
Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published: — without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guide. ance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Title: COMPOSITIONS AND METHODS FOR AMELIORATING TOXINS AND VIRULENCE FACTORS

Abstract: The invention provides methods of ameliorating or preventing toxicity associated with a virulence factor or toxin (e.g., a microbial virulence factor, e.g., from Clostridium) in a subject, comprising administering to the subject in need thereof, a therapeutically effective amount of polypeptide having a protease activity, e.g., an isolated, recombinant or synthetic protease or antibody having protease activity, wherein the polypeptide (e.g., protease) partially or completely neutralizes the virulence factor or a toxin, thereby ameliorating or preventing the toxicity associated with the virulence factor or toxin.
COMPOSITIONS AND METHODS FOR AMELIORATING
TOXINS AND VIRULENCE FACTORS

FIELD OF THE INVENTION

This invention relates to the fields of enzymology and medicine. The invention provides compositions and methods for ameliorating or preventing toxicity associated with a virulence factor or a toxin, e.g., a microbial virulence factor or toxin, such as Clostridium toxin, in a subject.

BACKGROUND

Clostridium difficile is an anaerobic bacillus that infects between 10 to 25 percent of hospitalized patients who undergo antibiotic treatment. The infection can be asymptomatic or can lead to chronic diarrhea, colitis and life-threatening pseudomembranous colitis (PMC). Antibiotic-induced clearance of the normal intestinal flora paves the way for C. difficile infection. Patients with severe risk factors tend to be the elderly and immunocompromised individuals who suffer from comorbidity factors.

Very little is known concerning the structural or functional aspects of the repeat domains of toxins A and B. The CWB-domain of LytA from Streptococcus pneumoniae (Figure 1a) is known to form a unique β-solenoid structure comprised of six repeats. The CWB-domains of toxin A and B have the added complexity of a 30 residue, non-standard repeat sequence which occurs approximately once every six classical repeats. The structural consequence of these 30 residue inserts is unknown. Additionally, the ability of the toxin CWB-domains to associate with mammalian cell surfaces via glycoprotein recognition or other means has often been implied from known functional aspects of related, non-toxin protein domains derived from bacteria other than C. difficile. While recombinantly derived C. difficile toxin CWB-domains have been reported for over ten years, the majority of the published data focuses on their use as vaccines, as immunogens for raising neutralizing antibodies, or even as potential blocking reagents against the full-length toxins.

SUMMARY

The invention provides compositions and methods of ameliorating or preventing toxicity associated with a virulence factor or toxin (e.g., a microbial virulence factor or toxin) in a subject, comprising administering to the subject in need thereof, a
therapeutically effective amount of polypeptide having a protease activity, e.g., an isolated, recombinant or synthetic protease, wherein the polypeptide (e.g., protease or antibody having protease activity) partially or completely neutralizes the virulence factor or toxin (e.g., a bacterial toxin, such as *Clostridium* toxin), thereby ameliorating or preventing the toxicity associated with the virulence factor or toxin. In one aspect, ameliorating or preventing toxicity associated with a virulence factor or toxin comprises ameliorating or preventing symptoms or conditions associated with the virulence factor or toxin, e.g., gastrointestinal problems, such as ulcers, diarrhea, pseudomembranous colitis or a cancer.

In one aspect, the virulence factor comprises a microbial, e.g., a bacterial, protozoan, fungal or viral virulence factor. In one aspect, the virulence factor comprises an adherence factor, a coat protein, an invasion factor, a capsule, or a toxin. The toxin can be an exotoxin or an endotoxin, e.g., a microbial exotoxin or an endotoxin, such as a bacterial exotoxin or an endotoxin, e.g., from *Clostridium*. In one aspect, the toxin is produced by a *Clostridium* bacteria, e.g., comprises a *Clostridium difficile* toxin A. In one aspect, the polypeptide having a protease activity cleaves a *C. difficile* toxin A within residues 1812-2710 of the toxin. In one aspect, the polypeptide having a protease activity cleaves a *C. difficile* toxin B within residues 1814-2366 of the toxin.

In one aspect, the protease cleaves a site within or flanking the cell wall binding domain of the virulence factor, or toxin, or cleaves a site that effects the function of the cell wall binding domain, e.g., abrogates or diminishes the ability of the domain to bind a cell wall. In one aspect, the cell wall binding domain comprises *Clostridial* repetitive oligopeptides or a P-motif. The cell wall binding domain polypeptide can be ToxA:2459-2710.

In one aspect, the protease comprises a polypeptide having at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to a *C. difficile* CWB-domain of the invention, e.g., residues 2459-2710 of toxin A, and as described below.

In one aspect, the protease is stable in acidic and/or alkaline conditions equivalent to the subjects gut, e.g., in the stomach or intestine. In one aspect, the protease is stable in cecal fluid, or equivalent, e.g., cecal fluid supernatant and/or homogenized cecal content *in vitro*, for example, is stable for about 30 minutes at 37°C. In one aspect, a protease is modified to improve its stability various acidic and/or alkaline conditions, or to improve
its stability in the presence of cecal fluid supernatant or homogenized cecal content *in vitro*. The protease can also be modified to provide, or increase, its resistance to proteolysis, e.g., to other enzymes, such as pepsin, pancreatin, trypsin, trypsinogen, chymo-trypsinogen, carboxy-peptidase, pro-carboxy-peptidase, elastase and/or pro-elastase. The modification can be an amino acid sequence modification, a chemical modification of a natural amino acid (e.g., phosphorylation, pegylation, carboxylation, attachments of carbohydrates and the like), polymerization, use of a synthetic amino acid, e.g., a peptidomimetic, and the like.

In one aspect, the administered isolated, recombinant or synthetic protease has the activity of a pepsin, pancreatin, trypsin, trypsinogen, chymo-trypsinogen, carboxy-peptidase, pro-carboxy-peptidase, elastase, pro-elastase, or a combination thereof, e.g., the proteolysis can be mediated by pepsin, pancreatin, trypsin, trypsinogen, chymo-trypsinogen, carboxy-peptidase, pro-carboxy-peptidase, elastase, pro-elastase, or a combination thereof.

In one aspect, the protease is modified to have an chemical stability superior to the unmodified protease; a thermotolerance or thermostability superior to the unmodified protease; a pH tolerance superior to the unmodified protease; a reduced immunogenicity; a stability in pharmaceutical formulation superior to that of the unmodified protease; an increased half-life relative to the unmodified protease; a stable or increased recombinant expression in a host cell; or a combination thereof. The modification can be an amino acid sequence modification, a chemical modification of a natural amino acid (e.g., phosphorylation, pegylation, carboxylation, attachments of carbohydrates and the like), polymerization, use of a synthetic amino acid, e.g., a peptidomimetic (and the like) or a combination thereof.

The polypeptide having protease activity can comprise in whole or in part an isolated polypeptide, a synthetic equivalent to a polypeptide, e.g., a peptidomimetic, or a recombinant polypeptide. In one aspect, the recombinant polypeptide is recombinantly expressed in a mammalian, yeast, bacterial, insect, fungal or plant cell.

In one aspect, the sequence of the polypeptide having protease activity is modified by an addition, deletion, or substitution of the amino acid sequence of the polypeptide (e.g., protease or antibody having enzymatic activity). The polypeptide having protease activity can be modified by one or more additions, deletions, or substitutions to a nucleic acid sequence encoding the polypeptide (e.g., protease or antibody having enzymatic activity). The additions, deletions, or substitutions to the nucleic acid sequence encoding
the polypeptide (e.g., protease or antibody having enzymatic activity) can be introduced by a method comprising error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, Gene Site Saturation Mutagenesis™ (GSSM™), synthetic ligation reassembly (SLR), or a combination thereof. The additions, deletions, or substitutions to the nucleic acid sequence encoding the polypeptide (e.g., protease or antibody having enzymatic activity) can be introduced by a method comprising recombination, recursive sequence recombination, phosphorothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation, or a combination thereof.

