Structure and mode of action of clostridial glucosylating toxins: the ABCD model

Thomas Jank and Klaus Aktories

Institut für Experimentelle und Klinische Pharmakologie und Toxikologie der Albert-Ludwigs-Universität Freiburg, Otto-Krayer-Haus, Albertstrasse 25, D-79104 Freiburg, Germany

Toxins A and B, which are the major virulence factors of antibiotic-associated diarrhea and pseudomembranous colitis caused by Clostridium difficile, are the prototypes of the family of clostridial glucosylating toxins. The toxins inactivate Rho and Ras proteins by glucosylation. Recent findings on the autcatalytic processing of the toxins and analysis of the crystal structures of their domains have made a revision of the current model of their actions on the eukaryotic target cells necessary.

Introduction
Antibiotic-associated diarrhea and pseudomembranous colitis induced by Clostridium difficile have emerged as important nosocomial infections. During the past decade, these diseases seem to have become more serious and more frequently refractory to therapy [1]. The occurrence of hypervirulent strains of C. difficile such as PCR ribotype 027/PFGE type NAP1 [2] with attributable mortality rates of up to 16.7% is concerning [3]. The major virulence factors of C. difficile are two protein toxins, named toxin A and toxin B (also designated TcdA and TcdB), which have been recognized for their important role in disease for the past 30 years. In addition, some strains of C. difficile produce a binary ADP-ribosylating toxin (C. difficile transferase, CDT) that modifies G-actin [4]. Notably, hypervirulent strains are characterized by production of 10- to 20-fold larger amounts of toxins A and B, the resistance to fluoroquinolones and production of the actin-ADP-ribosylating toxin [2].

Structure–function relationships of clostridial glucosylating toxins
C. difficile toxin A and toxin B are the prototypes of the family of clostridial glucosylating toxins [5,6]. Other members of this toxin family are the hemorrhagic and the lethal toxins from Clostridium sordellii and the α-toxin from Clostridium novyi. Moreover, several isoforms of toxin A and B have been described, adding to this family of cytotoxins [7–9]. All of these toxins share sequence identities ranging from 36% to 90% and have molecular masses between 250 and 308 kDa [10,11]. Because toxins A and B are of major clinical importance and because most data obtained are from these toxins, this review focuses on recent progress in the understanding of the structure–function relationship of these toxins.

Toxins with multimodular structure
Toxin A consists of 2710 amino acid residues with a molecular mass of 308 kDa and toxin B comprises 2366 residues with a mass of 269.6 kDa. These toxins are therefore also known as large clostridial cytotoxins [10]. On the basis of their amino acid sequences, a tripartite structure for the toxins had been suggested [10,12,13], with a biologically active N-terminal domain, a middle translocation section characterized by a small hydrophobic stretch (prediction of transmembrane regions), and a C-terminal receptor-binding domain. This prediction of the structure–function relationship was in line with the model of AB-toxins such as diphtheria toxin, consisting of a biologically active domain and a binding or translocation domain [14]. The binding domain can be separated into subdomains (e.g. for receptor binding and translocation) resulting in a tripartite structure. However, recent studies indicate that a multimodular structure more accurately describes the structure–function relationship of the clostridial glucosylating toxins (Figure 1). These toxins have a biologically active domain and a binding and translocation domain but in addition they have an autocatalytic, self-cutting protease domain for toxin processing. Thus, the AB-toxin model can be extended to an ABCD model (A, biological activity; B, binding; C, cutting; D, delivery).

