. Biol.,* 1974, vol. 43, 211-225 **[0081]**
- **IKEMURA, T.** *J. Mol. Biol.,* 1981, vol. 146, 1-21 **[0081]**
- **IKEMURA, T.** *J. Mol. Biol.,* 1981, vol. 151, 389-409 **[0081]**
- **DEVEREUX et al.** *Nucl. Acid Res.,* 1984, vol. 12 (1), 387 **[0083]**
- **ATSCHUL et al.** *J. Mol. Biol.,* 1990, vol. 215, 403 **[0083]**
- **NEEDLEMAN ; WUNSCH.** *J. Mol. Biol.,* 1970, vol. 48, 443-453 **[0083]**
- **KRUSE et al.** Tissue Culture. Academic Press, 1973 **[0090]**
- Ullmann's Encyclopedia of Industrial Chemistry. 2003 **[0121]**
- Remington's Pharmaceutical Sciences. Mack Publishing, 2005 **[0121]**
- Pharmaceutical Dosage Forms. Marcel Dekker, 1989 **[0121]**
- **ANSEL et al.** Pharmaceutical Dosage Forms and Drug Delivery Systems. Williams & Wilkins, 2005 **[0121]**
- **MEDZHITOV ; JANEWAY.** *Curr. Opin. Immunol.,* 1997, vol. 9 (1), 4-9 **[0143]**

C. difficile toxin A és toxin B proteinek izolált polipeptidei és alkalmazásaik

Szabadalmi igénypontok

1. Immunogén vagy vakcina készítmény, mely a 4. számú szekvenciának megfelelő aminosavszekvenciával legalább 85 %-os szekvencia azonossággal rendelkező izolált polipeptidet és egy gyógyászatilag elfogadható hordozót vagy kötőanyagot tartalmaz.

2. Az 1. igénypont szerinti készítmény, ahol a polipeptid 100 %-os túlélést biztosít a polipeptiddel vakcinált hörcsögökben *Clostridium difficile* spóra 10^2, 10^3 vagy 10^4 dózisban történő intragasztrális adagolását követően.

3. Az 1. vagy 2. igénypont szerinti készítmény, ahol a polipeptid a *Clostridium difficile* toxin A C-terminális doménjéből származó 19 ismétlődő egységet tartalmaz.

4. Az 1-3. igénypontok bármelyike szerinti készítmény, ahol a polipeptid a *Clostridium difficile* toxin B C-terminális doménjéből származó 23 ismétlődő egységet tartalmaz.

5. Az 1-4. igénypontok bármelyike szerinti készítmény, ahol a polipeptid legalább 90%-os, előnyösebben 95%-os, legelőnyösebben 99%-os szekvencia azonosságot mutat a 4. számú szekvenciának megfelelő aminosavszekvenciával.

6. Az 1-5. igénypontok bármelyike szerinti készítmény, ahol a polipeptid a 2. számú szekvenciával vagy 4. számú szekvenciával rendelkezik.

7. Az 1-6. igénypontok bármelyike szerinti készítmény, mely továbbá egy adjuvánst is tartalmaz.

8. A 7. igénypont szerinti készítmény, ahol az adjuváns timsó.

9. Egy 4. számú szekvenciának megfelelő aminosavszekvenciával legalább 85 %-os szekvencia azonossággal rendelkező izolált polipeptid vagy az 1-8. igénypontok bármelyike szerinti

készítmény *C. difficile*-vel összefüggő betegség (CDAD) megelőzésében vagy kezelésében történő alkalmazásra.

10. A polipeptid vagy készítmény a 9. igénypont szerinti alkalmazásra, CDAD veszélyének kitett alanyban a CDAD megelőzésében történő alkalmazásra.

11. A polipeptid vagy készítmény a 10. igénypont szerinti alkalmazásra, ahol a CDAD veszélyének kitett alany: i) idős alany; ii) AIDS-es alany; iii) immunszupresszánst kapó vagy annak szedésére tervezett alany; iv) kórházi kezelésre tervezett vagy kórházban levő alany; v) intenzív osztályon levő vagy intenzív osztályon elhelyezendő alany; vi) gyomor- és bélrendszeri sebészeti beavatkozás alatt álló vagy beavatkozásra tervezett alany; vii) szanatóriumban gondozott vagy hosszútávú szanatóriumi gondozásra tervezett alany; viii) gyakori és/vagy folyamatos antibiotikumos kezelésre szoruló társbetegséggel rendelkező alany; ix) visszaeső CDAD-vel rendelkező alany; x) 18 év alatti alany; xi) 18-65 év közötti korú alany; vagy xii) 65 évesnél idősebb alany.

12. A polipeptid vagy készítmény a 10. igénypont szerinti alkalmazásra, ahol a CDAD veszélyének kitett alany 65 évesnél idősebb.

13. A polipeptid vagy készítmény a 9. igénypont szerinti alkalmazásra, ahol a CDAD kezelése a betegség súlyosságának és/vagy időtartamának csökkentését, a betegség tüneteinek javítását vagy ezek kombinációit jelenti.

14. Az 1-13. igénypontok bármelyike szerinti készítmény, mely egy további antigént tartalmaz.

15. Az 1-14. igénypontok bármelyike szerinti készítmény, mely továbbá egy érzéstelenítőszerek, fájdalomcsillapítók, gyulladásgátlók, szteroidok, antibiotikumok, antiarthritikumok, anorektikumok, antihisztaminok és antineoplasztikumok közül választott terápiás szert tartalmaz.