In one aspect, the polypeptide having protease activity (e.g., protease or antibody having enzymatic activity) is formulated for oral delivery. The polypeptide having protease activity can be formulated as an enteral formulation. The polypeptide having protease activity can be formulated in any way, e.g., as a liquid, an emulsion, a suspension, a pill, a tablet, a capsule, a spray, a powder, a lozenge, a troche, in a liposome, a gel, a microsphere, a multiparticulate core particle, and the like.

In one aspect, the method further comprises administering more than one polypeptide having protease activity, or, one or more additional anti-virulence factor or anti-toxin agents. The additional polypeptides having protease activity or anti-virulence factor or anti-toxin agents can be administered simultaneously or sequentially with the polypeptide having protease activity. For example, the additional anti-virulence factor agent or anti-toxin can be one or more monoclonal or polyclonal antibodies, a small molecule having anti-virulence factor agent or anti-toxin activity, a drug (e.g., having anti-microbial activity).

In one aspect, the method further comprises co-treating or pre-treating the subject to lower the pH of the intestine or increase the pH of the stomach. The method can also further comprise co-treating or pre-treating the subject with at least one antimicrobial agent, e.g., a drug, e.g., an antibiotic or other antimicrobial agent, such as vancomycin, bacitracin or metronidazole, an antibody, and the like.
In one aspect, the toxicity ameliorated, i.e., the toxicity associated with the virulence factor or toxin, comprises gastrointestinal problems, e.g., a gastrointestinal disease or disorder. The gastrointestinal disease or disorder ameliorated or otherwise treated by the compositions and methods of the invention can include an ulcer, diarrhea, pseudomembranous colitis or cancer.

The pharmaceutical compositions of the invention can be formulated for enteric delivery, e.g., comprising an enteric coating. The pharmaceutical composition can be formulated as a suspension, an emulsion, a liquid, a capsule, a tablet, a lozenge, a troche, a powder, a gel, a microsphere, a liposome, a multiparticulate core particle, a spray and the like. In one aspect, the pharmaceutical composition comprises an antibody or a protease or a combination thereof, e.g., comprising from about 50% to about 95% of the batch size (weight/weight) of the composition.

The invention provides kits comprising the pharmaceutical compositions of the invention, and in one aspect further comprises instructions for administering the pharmaceutical composition.

The invention provides methods of identifying a polypeptide, e.g., a protease or antibody, useful in the amelioration or prevention of virulence factor-associated toxicity comprising contacting a virulence factor with a test polypeptide, e.g., a protease; assessing the proteolytic cleavage of the virulence factor by the polypeptide; and identifying the polypeptide that cleaves or otherwise modifies the virulence factor as the polypeptide (e.g., protease) that is useful in the amelioration or prevention of virulence factor-associated toxicity. In one aspect, the method comprises incubating the virulence factor and test polypeptide (after the contacting step), e.g., in an ELISA binding plate. In one aspect, the method comprises probing the proteins bound to the ELISA plate with an antibody that binds the virulence factor; and determining the amount of bound antibody, whereby the absence of bound antibody indicates that the polypeptide, e.g., protease or antibody, cleaved bound to or otherwise modified the virulence factor. In one aspect, the virulence factor comprises an adherence factor, a coat protein, an invasion factor, a capsule, or a toxin, e.g., a Clostridium difficile toxin, such as a Clostridium difficile toxin A. In one aspect, the contacting step is performed in conditions resembling (e.g., the same as or equivalent to) those present in the digestive tract, e.g., stomach or intestine. For example, in one aspect, the step is performed in the presence of cecal content homogenate.
In one aspect, the method further comprises a step for selecting a polypeptide having virulence factor-ameliorating activity (e.g., protease or antibody activity) that: retains greater than about 75%, 80%, 85% or 90% activity (e.g., proteolytic activity) after about 15, 20, 25 or 30 or more minutes in a cecal fluid supernatant (or equivalent) at about equivalent body temperature (e.g., 37°C); retains greater than about 40% virulence factor-ameliorating activity (e.g., protease or antibody activity) after about 30 minutes in cecal content homogenate (or equivalent) at about equivalent body temperature (e.g., 37°C); exhibits sufficient activity to degrade greater than about 75%, 80%, 85% or 90% of the toxin in less than about 15, 20, 25 or 30 minutes; increases the percentage of cells surviving after co-incubation with toxin to greater than about 50%, or more; or, partially or completely neutralizes the toxin-induced toxicity in the rat ileal loop model, or equivalent (see, e.g., U.S. Patent No. 5,773,000; Charney (1976) J. Clin. Invest. 57(6): 1590-1599). In one aspect, the neutralizing activity is similar to or greater than the neutralizing activity of cholestyramine.

In one aspect, the method further comprises a step for modifying the identified protease to have: an increased stability in the presence of cecal fluid supernatant or homogenized cecal content in vitro; an chemical stability superior to the unmodified protease; a thermostolerance or thermostability superior to the unmodified protease; a pH tolerance superior to the unmodified protease; a reduced immunogenicity; a stability in pharmaceutical formulation superior to that of the unmodified protease; an increased half-life relative to the unmodified protease; a stable or increased recombinant expression in a host cell; or a combination thereof.

In one aspect, the method further comprises selecting a protease that has: an increased stability in the presence of cecal fluid supernatant or homogenized cecal content in vitro; an chemical stability superior to the unmodified protease; a thermostolerance or thermostability superior to the unmodified protease; a pH tolerance superior to the unmodified protease; a reduced immunogenicity; a stability in pharmaceutical formulation superior to that of the unmodified protease; an increased half-life relative to the unmodified protease; a stable or increased recombinant expression in a host cell; or a combination thereof. In one aspect, the modification is an addition, deletion, or substitution in the amino acid sequence of the protease. The modification of protease can be introduced by one or more additions, deletions, or substitutions to the nucleic acid sequence encoding the protease polypeptide. The additions, deletions, or substitutions to the nucleic acid sequence encoding the protease can be introduced by a method
comprising error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly
PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive
ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis,
gene reassembly, Gene Site Saturation Mutagenesis™ (GSSM™), synthetic ligation
reassembly (SLR), or a combination thereof. The additions, deletions, or substitutions to
the nucleic acid sequence encoding the protease can be introduced by a method
comprising recombination, recursive sequence recombination, phosphothioate-modified
DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis,
point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical
mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection
mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble
mutagenesis, chimeric nucleic acid multimer creation, or a combination thereof.

In one aspect, the identified protease cleaves *C. difficile* toxin A within residues
1812-2710 of the toxin. In one aspect, the identified protease cleaves *C. difficile* toxin B
within residues 1814-2366 of the toxin.

The invention provides methods for identifying a compound that inhibits the
binding of a *Clostridium sp.* toxin with a mammalian (e.g., a human) cell, comprising
contacting a mammalian cell with a recombinant, functional cell wall binding domain
polypeptide from the *Clostridium sp.* toxin and a test compound in the presence of Ca$^{2+}$;
determining the amount of cell binding by the cell wall binding domain polypeptide; and
identifying the test compound as an inhibitory compound if it reduces or eliminates
binding of the cell wall binding domain polypeptide to the cell. The invention provides
methods for identifying an inhibitory compound that inhibits the binding of a *Clostridium
sp.* toxin with a mammalian (e.g., a human) cell, comprising contacting a mammalian cell
with a *Clostridium sp.* toxin and a test compound in the presence of Ca$^{2+}$; determining the
amount of cell binding by the cell wall binding domain polypeptide; and identifying the
test compound as an inhibitory compound if it reduces or eliminates binding of the cell
wall binding domain polypeptide to the cell. In one aspect if these identifying methods,
the toxin comprises a *C. difficile* toxin. The cell wall binding domain polypeptide can be
ToxA:2459-2710. The compound can reduces or eliminates binding of the full length
toxin to the cell. In one aspect, the contacting occurs in the presence of ImM CaCl$_2$. The
compound can reduce the toxicity of the full length toxin to mammalian cells.

The invention provides methods for ameliorating or preventing toxicity associated
with *Clostridium sp.* toxin in a subject, comprising administering to a subject in need
thereof, a therapeutically effective amount of a compound of the invention, e.g., a
identified by a method of the invention.