The C terminus binds to target cell membranes
The receptor-binding domain located at the C terminus of the clostridial glucosylating toxins consists of repeating polypeptide units [12,15–17]. Recently, two C-terminal fragments (terminal 127 and 255 residues) of toxin A (toxinotype VI) were crystallized [18,19]. These studies gave new insights into the overall structure of the C terminus of the toxin (Figure 1). Toxin A folds in a solenoid-like structure with 32 small repeats consisting of 15–21 residues and seven large repeats of 30 residues. Each repeat forms a single β-hairpin that is rotated by 120° to each other, thereby forming a screw-like superfold (Figure 1). Solenoid structures are frequently found in bacterial virulence proteins [20]. They increase the surface area of proteins and enable protein–protein or protein–carbohydrate interactions. Some years ago, Krivan and Tucker showed binding of the trisaccharide Galα1–3Galβ1–4GlcNAc to toxin A [21,22]. This was confirmed and structurally explained by co-crystallization of toxin A with an artificial trisaccharide containing the...
Gal\(\alpha_1-3\)Gal\(\beta_1-4\)GlcNAc-glycan [19]. The carbohydrate-binding groove is formed between the junction of a large repeat and the hairpin turn of the following small repeat by several amino acids. These residues are not conserved in the other clostridial glucosylating toxins and might have a crucial role in carbohydrate receptor specificity. However, human tissue generally does not produce \(\alpha\)-anomeric galactose bonds [23], indicating that the carbohydrate structure Gal\(\alpha_1-3\)Gal\(\beta_1-4\)GlcNAc cannot be part of intestinal receptors in humans. Therefore the disaccharide Gal\(\beta_1-4\)GlcNAc, which is present in humans, has been suggested to be part of a possible glycan receptor. Whether a glycoprotein or glycosphingolipid [24] (either with or without an additional protein receptor) comprises the intestinal receptor remains to be clarified. The C-terminal part of toxin B is probably similar to that of toxin A. However, structural data are not available to date and the nature of the receptor of \(C.\) difficile toxin B is even further from being defined.

**Figure 1.** Domain structure of clostridial glucosylating toxins. The ABCD model (labeled in bold at top of figure) of clostridial glucosylating toxins is shown with *Clostridium difficile* toxin B as an example. The biologically active glucosyltransferase A-domain (bold) is located at the N terminus (amino acids 1–543) and has been crystallized recently. The DXD motif, which is involved in Mn\(^{2+}\) coordination, is located in this domain. The C terminus of the clostridial glucosylating toxins, which consists of polypeptide repeats, is involved in receptor binding (B-domain, bold). A fragment of the C-terminal part of toxin A has been deduced, which consists of 39 repetitive elements [19]. According to this model the repeats have a \(\beta\)-hairpin structure. Rotation of the hairpins by 120° relative to the neighboring hairpin forms a screw-like structure. The cysteine protease domain (residues 544–767), which is similar to *Vibrio cholerae* RTX toxin, is located adjacent to the glucosyltransferase domain. This C-domain (bold) is involved in processing and cutting of the toxin. The cysteine protease C-domain can be characterized by the catalytic triad consisting of Asp587, His653 and Cys698 (DHC). Alternatively, the DXG (D1665) motif was suggested to be part of an aspartate protease domain, possibly involved in processing of the toxins. In the middle part of the protein is a short hydrophobic region (residues 956–1128), which might be involved in pore formation and delivery of the catalytic domain into the cytosol (D-domain, bold). Images were created using PyMOL (www.pymol.org).

**Toxin uptake: a question of cutting and delivery**

Following receptor binding, the clostridial glucosylating toxins are endocytosed [25] (Figure 2) – the precise endocytosis pathway is not known. After endocytosis, the toxins translocate through the early endosomal membrane into the cytosol. This process depends on the acidification of endosomes by vesicular H\(^{+}\)-ATPase. Bafilomycin, which blocks the H\(^{+}\)-ATPase, inhibits cytosolic entry of the toxin and intoxication of cells [26]. Therefore, *C. difficile* toxins A and B belong to the short trip toxins group (which includes diptheria toxin) that enter the cytosol of host cells from an early endosomal compartment. These protein toxins can be
distinguished from the long trip toxins (which includes cholera toxin) that travel retrograde from endosomes to the Golgi and from there to the ER, where they eventually enter the cytosol [27]. The low pH of early endosomes induces conformational changes of the clostridial toxins, resulting in exposure of a hydrophobic region of the protein toxins, enabling membrane insertion [28]. It has been suggested that residues 956–1128 of toxin B are part of the hydrophobic region, which is involved in membrane insertion [13]. A short-term decrease in extracellular pH of the medium of cell cultures mimics the low pH of endosomes and enables the toxins to enter host cells directly through the cell membrane [26,28]. Membrane insertion is paralleled by formation of pores. This can be shown by 86Rb+ ion release from 86Rb+-loaded host cells when the pH of the cell medium is reduced to pH 5 [26]. Toxin B mutants and N-terminal truncated toxins consisting of the C terminus and the middle part of the protein (including the hydrophobic residues 956–1128) are sufficient for pore formation [26]. It has been suggested that the C terminus, which is involved in binding, is not essential for pore formation but enhances toxin–membrane interaction. The efficacy of pore formation induced by toxin A is largely cholesterol-dependent. Cholesterol depletion of membranes with methyl-β-cyclodextrin inhibits 86Rb+ efflux and cholesterol repletion reconstituted pore-forming activity of toxin A [29]. Whether pore formation is directly involved in delivery of the toxin into the cytosol remains unclear, but it is thought that the hydrophobic region (residues 956–1128 of toxin B) is part of the delivery domain.

The precise mechanism of the translocation process remains one of the most enigmatic puzzles of the action of these toxins. Recent studies indicate that the toxins must be processed to reach the cytosol. Only the catalytic domain of the toxins, including the N-terminal 543 amino acids, is delivered into the cytosol [30,31]. The search for a host protease possibly involved in toxin processing resulted in the unexpected finding that the clostridial glucosylating toxins are auto-proteolytically processed [32]. Even more surprising, inositol hexakisphosphate (InsP6, phytic acid) was found to be an essential factor for activation of proteolysis [32], however its functional role is not clear. Polyphosphorylated inositol, a common inositol metabolite in mammalian cells, is highly charged and has diverse biological functions, including mRNA traffic king, binding to clathrin-assembly protein AP-2, inhibition of protein phosphatases and stimulation of protein kinases [33,34]. InsP6 might be involved in stabilization of toxin protein conformation, which is essential for protease activity and/or proper cleavage.

Recently it was suggested that toxin B possesses aspartate protease activity, which is responsible for toxin
cleavage because the aspartate protease inhibitor EPNP (1,2-epoxy-3-p-nitrophenoxypropane) blocked the processing of toxin B and labeled aspartate 1665 as part of a short Asp-Xaa-Gly (DXG) motif observed in many aspartate proteases [32]. Another hypothesis describes a cysteine protease activity as being responsible for the processing of clostridial glucosylating toxins. There are several lines of evidence supporting this hypothesis. First, as observed frequently for cysteine proteases, dithiothreitol activates the auto-catalytic cleavage of the toxin. Second, the autocatalytic activity is blocked by N-ethylmaleimide, a typical inhibitor of cysteine proteases. Third, the clostridial glucosylating toxins share sequence similarity with a novel, recently identified autocatalytic cysteine protease domain within the structure of the RTX (repeats in toxins) toxin of Vibrio cholerae. Several essential catalytic residues, including the putative catalytic triad D587, H653 and C698, are conserved [35,36] (Figure 1). According to this model, a cysteine protease cutting domain, which is adjacent to the glucosyltransferase domain, is responsible for processing of the toxin.

The N-terminal enzyme domain

The N terminus harbors the glucosyltransferase activity of the toxins and is the biological activity domain [37]. The 543 amino acid residues, which are delivered into the cytosol of host cells, form the glucosyltransferase domain. Recently, the 3D-structure of the catalytic domain (residues 1–543) of toxin B has been solved [38]. The catalytic core is formed by a mixed $\alpha/\beta$-fold with mostly parallel $\beta$-strands (Figure 3a). The overall structure of the catalytic core is similar to other bacterial glycosyltransferases, such as Neisseria meningitidis galactosyltransferase LgtC (lipooligosaccharide glycosyl transferase C) [39] and Bacillus subtilis glycosyltransferase SpsA [40], but also to eukaryotic glycosyltransferases, such as bovine galactosyltransferase $\alpha$GalT [41]. All of these transferases belong to the glycosyltransferase A (GT-A) family [42]. Mainly based on the sequence homology described here, the family of clostridial glucosylating toxins has been designated glycosyltransferase family 44 as defined by Henrissat et al. (http://www.cazy.org). This family also includes genes from Escherichia coli and Chlamydia trachomatis coding for putative glycosyltransferases.