The invention provides methods for ameliorating or preventing toxicity associated
with *Clostridium difficile* toxin A in a subject, comprising administering to a subject in
need thereof, a therapeutically effective amount of a compound which disrupts the Ca$^{2+}$-
dependent cell surface association of toxin A cell wall binding domain with the surface of
a mammalian cell.

The invention provides novel structural and functional *C. difficile* CWB-domains,
e.g., residues 2459-2710 of toxin A, and as described below, and compositions and
methods using these novel *C. difficile* CWB-domains.

The details of one or more aspects of the invention are set forth in the accompa¬
nying drawings and the description below. Other features, objects, and advantages of the
invention will be apparent from the description and drawings, and from the claims.

All publications, patents, patent applications, GenBank sequences and ATCC
deposits, cited herein are hereby expressly incorporated by reference for all purposes.

**BRIEF DESCRIPTION OF DRAWINGS**

The following drawings are illustrative of aspects of the invention and are not
meant to limit the scope of the invention as encompassed by the claims.

The patent or application file contains at least one drawing executed in color.
Copies of this patent or patent application publication with color drawing(s) will be
provided by the Office upon request and payment of the necessary fee.

Figure 1 illustrates an alignment of the consensus repeat sequences of CWB-
domains from *Streptococcus mutans* GtfC and GtfB, *Streptococcus downei* GtfI,
*Streptococcus pneumoniae* LytA, pneumococcal phage Cpl-I, *C. difficile* toxins A and
B and *C. Sordelii* Toxin L, as described in detail in Example 1, below.

Figure 2 illustrates a far UV CD spectra of toxin CWB-domains, as described in
detail in Example 1, below.

Figure 3 illustrates data showing Ca$^{2+}$-binding to ToxA:llR and Ca$^{2+}$-accelerated
CHO cell killing by full length toxin A, as described in detail in Example 1, below.

Figures 4a-f illustrate data showing calcium dependent binding of ToxA:llR to
CHO cell surfaces determined by flow cytometry, as described in detail in Example 1,
below.
Like reference symbols in the various drawings indicate like elements.

DETAILED DESCRIPTION

The invention provides compositions and methods for ameliorating or preventing toxicity associated with a virulence factor or a toxin in a subject. The methods can comprise administering to the subject in need thereof a therapeutically effective amount of an isolated, synthetic or recombinant polypeptide having protease activity, wherein the polypeptide having protease activity partially or completely neutralizes the virulence factor or toxin, thereby ameliorating or preventing the toxicity associated with the virulence factor or toxin.

In aspect, the invention provides novel structural and functional C. difficile CWB-domains. The full length CWB-domains of toxins A and B and a number of smaller constructs (Figure 1b) were cloned, expressed and characterized. A difference in the minimal sequence requirements for forming tertiary structure within the CROP-domains compared to other bacterial CWB-domains was found, and the invention for the first time demonstrates a dependence on Ca\(^{2+}\) for mammalian cell surface recognition of the C. difficile CWB-domains.

The polypeptides and peptides of the invention, and polypeptides and peptides used in the compositions and methods of the invention, can comprise amino acids joined to each other by peptide bonds or modified peptide bonds, e.g., peptide isosteres and may contain modified amino acids other than the 20 gene-encoded amino acids. These polypeptides can also include polypeptides modified by either natural processes, such as post-translational processing, or by chemical modification techniques. Modifications can occur anywhere in the polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also a given polypeptide may have many types of modifications. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of a phosphatidylinositol, cross-linking cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation,

The polypeptides and peptides of the invention, and polypeptides and peptides used in the compositions and methods of the invention, can be glycosylated. The glycosylation can be added post-translationally either chemically or by cellular biosynthetic mechanisms, wherein the later incorporates the use of known glycosylation motifs, which can be native to the sequence or can be added as a peptide or added in the nucleic acid coding sequence. The glycosylation can be O-linked or N-linked.

The polypeptides and peptides of the invention, and polypeptides and peptides used in the compositions and methods of the invention, can comprise "mimetic" and "peptidomimetic" forms. The terms "mimetic" and "peptidomimetic" refer to a synthetic chemical compound which has substantially the same structural and/or functional characteristics of the polypeptides of the invention. The mimetic can be either entirely composed of synthetic, non-natural analogues of amino acids, or, is a chimeric molecule of partly natural peptide amino acids and partly non-natural analogs of amino acids. The mimetic can also incorporate any amount of natural amino acid conservative substitutions as long as such substitutions also do not substantially alter the mimetic’s structure and/or activity. As with polypeptides of the invention which are conservative variants or members of a genus of polypeptides of the invention (e.g., having about 50% or more sequence identity to an exemplary sequence of the invention), routine experimentation will determine whether a mimetic is within the scope of the invention, i.e., that its structure and/or function is not substantially altered. Thus, in one aspect, a mimetic composition is within the scope of the invention if it has activity equivalent to a novel C. difficile CWB-domain of the invention.

Polypeptide mimetic compositions of the invention can contain any combination of non-natural structural components. In alternative aspect, mimetic compositions of the invention include one or all of the following three structural groups: a) residue linkage groups other than the natural amide bond ("peptide bond") linkages; b) non-natural residues in place of naturally occurring amino acid residues; or c) residues which induce
secondary structural mimicry, i.e., to induce or stabilize a secondary structure, e.g., a beta turn, gamma turn, beta sheet, alpha helix conformation, and the like. For example, a polypeptide of the invention can be characterized as a mimic when all or some of its residues are joined by chemical means other than natural peptide bonds. Individual peptidomimetic residues can be joined by peptide bonds, other chemical bonds or coupling means, such as, e.g., glutaraldehyde, N-hydroxysuccinimide esters, bifunctional maleimides, N,N'-dicyclohexylcarbodiimide (DCC) or N,N'-diisopropylcarbodiimide (DIC). Linking groups that can be an alternative to the traditional amide bond ("peptide bond") linkages include, e.g., ketomethylene (e.g., -C(=O)-CH₂- for -C(=O)-NH-), aminomethylene (CH₂-NH), ethylene, olefin (CH=CH), ether (CH₂-O), thioether (CH₂-S), tetrazole (CN₄⁻), thiazole, retroamide, thioamide, or ester (see, e.g., Spatola (1983) in Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, pp 267-357, "Peptide Backbone Modifications," Marcell Dekker, NY).

A polypeptide or peptide can also be characterized as a mimic by containing all or some non-natural residues in place of naturally occurring amino acid residues. Non-natural residues are well described in the scientific and patent literature; a few exemplary non-natural compositions useful as mimetics of natural amino acid residues and guidelines are described below. Mimetics of aromatic amino acids can be generated by replacing by, e.g., D- or L- naphylalanine; D- or L- phenylglycine; D- or L-2 thienylalanine; D- or L- 1, -2, -3, or 4- pyreneylalanine; D- or L-3 thieneylealanine; D- or L- (2-pyridinnyl)-alanine; D- or L- (3-pyridynyl)-alanine; D- or L- (2-pyraznyl)-alanine; D- or L- (4-isopropyl)-phenylglycine; D- (trifluoromethyl)-phenylglycine; D- (trifluoromethyl)-phenylalanine; D- or L-p-biphenylphenylalanine; D- or L-p-methoxy-biphenylphenylalanine; D- or L-2-indole(alkyl)alanines; and, D- or L-alkylarines, where alkyl can be substituted or unsubstituted methyl, ethyl, propyl, hexyl, butyl, penty, isopropyl, iso-butyl, sec-isotyl, iso-pentyl, or a non-acidic amino acids. Aromatic rings of a non-natural amino acid include, e.g., thiazoyl, thiophenyl, pyrazoyl, benzimidazoyl, naphthyl, furanyl, pyrrolyl, and pyridyl aromatic rings.