Figure 3. The glucosyltransferase domain of Clostridium difficile toxin B. (a) Structural model of the glycosyltransferase domain of C. difficile toxin B, where UDP-glucose (blue) is attached to the catalytic cleft via $\text{Mn}^{2+}$ (brown) and to the residues of the DXD motif (ball-and-stick model); water molecules are shown as small blue spheres. The GT-A type fold catalytic core is presented in white and the additional subdomains are shown in brown. The putative “flexible loop” region is shown in red with Trp520 hydrogen bonding to the $\beta$-phosphate of UDP-glucose. (b) The catalytic cleft of toxin B, as in (a) but rotated 45°. Several amino acids involved in co-substrate binding are shown. (c) Schematic representation of the catalytic cleft shown in (b). Images were created using PyMOL (www.pymol.org).
The catalytic core of toxin B consists of 234 residues plus >300 additional residues. These additional residues are mainly helices. The four N-terminal helices are particularly prominent and seem to be an independent subdomain, possibly involved in membrane interaction. Mesmin and coworker showed for lethal toxin from *C. sordellii* (the closest homolog of toxin B) that the first 18 amino acids (correlating to the first helix of the N-terminal subdomain) are crucial for lipid bilayer attachment [43]. Also, truncations beyond amino acid Lys65 [44] lead to inactivation of the glucosyltransferase and glucohydrolase activity, demonstrating the importance of the N-terminal helical bundle for catalytic activity. Recently, it was shown that domain C1 of *Pasteurella multocida* toxin (PMT) is structurally similar (41% similarity with toxin B) to this subdomain of toxin B. A role in membrane interaction has also been proposed for PMT [45].

The crystal structure of the catalytic domain of toxin B revealed a set of essential amino acid residues involved in glucosyltransferase reaction or in substrate binding. The Asp-Xaa-Asp (DXD) motif is characteristic of GT-A family members (Asp286 and Asp288 in toxin B) [46,47]. This motif is involved in Mn2+, UDP and glucose binding. Whereas Asp288 complexes directly with Mn2+, interaction of Asp286 with Mn2+ occurs via a water molecule. Asp286 also interacts with the 3’ hydroxyl group of the UDP-ribose and with the 3’ hydroxyl group of glucose, so it is of major importance for proper positioning of UDP-glucose in the catalytic cleft of the enzyme. Trp102, which is also essential for enzyme activity, stabilizes the uracil ring of UDP by aromatic stacking. In addition to the DXD motif and Trp102, Asp270, Arg273, Tyr284, Asn384 and Trp520 were identified by alanine scanning as being essential for enzyme activity [48]. Asp270 and Arg273 are conserved in several GT-A type glycosyltransferases and form with the help of Asp286, a highly defined hydrogen-bond network for proper adjustment of the glucose moiety of the donor substrate [41,49]. A common structural feature of all GT-A type glycosyltransferase is the ‘flexible loop’, which switches from an open, disordered conformation (apo-enzyme) to a closed, ordered conformation on UDP-sugar binding. Co-substrate binding creates a deep pocket that serves as a binding site for the acceptor substrate [50] (Figure 3a). The deep burial of the UDP-sugar could prevent water molecules from acting as acceptors and hydrolyzing the nucleotide sugar. It was proposed that the ‘flexible loop’ helped release the reaction products, leading to suitable turnover of the enzyme [42,51]. Trp520 is well conserved in all of the clostridial glucosylating toxins and resides on the putative ‘flexible loop’ formed by amino acids 510 to 523 and interacts with the scissile bond of the donor substrate (β-phosphate oxygen of UDP). It was shown for other glycosyltransferases that the corresponding tryptophan residue swings out thereby opening the conformation while releasing UDP [52,53].