Mimetics of acidic amino acids can be generated by substitution by, e.g., non-carboxylate amino acids while maintaining a negative charge; (phosphonono)alanine; sulfated threonine. Carboxyl side groups (e.g., aspartyl or glutamyl) can also be selectively modified by reaction with carbodiimides (R'-N-C-N-R') such as, e.g., 1-cyclohexyl-3(2-morpholinyl-(4-ethyl) carbodiimide or 1-ethyl-3(4-azonia- 4,4-
dimetholpentyl) carbodiimide. Aspartyl or glutamyl can also be converted to asparaginyl and glutaminyl residues by reaction with ammonium ions. Mimetics of basic amino acids can be generated by substitution with, e.g., (in addition to lysine and arginine) the amino acids ornithine, citrulline, or (guanidino)-acetic acid, or (guanidino)alkyl-acetic acid, where alkyl is defined above. Nitrile derivative (e.g., containing the CN-moiety in place of COOH) can be substituted for asparagine or glutamine. Asparaginyl and glutaminyl residues can be deaminated to the corresponding aspartyl or glutamyl residues. Arginine residue mimetics can be generated by reacting arginyl with, e.g., one or more conventional reagents, including, e.g., phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanediol, or ninhydrin, in one aspect under alkaline conditions. Tyrosine residue mimetics can be generated by reacting tyrosyl with, e.g., aromatic diazonium compounds or tetranitromethane. N-acetylimidazolyl and tetranitromethane can be used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively. Cysteine residue mimetics can be generated by reacting cysteinyl residues with, e.g., alpha-haloacetates such as 2-chloroacetic acid or chloroacetamide and corresponding amines; to give carboxymethyl or carboxyamidomethyl derivatives. Cysteine residue mimetics can also be generated by reacting cysteinyl residues with, e.g., bromo-trifluoracetone, alpha-bromo-beta-(5-imidozoyl) propionic acid; chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide; methyl 2-pyridyl disulfide; p-chloromercuribenzoate; 2-chloromercuri-4-nitrophenol; or, chloro-V-nitrobenzo-oxa-l^-diazole. Lysine mimetics can be generated (and amino terminal residues can be altered) by reacting lysinyl with, e.g., succinic or other carboxylic acid anhydrides. Lysine and other alpha-amino-containing residue mimetics can also be generated by reaction with imidoesters, such as methyl picolinimidate, pyridoxal phosphate, pyridoxal, chloroborohydride, trinitrobenzenesulfonic acid, O-methylisourea, 2,4, pentanediol, and transamidase-catalyzed reactions with glyoxylate. Mimetics of methionine can be generated by reaction with, e.g., methionine sulfoxide. Mimetics of proline include, e.g., piperolic acid, thiazolidine carboxylic acid, 3- or 4- hydroxy proline, dehydroproline, 3- or 4-methylproline, or 3,3-dimethylproline. Histidine residue mimetics can be generated by reacting histidyl with, e.g., diethylprocarbonate or para-bromophenacyl bromide. Other mimetics include, e.g., those generated by hydroxylation of proline and lysine; phosphorylation of the hydroxyl groups of seryl or threonyl residues; methylation of the alpha-amino groups of lysine, arginine and histidine; acetylation of the N-terminal amine; methylation of main chain
amide residues or substitution with N-methyl amino acids; or amidation of C-terminal carboxyl groups.

A residue, e.g., an amino acid, can also be replaced by an amino acid (or peptidomimetic residue) of the opposite chirality. Thus, any amino acid naturally occurring in the L-configuration (which can also be referred to as the R or S, depending upon the structure of the chemical entity) can be replaced with the amino acid of the same chemical structural type or a peptidomimetic, but of the opposite chirality, referred to as the D-amino acid, but also can be referred to as the R- or S-form.

The invention also provides methods for modifying the polypeptides or peptides of the invention, or polypeptides or peptides used in the compositions and method of the invention, by either natural processes, such as post-translational processing (e.g., phosphorylation, acylation, etc), or by chemical modification techniques, and the resulting modified polypeptides. Modifications can occur anywhere in the polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also a given polypeptide may have many types of modifications. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of a phosphatidylinositol, cross-linking cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, and transfer-RNA mediated addition of amino acids to protein such as arginylation. See, e.g., Creighton, T.E., Proteins - Structure and Molecular Properties 2nd Ed., W.H. Freeman and Company, New York (1993); Posttranslational Covalent Modification of Proteins, B.C. Johnson, Ed., Academic Press, New York, pp. 1-12 (1983).

Solid-phase chemical peptide synthesis methods can also be used to synthesize the polypeptides or peptides of the invention or used to practice the invention, e.g., Merrifield (1963) J. Am. Chem. Soc. 85:2149-2154; Stewart, J.M. and Young, J.D., Solid Phase Peptide Synthesis, 2nd Ed., Pierce Chemical Co., Rockford, pp. 11-12) and are employed in commercially available laboratory peptide design and synthesis kits.
Such commercially available laboratory kits have generally utilized the teachings of H. M. Geysen et al, Proc. Natl. Acad. ScL, USA, 81:3998 (1984) and provide for synthesizing peptides upon the tips of a multitude of "rods" or "pins" all of which are connected to a single plate. When such a system is utilized, a plate of rods or pins is inverted and inserted into a second plate of corresponding wells or reservoirs, which contain solutions for attaching or anchoring an appropriate amino acid to the pin's or rod's tips. By repeating such a process step, i.e., inverting and inserting the rod's and pin's tips into appropriate solutions, amino acids are built into desired peptides. In addition, a number of available Fmoc peptide synthesis systems are available. For example, assembly of a polypeptide or fragment can be carried out on a solid support using an Applied Biosystems, Inc. Model 431A™ automated peptide synthesizer. Such equipment provides ready access to the peptides of the invention, either by direct synthesis or by synthesis of a series of fragments that can be coupled using other known techniques.

EXAMPLES

Example 1: C. difficile Cell Wall Binding (CWBVDomains of the Invention)

In one aspect, the invention provides novel structural and functional C. difficile CWB-domains, e.g., residues 2459-2710 of toxin A, and compositions and methods using these novel C. difficile CWB-domains.

Physical Properties of the C. difficile Toxin A and B CWB-domains

Unlike the repeat regions of LytA (13), the C. difficile CWB-repeats did not form unique tertiary structures with a minimum of six repeats. Three separate toxin CWB-domain constructs (ToxA:6R = residues 1800-1945 of toxin A; ToxA:7R = 2078-2234 of toxin A and ToxB:7R = 2208-2366 of toxin B, Figure 1b) composed of six or seven repeats did not form a tertiary fold. All three domains did maintain non-random (perhaps β-solenoid-type) secondary structures based on far UV CD measurements. The spectra of all three constructs have a minimum at 212 nm unlike random coil structures (Figure 2a) and similar to what has been observed for the CWB-containing phage Cpl-I lysozyme (22). However, the three domains all lacked a near UV CD signal and thermal unfolding transition indicative of a tertiary fold. The far UV CD spectra of the three toxin domains did not change appreciably when heated to 95 °C suggesting that the non-random secondary structure observed is an aspect of the unfolded state of the toxin A and
B CWB-domains. The majority of the expressed protein was found in inclusion bodies, another indication that all three domains were not well folded. ToxA:6R and ToxA:7R isolated from inclusion bodies using Ni²⁺ capture in urea and/or reverse phase HPLC failed to refold to the unique secondary structure of the soluble protein fraction and had far UV CD spectra indicative of random coil (data not shown).

An eleven repeat construct, ToxA:11R (residues 2459-2710 of toxin A), is cooperatively folded with a far UV CD spectrum identical to that of the full-length CWB-domains of both toxin A and B (denoted ToxA:40R = residues 1800-2710 of toxin A; ToxB:25R = residues 1807-2366 of toxin B). The far UV CD spectra of the folded domains were distinguishable from the 6 and 7 repeat domains containing an additional, strongly positive band at 230 nm (Figure 2a). Unlike the 6 and 7 repeat CWB-domains, ToxA:11R displayed a negative near UV CD peak with a minimum at 284 nm indicating the burial of aromatic residues within a tertiary fold (Figure 2b). The full-length repeat domains, ToxA:40R and ToxB:25R, as well as ToxA:1IR were monomeric as judged by gel filtration. The three toxin domains with 6 or 7 repeats associated with the gel filtration matrix, typical of partially folded proteins with exposed hydrophobic residues. ToxA:1IR underwent a single unfolding transition when heated. Its melting temperature (T_M) as measured by both near (284 nm) and far (230 nm) UV CD was 49 °C (Figure 3a). This unfolding temperature agrees with early literature demonstrating that toxins A and B are no longer active at 56 °C (1). The spectrum of thermally unfolded ToxA:HR resembled the spectra of the smaller, 6-7 repeat domains indicating the presence of similar secondary structure in the denatured state of ToxA:1IR (data not shown).

Similar to pneumococcal LytA, ToxA:HR binds choline (23,24), but did not bind Galα1-3Galβ1-4GlcNac as has been reported for full-length toxin A (15,25). Binding was assessed by the addition of various amounts of choline or Galα1-3Galβ1-4GlcNac and testing for an increase in the T_M of the protein upon binding (Figure 3b). ToxA:1IR bound choline with an apparent dissociation constant of 13 ± 5 mM assuming a 1:1 interaction. The small CWB-domains of LytA and Cpl-I both bind numerous choline moieties (13,42). Adding up to 10 mM Galα1-3Galβ1-4GlcNac did not affect the T_M of TOXA:1IR. Additionally, surface plasmon resonance experiments failed to detect any interaction between all toxin A CWB-domain constructs and a sensor chip surface coated with Galα1-3Galβ1-4GlcNac conjugated BSA (data not shown). The toxin A CWB-domains were applied to the surface at concentrations as high as 1 µM.
While there is no published evidence citing the necessity of metals for CWB-domain cell surface recognition, it is common for carbohydrate or phospholipid binding proteins to utilize metal ions when binding to their respective ligands (26, 27). The importance of metal ion binding has been established for the intracellular UDP-glucose hydrolysis activity of toxins A and B (28) and extracellular calmodulin has been shown to have an effect on toxin internalization (29). Intracellular Ca\(^{2+}\) levels have been demonstrated not to have an effect on toxin activity (30). Based on these factors, we tested the CWB-domains for their ability to coordinate Ca\(^{2+}\) or Mg\(^{2+}\) and for the potential requirement of these metal ions for CWB-domain association with mammalian cells.

Ca\(^{2+}\), but not Mg\(^{2+}\), was shown to bind ToxA:1IR as judged by its thermostability in the presence and absence of each ion. Addition of 10 mM Mg\(^{2+}\) had no effect on the stability of the domain; whereas a strong Ca\(^{2+}\)-dependent increase in the protein's T\(_M\) was observed with increases in CaCl\(_2\) concentrations from 1 to 20 mM. An apparent dissociation constant, K\(_D\) = 8 ± 2 mM at 25 °C, was calculated based on the assumption of a 1:1 binding model (Figure 3b). This level of affinity would be insignificant for intracellular Ca\(^{2+}\)-binding proteins; however, extracellular Ca\(^{2+}\)-binding proteins tend to have K\(_D\)-values between 10\(^{-4}\) and 10\(^{-3}\) M due to much higher Ca\(^{2+}\) levels within the extracellular space (31). Oligomerization was not a function of Ca\(^{2+}\) or choline binding as ToxA:HR, ToxA:40R and ToxB:25R were all monomeric in the absence and presence of 10 mM CaCl\(_2\) or 5 mM choline as judged by gel filtration. Gel filtration studies were performed as described previously (32). The addition of Ca\(^{2+}\) also did not effect the inability of the repeat domains to bind Gal\(\alpha\)1-3Gal\(\beta\)1-4GlcNac.

**Cell Surface Association of Toxin A CWB-domains**

Using flow cytometry, we demonstrated Ca\(^{2+}\)-dependent toxin A CWB-domain binding to CHO cell surfaces. CHO cells have been used previously to study the toxigenic effects and/or binding capabilities of toxins A and B (33-36). In the absence of Ca\(^{2+}\), the full length CWB-domain of toxin A, ToxA:40R, and the folded 11 repeat domain of toxin A, ToxA:HR, did not bind CHO cells (Figure 4c). Titration of CaCl\(_2\) from approximately 100 µM up to 4 mM induced a strong interaction between ToxA:HR and CHO cell surfaces (Fig 4b-e). Binding of ToxA:1 IR to CHO cells was also dependent on the concentration of the protein. Doubling the ToxA:1 IR concentration from 50 to 100 µg/mL increased the population of fluorescently-labeled cells by ~30% in the presence of 1 mM Ca\(^{2+}\) (Figure 4f). Ca\(^{2+}\)-dependent cell surface binding of
ToxA:40R was also detectable, but the fluorescence signal was limited by the comparatively low molar amount of protein available for the assay (data not shown). The control His-Tagged protein, bovine IgGl C1H3 (32), did not associate with CHO cells at protein concentrations as high as 200 µg/mL and Ca2+ concentrations as high as 10 mM.

The smaller, 6 to 7 repeat toxin A domains, ToxA:6R and ToxA:7R, associated with CHO cell surfaces weakly in the presence and absence of Ca2+ (data not shown). This binding may be a result of their partially folded nature and not an intrinsic function of the domains (37,38). Addition of 10 mM CaCl2 to solutions of ToxA:6R and ToxA:7R did not induce folding or structural changes as judged by CD. At 1 mM CaCl2, the fluorescent cell population was 28.1% and 25.0% for 50 µg/mL ToxA:6R and ToxA:7R, respectively; both weaker than the fluorescent shift induced by ToxA:1 IR (data not shown). Increasing the Ca2+ concentration to 4 mM weakly increased the fluorescent population shift generated by ToxA:6R and ToxA:7R (-2-10%); however, this increase in the fluorescent cell population was much weaker than what was observed when 4 mM Ca2+ was added to ToxA:1IR in the assay.

We next tested whether extracellular Ca2+ played a role in toxin-mediated cell killing. CHO cells were treated with 0.4 and 0.08 µg of full length toxin A and incubated at 37 °C. Both toxin concentrations induced 100% cell rounding after 48 hour incubations using non-Ca2+-depleted DMEM media. In our experience, cell rounding inevitably led to the detachment or death of adherent CHO cells and was a strong indicator of toxin A cytotoxicity. In the presence of 0.4 µg toxin A, cells rounded rapidly independent of the Ca2+ concentration, perhaps due to a saturating level of toxin (data not shown). At 0.08 µg toxin A, Ca2+ had a repeatable and definitive effect on the kinetics of cell rounding. The largest difference in cell rounding between cells incubated in Ca2+-depleted versus Ca2+-rich media appeared at 5 hours (Figure 3c). Toxin A mediated cell rounding was more rapid at Ca2+ concentrations above 1 mM which correlates well with the apparent affinity of Ca2+ for the CWB-domain of toxin A and the Ca2+ levels necessary to induce CWB-domain binding to CHO cell surfaces. Low levels of Ca2+ (100 µM and below) also appeared to enhance toxin susceptibility (data not shown). Whether this enhancement at low Ca2+ levels is a direct response by the toxin or complicated by cellular responses to low Ca2+ is undetermined. In the absence of toxin, low Ca2+ levels altered the normal elongated CHO cell morphology and retard their ability to reach confluence compared to cells incubated at higher Ca2+ concentrations. This may explain
the apparently greater susceptibility of CHO cells to toxin A at very low extracellular Ca$^{2+}$ levels.

**Discussion**

Contrary to what has been observed for LytA (13), none of the 6-7 contiguous toxin A or B repeat domains in this study formed tertiary structure similar to what was observed for the full-length toxin CWB-domains, even though they did form non-random, presumably β-solenoid-like secondary structure. Both toxins A and B have unique 30 residue peptide stretches after approximately every sixth 20 residue P-motif. These unique stretches likely add complexity to the tertiary fold of the CWB-domains of toxins A and B, differentiating these domains from the CWB-domain of LytA. The necessity of several additional repeats to form a tertiary fold may explain why no structure has been published for the toxin CWB-domains. Interestingly, the nuclear magnetic resonance assignments of a 5-repeat stretch of toxin A were recently deposited into the BMRB (39). The spectrum of this construct is well dispersed and suggests that the CROP polypeptide is folded. However, the CD data reported here indicates that there is likely a tertiary structure consequence to having additional CROPs. The fact that all 6 and 7 repeat CWB-domains analyzed in this study lack the ability to fold (including the C-terminal ToxB:2208-2366 which contains two more repeats than the related C-terminal toxin A CWB-construct published in the NMR study) suggests that larger constructs will be necessary to properly define the toxin CWB-fold.