Glycosyltransferases are strictly stereospecific enzymes. The known clostridial glucosylating toxins are retaining glycosyltransferases, because the sugar is attached to the target protein in the same α-anomeric configuration as in the substrate UDP-glucose [54,55]. A S_Ni (internal return) mechanism has been suggested for the transferase reaction. This mechanism is characterized by an oxocarbenium-like transition state [38].

The clostridial glucosylating toxins exhibit high co-substrate specificity. Whereas toxin A and B and also the related lethal toxin from *C. sordellii* use UDP-glucose as a donor substrate, the co-substrate of α-toxin from *Clostridium novyi* is UDP-N-acetylglucosamine (UDP-GlcNAc) [56]. Biochemical studies revealed that Ile383 and Glu385 are responsible for co-substrate specificity. These residues limit the space of the catalytic cleft for binding of the co-substrate in toxin B. Exchange of these residues changes the co-substrate specificities of the toxins. Thus, the toxin B double mutant Ile383Ser, Gln385Ala accepts UDP-GlcNAc as a co-substrate [57] and changing Ser385 and Ala387 to Ile and Glu, respectively, turns α-toxin into a UDP-glucose-accepting transferase.

**Interaction of toxin B with Rho GTPases**

Glycosyltransferases are region-selective enzymes. In the case of the clostridial glucosylating toxins a specific threonine residue in the substrate proteins (small GTP-binding proteins) is monoglucosylated [54,58]. This threonine senses the integrity of the nucleotide GTP, which is bound to the GTPase in the active conformation and to the hydrolyzed GDP form in the inactive conformation. Sensing occurs via binding to a Mg2+ ion, which is also complexed with the phosphates of the nucleotide. Glucosylation prevents the sensor function and thereby prevents the fundamental conformational switch of the GTP-binding proteins necessary for interaction with diverse effectors or regulator proteins [59] (Box 1).

Substrate specificity of the GTP-binding proteins for the toxins has not been structurally defined to date. Nevertheless, distinct amino acids or regions on Rho GTPases have been shown to have a role in defining specificity. The substrate properties of RhoA versus RhoD are defined by the residues Ser73 and Phe85, respectively, located on the switch II region [60]. The differential recognition of Rac and RhoA by lethal toxin, toxin B1470 and toxin B8864, is attributed to the N-terminal region around amino acids 22–27 of the GTPases where Lys27 of Rho has a major role [61].

Earlier studies showed that the C-terminal part of the catalytic domain of the toxins (residues 364–516) confers substrate specificity [62]. Recent data indicate that amino acids Arg455, Asp461, Lys463 and Glu472, and residues of helix α17 (e.g. Glu449) of toxin B are essential for enzyme–protein substrate recognition [48]. Introduction of helix α17 of toxin B into *C. sordellii* lethal toxin prevents the modification of Ras subfamily proteins but enables glucosylation of RhoA. Crystallographic and biochemical results led to a docking model in which the GTPase consensus binding region (suggested for effector or regulator binding [63]) interacts in a similar manner with the glucosyltransferase toxins (Figure 4). In this model, the membrane-associated regions of the GTPase and the region of the toxin, which is supposed to interact with the membrane, are located on the same side [48].
Box 1. Functional consequences of glucosylation of Rho GTPases by clostridial toxins