The toxin A CWB-domains bound CHO cell surfaces in a Ca$^{2+}$-dependent fashion. There is little published evidence to our knowledge that definitively links cell surface association of toxin A to the CWB-domains. A recent study by Pfeifer et al. (40) demonstrates the localization of residues 547-2366 of toxin B in membrane fractions of Vero cells using radioactively labeled-protein. This toxin B construct lacks the cytotoxic domain of the molecule but includes the putative transmembrane region, the 700 residue domain of unknown function, as well as the CWB-domain. This fragment of toxin B has been reported to form pores within membranes by a pH inducible mechanism (33). Another study by Aktories and coworkers demonstrated that a construct similar to ToxA: 1IR, named REP23 1, could bind F9-cell surfaces, but only at relatively high concentrations, 200 µg/mL, and after pretreatment of cells with 4% paraformaldehyde (19). No binding was detected in their assay with CHO cells (Ca$^{2+}$ conditions not discussed). At very high concentrations (200 µg/mL), REP231 partially inhibited toxin A
mediated cell killing. This result indirectly supports the ability of CWB-domains to bind and saturate toxin A receptors on the surfaces of F9 cells. Just and coworkers have shown that a toxin A fragment containing the entire CWB-domain can undergo endocytosis in HT29 cells; however, addition of N-terminal amino acids from the central 700 residue domain more effectively neutralized toxin A suggesting a potential role for this central domain in endocytosis/cell recognition (41). Deletion of the C-terminal half of the toxin A CWB-domain eliminated its capability for endocytosis (41). Other studies also report the effect of deleting the CWB-domains from the toxins in an attempt to more fully characterize their function. Deletion of the entire CWB-domain of toxin B only attenuated its toxicity 10-fold (3). Similar to what was observed for toxin B, removal of one or two 20-residue repeats from the LytA choline binding domain attenuated its function, but did not delete it. Thus, the in vivo function of the CWB-domain and the necessity for various numbers of repeats for the potency of toxins A and B is still not entirely clear.

The importance of Ca$^{2+}$ for the function of the CWB-domain of toxin A is a novel property associated with this domain. Here we report that Ca$^{2+}$-binding was important for CWB-domain binding to CHO cell surfaces and appeared to influence the kinetics of toxin A cytotoxicity. This effect correlated well with the apparent affinity of Ca$^{2+}$ for the CWB-domains themselves. It will be interesting to determine whether in vivo Ca$^{2+}$ levels in the gastro-intestinal tract influence the virulence of C. difficile infection.

The CWB-domain of toxin A also bound choline as has been described for other proteins with similar repeats (25). The binding was weak, i.e. KD $\sim$ 10$^{-7}$ M, and likely leads to only transient associations of the domain with choline moities. The fact that 2% free choline (~300 mM) is necessary to inhibit the cell wall binding of LytA to pneumonococci (24) indicates that the relative affinity of LytA for choline is low, similar to what was observed here for the CWB-domain of toxin A. It has been suggested that choline induced dimerization/oligomerization of the LytA CWB-domain is important for positioning its amidase domain into the peptidoglycan layer (43). On the contrary, ToxA:40R, ToxB:25R and ToxA:1 IR were all monomeric in the presence of 10 mM Ca$^{2+}$ and in the presence of 5 mM choline suggesting oligomerization is not a function associated with Ca$^{2+}$ or choline binding. Choline binding has been reported to stabilize the structures of the CWB-domains of the pneumonococcal phage Cp-I lysin (CpI-I, 42) and LytA proteins (43), similar to what we observed for the toxin A CWB-construct.
Interestingly, the strongly positive CD band observed for the folded toxin A and B constructs only appears in the CpI-I far UV spectrum when choline is introduced into the solution (22). The choline moiety is completely surrounded by aromatic residues of the β-solenoid fold and it has been suggested that choline orients these aromatics in CpI-I in a conformation that induces the strongly positive far UV CD band (42). Choline did not induce changes in the CD spectra of the toxin A and B CWB-domains. Instead, these domains appear to intrinsically contain the strongly positive CD band at 230 nm.

The potential ability of the toxin A CWB-domains to associate with the phosphatidylcholine within mammalian cells could allow the toxin molecule to translate in 2-dimensions along the membrane surface to find its glycoprotein receptor increasing its apparent affinity. Phosphatidylcholine is generally the most abundant phospholipid in mammalian bilayers (>50%). This mechanism is consistent with the fact that toxin B and other CWB-domain containing proteins exhibit attenuations in their apparent function upon truncation or deletion of their CWB-domain as opposed to a total loss of function (3). The ability of toxin A to bind choline also suggests the functionality of its CWB-domain may not be limited to the recognition of mammalian cell surface receptors. These domains may also be important for toxin association to the lipoteichoic acid layer on the surface of C. difficile before toxin delivery to mammalian cells. Association of toxins with the C. difficile coat can help localize the toxins to the site of bacterial adherence and/or infection aiding the infiltration of the colonic mucosa.

Figures

Figure 1. a. Alignment of the consensus repeat sequences of CWB-domains from Streptococcus mutans GtfC and GtfB, Streptococcus downei GtfI, Streptococcus pneumoniae LytA, pneumococcal phage CpI-I, C. difficile toxins A and B and C. Sordelii Toxin L. Repeat residues that are completely conserved in every repeat of each protein are underlined. Toxins A and B were cloned from genomic DNA preparations of C. difficile (ATCC #51695). Cultures were grown in a thioglycollate anaerobic broth (BBU273127) at 37 °C for 48-72 hours prior to DNA extraction. Inserts were generated for toxin A and toxin B CWB-repeat regions using primers matched to the NCBI deposited sequences (Toxin B accession no. CAA80815.1; toxin A accession no. AAA23283.1-C. difficile strain VP10463). Inserts were cloned into the pSE420 bacterial expression plasmid modified in-house to include a C-terminal hexa-histidine tag (Invitrogen, Cat#V4020). Cloned toxin B sequences matched the NCBI reference while
toxin A clones consistently had 4 amino acid mutations (N1939D, L2080W, D2426H and A2427N) which were likely strain specific. b. Schematic diagram of the CWB-domain constructs that were subcloned. Each toxin containing plasmid was transformed into BL21(DE3) (Invitrogen, cat#C60000-03) or recA deficient XL1 -blue (Stratagene, cat#200236) for improving the insert stability of the repeat domains. Another strategy for cloning and expressing the toxin domains was recently reported (44). Protein expression was performed following standard protocols (45). Briefly, transformed cells were cultured in 2-6 L Luria Broth with 100 µg/mL carbenicillin at 37 °C and induced with 1 mM IPTG upon reaching an OD600 between 0.6 and 0.8 AU. Cultures were grown for 12 hours at 25 °C, pelleted and stored at -20 °C before resuspension in PBS (Sigma, Cat#P-3813) and sonication. Soluble protein fractions were passed over a Ni²⁺-bound HiTrap chelating column and eluted with an imidazole gradient (Amersham, Cat#17-0409-01). Toxin B domains were dialyzed against a 50 mM MES, pH 7.9 buffer and further purified using a HiTrap-SP prepacked ion exchange resin (Amersham, Cat#17-1 151-01) and a NaCl gradient. Ni²⁺-purified toxin A domains were dialyzed against 50 mM Tris, pH 7.0 before passage over a HiTrap-CM prepacked ion exchange resin (Amersham, Cat#17-5056-01) and eluted with salt. Stock concentrations were determined by UV absorbance using the method of Pace and coworkers (46).

Figure 2. a. Far UV CD spectra of toxin CWB-domains at pH 7.4, 25 °C. CD spectra were taken on an Aviv model 215 spectrophotometer. All Far UV spectra were the average of at least three scans utilizing a signal averaging time of 3s/λ and a 1 nm bandwidth. Protein solutions were maintained in a 5 mM phosphate, 10 mM NaCl buffer at pH 7.4. b. Near UV CD spectra of toxin CWB-domains under the same conditions. Near UV spectra were the average of 5 scans with a 3s/λ signal averaging time and a 2 nm bandwidth.