The clostridial glucosylating toxins modify Rho proteins (Figure I). These are low molecular mass GTPases, which control numerous signaling pathways [64,65] (Figure Ia). Approximately 20 Rho GTPases have been described, including RhoA, Rac1 and Cdc42, the best studied targets of the toxins. Rho GTPases are inactive in the GDP-bound form and associated with guanine nucleotide dissociation inhibitors (GDI), which keep the GTPases in the cytosol. Guanine nucleotide exchange factors (GEFs) activate Rho GTPases. This enables interaction with different effectors to control numerous signaling processes. The active state of Rho GTPases is terminated by hydrolysis of bound GTP facilitated by GTPase-activating proteins (GAP). Rho proteins regulate the actin cytoskeleton, enzyme activation (e.g. protein kinases, phospholipases), cell polarity, gene transcription and cell proliferation (for review see [65]). However, considering host-pathogen interactions and immune cell activities, it is of note that Rho proteins are essential for epithelial barrier functions, immune cell migration, adhesion, phagocytosis, superoxide production, cytokine secretion and immune cell signaling. The clostridial glucosylating toxins modify Rho GTPases at threonine 35 or 37, which is located in the switch-I region of the GTPases [58,66] (Figure Ib). The toxin-catalyzed glucosylation of Rho proteins results in inhibition of effector coupling and subsequent blocking of signal transduction pathways [67]. It also blocks nucleotide exchange by GEFs [67] and inhibits intrinsic and GAP-stimulated GTPase activity [67]. Glucosylated Rho is no longer able to interact with GDI and is therefore found at the plasma membrane [68]. However, crystallographic [54] and NMR [55] data obtained from the Rho-related Ras protein, which has been glucosylated by the lethal toxin of C. sordellii, indicate that glucosylation prevents formation of the active ‘GTP conformation’ of the GTPases, whereas binding to GTP itself is possible. Surprisingly, the glucosylating toxins cause up-regulation of Rho B, which is an immediate-early gene product [69]. Some Rho B seems to be able to escape modification by toxins A or B and might have important signaling functions [70].

Figure I. Regulation and signaling of Rho proteins (a) and functional consequences of the modification of Rho proteins by clostridial glucosylating toxins (b).
Concluding remarks

Studies from recent years have yielded important progress in our knowledge of the structure and mode of action of clostridial glucosylating toxins. These results indicate the need for a revision of the current AB-model of clostridial glucosylating toxins. Characterization of clostridial glucosylating toxins according to this model seems to be a scientific Procrustes’ bed. Thus, we suggest that an ABCD model may be more appropriate for describing the structure and function of these toxins. However, clostridial toxins are large and their full 3D-structure remains unknown, we also have yet to understand the precise functions of all toxin domains. This model will probably need to be revised again in the future. Nevertheless, recent results on the structural analysis of the toxins have increased our knowledge about this medically important toxin family, and also offer multiple options for new therapeutic strategies. For example, scavenging the toxins using compounds interacting specifically with the receptor-binding domain is a feasible approach [19]. Another option is the development of specific inhibitors of the autocatalytic cleavage or the search for compounds that specifically inhibit the glucosyltransferase activity of these toxins. To this end, solving the 3D-structure of other members of the family of glucosylating toxins in addition to the crystal structure of the enzyme–substrate–cosubstrate complex is highly desirable.

References

3 Pepin, J. et al. (2005) Mortality attributable to nosocomial Clostridium difficile-associated disease during an epidemic caused by a hypervirulent strain in Quebec. CMAJ 173, 1037–1042
8 Chaves-Olarte, E. et al. (1999) A novel cytotoxin from Clostridium difficile serogroup F is a functional hybrid between two other large clostridial cytotoxins. J. Biol. Chem. 274, 11046–11052


Larsen, R.D. et al. (1990) Frameshift and nonsense mutations in a human genomic sequence homologous to a murine UDP-Gal:β1,4Gal:1,4→GlcNAc d1,3-galactosyltransferase cDNA. J. Biol. Chem. 265, 7055–7061


Barth, H. et al. (2001) Low pH-induced formation of ion channels by Clostridium difficile toxin B in target cells. J. Biol. Chem. 276, 10670–10676


Sehr, P. et al. (1998) Glucosylation and ADP-ribosylation of Rho proteins - effects on nucleotide binding, GTPase activity, and effector-coupling. Biochemistry 37, 5296–5305