Figure 3: Ca²⁺-binding to ToxA:1 IR and Ca²⁺-accelerated CHO cell killing by full length toxin A. a. Thermal denaturation of ToxA: 1IR monitored by near and far UV CD. b. Thermal unfolding transitions of ToxA: 1IR in the presence of EDTA, excess calcium and/or choline. All thermal transitions were fit to a two-state unfolding model (47) using a theoretical ΔCp* , 4064 cal/mol K (1 cal = 4.184 J) based on the size of the protein (48). Equilibrium dissociation constants at 25 °C for small molecules bound to ToxA: 1IR were determined by measuring the equilibrium unfolding constant in the absence and presence of ligand (49):
where $K_u$ is the equilibrium unfolding constant of ToxA:llR in the absence of ligand, $K_{app}$ is the apparent equilibrium unfolding constant in the presence of ligand and $[L]$ is the ligand concentration. Choline dihydrogen citrate was purchased from Sigma and unconjugated linear B2-trisaccharide (Galα1-3Galβ1-4GlcNac) from V-LABS. c. The percentage of rounded CHO cells (\#round/[\#round+\#flat] * 100\%) after incubation with a killing dose of 80 μg toxin A was determined in the presence of various concentrations of CaCl$_2$. Kinetic measurement of toxin A mediated cell rounding at various calcium concentrations was performed by culturing the CHO cells in calcium depleted Dulbecco’s modified Eagle’s medium (DMEM, Sigma cat#M8167) doped with L-glutamine and 5% fetal bovine serum (FBS) at 37 °C in a CO$_2$ incubator. Prior to the experiment, cells were split into T75 flasks at 1.5 X 10$^6$ cells/flask. 96-well cell culture plates were seeded from these stocks (Costar, cat#3598) and incubated for 4-6 hours before addition of 80 ng/mL Toxin A (List Biological Laboratories). CaCl$_2$ was doped into the DMEM solution to create a gradient consisting of 300 μM, 700 μM, 1 mM, 2 mM, 5 mM Ca$^{2+}$. The residual Ca$^{2+}$ introduced by FBS is below 100 μM (Invitrogen, personal communication). Above 5 mM, CaCl$_2$ began to precipitate in the DMEM media precluding higher calcium concentrations. Time points were taken every hour after the addition of toxin A (up to 5 hrs) and after overnight incubation. At least 150 cells were counted at three or more positions within each well to account for variability in cell rounding due to differential confluency within each well.

Figure 4: a-f. Calcium dependent binding of ToxA: 11R to CHO cell surfaces determined by flow cytometry. CHO-K1 cells (ATCC CCL-61) were maintained in DMEM (Gibco, cat#12430-054) supplemented with 10% fetal bovine serum at 37 °C in a CO$_2$ incubator. CHO cells were cultured at a ratio of 1:10 and grown in 10 cm dishes, washed two times with PBS, scraped, and pelleted at 1100 rpm at 4 °C for 5 minutes. Cell staining was performed with 5X10$^5$ cells/tube in wash buffer (PBS supplemented with 2.5 mM Hepes, 0.1% sodium azide, and 2% FBS). ToxA:6R, ToxA:7R, ToxA:11R and ToxA:40R were incubated with 0.5X10$^6$ CHO cells for 20 minutes on ice with various protein and Ca$^{2+}$ concentrations. Cells were spun down and washed two times with 2 mL wash buffer. Adhered toxins were detected via their histidine tags by
incubation with the Alexafluor 488 conjugated PENTA-HIS monoclonal antibody (1:500 dilution, Qiagen, cat#35310) for 20 minutes. The cells were washed twice more and resuspended in 500 µL wash buffer. Flow cytometry analyses were performed on a DakoCytofation MoFIO flow cytometer (Fort Collins, CO) equipped with a Coherent Enterprise II (Santa Clara, CA) water-cooled argon ion multi-line laser. The 488 nm line was used as the excitation source. Forward scatter (FSC), side scatter (SSC) and fluorescent properties were detected by R928 photomultiplier tubes (Hammamatsu, Shizuoka-ken, Japan). Fluorescence was detected between 510 and 550 nm. Data was collected for 10,000 events and was analyzed using DakoCyotmation Summit v3.1 software, (a) Forward- and side-scatter profiles of the CHO cell suspension. The gate used to discriminate live cells was based on photomultiplier counts between $10^3$ and $10^4$ for both side-scatter and forward-scatter, (b-f) Fluorescence profile of the gated CHO cells incubated in the presence of various Ca$^{2+}$ and ToxA:1 IR concentrations.

References


A number of aspects of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. Accordingly, other embodiments are within the scope of the following claims.
WHAT IS CLAIMED IS:

1. A method of ameliorating or preventing toxicity associated with a virulence factor in a subject, comprising administering to the subject in need thereof, a therapeutically effective amount of an isolated or recombinant protease, wherein the protease partially or completely neutralizes the virulence factor, thereby ameliorating or preventing the toxicity associated with the virulence factor.

2. The method of claim 1, wherein the virulence factor is an adherence factor, a coat protein, an invasion factor, a capsule, or a toxin.

3. The method of claim 2, wherein the toxin comprises an exotoxin or an endotoxin.

4. The method of claim 3, wherein the protease cleaves the cell wall binding domain of the toxin.

5. The method of claim 4, wherein the cell wall binding domain comprises clostridial repetitive oligopeptides or a P-motif.

6. The method of claim 3, wherein the toxin produced by a Clostridium bacteria.

7. The method of claim 6, wherein the toxin is Clostridium difficile toxin A.

8. The method of claim 1, wherein the protease is stable in cecal fluid supernatant and homogenized cecal content in vitro for 30 minutes at 37°C.

9. The method of claim 1, wherein the protease is modified to improve its stability in the presence of cecal fluid supernatant or homogenized cecal content in vitro.

10. The method of claim 1, wherein the protease is modified to increase its resistance to proteolysis.

11. The method of claim 10, wherein the proteolysis is mediated by pepsin, pancreatin, trypsin, trypsinogen, chymo-trypsinogen, carboxy-peptidase, procarboxy-peptidase, elastase, pro-elastase, or a combination thereof.
12. The method of claim 1, wherein the protease is modified to have
   a) an chemical stability superior to the unmodified protease;
   b) a thermotolerance or thermostability superior to the unmodified protease;
   c) a pH tolerance superior to the unmodified protease;
   d) a reduced immunogenicity;
   e) a stability in pharmaceutical formulation superior to that of the unmodified protease;
   f) an increased half-life relative to the unmodified protease;
   g) a stable or increased recombinant expression in a host cell; or
   h) a combination thereof.

13. The method of claim 1, wherein the protease is recombinantly expressed in
   a mammalian, yeast, bacterial, insect, or plant cell.

14. The method of claim 12, wherein the protease is modified by an addition,
   deletion, or substitution in the amino acid sequence of the protease.

15. The method of claim 12, wherein the protease is modified by one or more
   additions, deletions, or substitutions to the nucleic acid sequence encoding the protease
   polypeptide.

16. The method of claim 15, wherein the additions, deletions, or substitutions
   to the nucleic acid sequence encoding the protease are introduced by a method
   comprising error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly
   PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive
   ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis,
   gene reassembly, Gene Site Saturation Mutagenesis (GSSM), synthetic ligation
   reassembly (SLR), or a combination thereof.

17. The method of claim 15, wherein the additions, deletions, or substitutions
   to the nucleic acid sequence encoding the protease are introduced by a method
   comprising recombination, recursive sequence recombination, phosphothioate-modified
   DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis,
   point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical
mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation, or a combination thereof.

18. The method of claim 1, wherein the protease is formulated for oral delivery, or an enteral formulation.

19. The method of claim 1, further comprising administering one or more additional anti-virulence factor agents.

20. The method of claim 19, wherein the additional agent is administered simultaneously or sequentially with the protease.

21. The method of claim 19, wherein the additional anti-virulence factor agent is one or more monoclonal antibodies.

22. The method of claim 1, wherein the subject is pretreated to lower the pH of the intestine or increase the pH of the stomach.

23. The method of claim 1, wherein the subject is pretreated with at least one antibiotic.

24. The method of claim 1, wherein the toxicity associated with the virulence factor causes a gastrointestinal disease or disorder.

25. The method of claim 1, wherein the gastrointestinal disease or disorder is an ulcer, diarrhea, pseudomembranous colitis, or cancer.

26. A method of identifying a protease useful in the amelioration or prevention of virulence factor-associated toxicity, which comprises
   a) contacting a virulence factor with a test protease;
   b) assessing the proteolytic cleavage of the virulence factor by the protease; and
   c) identifying the protease that cleaves the virulence factor as the protease that is useful in the amelioration or prevention of virulence factor-associated toxicity.
27. The method of claim 26, wherein step b) comprises
   i) incubating the virulence factor-protease mixture from step a) in an ELISA
       binding plate;
   ii) probing the proteins bound to the ELISA plate with an antibody that binds
       the virulence factor; and
   iii) determining the amount of bound antibody, whereby the absence of bound
       antibody indicates that the protease cleaved the virulence factor.

28. The method of claim 26, wherein the virulence factor is an adherence
    factor, a coat protein, an invasion factor, a capsule, or a toxin.

29. The method of claim 28, wherein the toxin is Clostridium difficile toxin.

30. The method of claim 26, wherein step a) is performed in conditions
    resembling those present in the digestive tract.

31. The method of claim 26, wherein step a) is performed in the presence of
    cecal content homogenate.

32. The method of claim 26, further comprising a step d) for selecting a
    protease from step c) that
    i) retains greater than 75% proteolytic activity after 30 minutes in cecal fluid
        supernatant at 37°C;
    ii) retains greater than 40% proteolytic activity after 30 minutes in cecal
        content homogenate at 37°C;
    iii) exhibits sufficient activity to degrade greater than 90% of the toxin in less
        than 30 minutes;
    iv) increases the percentage of cells surviving after coincubation with toxin to
        greater than 50%; or
    v) partially or completely neutralizes the toxin-induced toxicity in the rat ileal
       loop model.

33. The method of claim 26, wherein the neutralizing activity in step v) is
    similar to or greater than the neutralizing activity of cholestyramine.
34. The method of claim 26, further comprising a step d) for modifying the identified protease to have
   a) an increased stability in the presence of cecal fluid supernatant or homogenized cecal content \textit{in vitro};
   b) an chemical stability superior to the unmodified protease;
   c) a thermotolerance or thermostability superior to the unmodified protease;
   d) a pH tolerance superior to the unmodified protease;
   e) a reduced immunogenicity;
   f) a stability in pharmaceutical formulation superior to that of the unmodified protease;
   g) an increased half-life relative to the unmodified protease;
   h) a stable or increased recombinant expression in a host cell; or
   i) a combination thereof.

35. The method of claim 34, wherein step d) further comprises selecting a protease from step c) that has
   a) an increased stability in the presence of cecal fluid supernatant or homogenized cecal content \textit{in vitro};
   b) an chemical stability superior to the unmodified protease;
   c) a thermotolerance or thermostability superior to the unmodified protease;
   d) a pH tolerance superior to the unmodified protease;
   e) a reduced immunogenicity;
   f) a stability in pharmaceutical formulation superior to that of the unmodified protease;
   g) an increased half-life relative to the unmodified protease;
   h) a stable or increased recombinant expression in a host cell; or
   i) a combination thereof.

36. The method of claim 34, wherein the modification is an addition, deletion, or substitution in the amino acid sequence of the protease.

37. The method of claim 34, wherein the modification of protease is introduced by one or more additions, deletions, or substitutions to the nucleic acid sequence encoding the protease polypeptide.
38. The method of claim 37, wherein the additions, deletions, or substitutions to the nucleic acid sequence encoding the protease are introduced by a method comprising error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, Gene Site Saturation Mutagenesis (GSSM), synthetic ligation reassembly (SLR), or a combination thereof.

39. The method of claim 37, wherein the additions, deletions, or substitutions to the nucleic acid sequence encoding the protease are introduced by a method comprising recombination, recursive sequence recombination, phosphothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation, or a combination thereof.

40. The method of claim 26, wherein the identified protease cleaves *C. difficile* toxin A within residues 1812-2710 of the toxin.

41. The method of claim 26, wherein the identified protease cleave *C. difficile* toxin B within residues 1814-2366 of the toxin.

42. A method of identifying a compound that inhibits the binding of a *Clostridium sp.* toxin with mammalian cells, which comprises

a) contacting a mammalian cell with a recombinant, functional cell wall binding domain polypeptide from the *Clostridium sp.* toxin and a test compound in the presence of Ca$^{2+}$;

b) determining the amount of cell binding by the cell wall binding domain polypeptide; and

c) identifying the test compound as an inhibitory compound if it reduces or eliminates binding of the cell wall binding domain polypeptide to the cell.
43. A method of identifying an inhibitory compound that inhibits the binding of a _Clostridium sp._ toxin with mammalian cells, which comprises
   a) contacting a mammalian cell with a _Clostridium sp._ toxin and a test compound in the presence of Ca\(^{2+}\);
   b) determining the amount of cell binding by the cell wall binding domain polypeptide; and
   c) identifying the test compound as an inhibitory compound if it reduces or eliminates binding of the cell wall binding domain polypeptide to the cell.

44. The method of claim 42 or 43, wherein the toxin is _C. difficile_ toxin.

45. The method of claim 42 or 43, wherein the cell wall binding domain polypeptide is ToxA:2459-2710.

46. The method of claim 42 or 43, wherein the compound reduces or eliminates binding of the full length toxin to the cell.

47. The method of claim 42 or 43, wherein the contacting occurs in the presence of ImM CaCl\(_2\).

48. The method of claim 42 or 43, wherein the compound reduces the toxicity of the full length toxin to mammalian cells.

49. A method of ameliorating or preventing toxicity associated with _Clostridium sp._ toxin in a subject, comprising administering to a subject in need thereof, a therapeutically effective amount of a compound identified by the method of claim 42 or 43.

50. A method of ameliorating or preventing toxicity associated with _Clostridium difficile_ toxin A in a subject, comprising administering to a subject in need thereof, a therapeutically effective amount of a compound which disrupts the Ca\(^{2+}\)-dependent cell surface association of toxin A cell wall binding domain with the surface of a mammalian cell.
Figure 1

a.
StreptmutansGtfC
StreptmutansGtfB
StreptsubrinasGtfI
PhageCpl-1
StreptpneumoniaLytA
CDiffToxA
CDiffToxB
CSordeliiToxL
CDiffToxA
CDiffToxB
CSordeliiToxL

b.

Toxin A

<table>
<thead>
<tr>
<th>Glucosyltransferase</th>
<th>putative function</th>
<th>unknown function</th>
<th>CWB-domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>TM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ToxA:40R = 1800
ToxA:11R = 2459
ToxA:7R = 2078
ToxA:6R = 1800

Toxin B

<table>
<thead>
<tr>
<th>N</th>
<th></th>
</tr>
</thead>
</table>

ToxB:25R = 1807
ToxB:7R = 2208
Figure 2

(a) [\theta] = 10^6 (deg cm^2 dmol^-1)

(b) [\theta] (deg cm^2 dmol^-1)
Figure 4

(a) SSC vs. FSC plot with data points.

(b) FL1 vs. SSC plot with a percentage of 14.4%.

(c) FL1 vs. SSC plot with ~100 μM CaCl₂, 50 μg/mL ToxA:11R.

(d) FL1 vs. SSC plot with 1 mM CaCl₂, 50 μg/mL ToxA:11R.

(e) FL1 vs. SSC plot with 4 mM CaCl₂, 50 μg/mL ToxA:11R.

(f) FL1 vs. SSC plot with 1 mM CaCl₂, 100 μg/mL ToxA:11R.

R2 percentages for each graph are as follows:

- (b) 14.4%
- (c) 14.1%
- (d) 51.7%
- (e) 94.4%
- (f) 79.6%