


Article

# An In-Silico Sequence-Structure-Function Analysis of the N-Terminal Lobe in CT Group Bacterial ADP-Ribosyltransferase Toxins

Miguel R. Lugo and A. Rod Merrill \* 

Department of Molecular and Cellular Biology, University of Guelph, Guelph N1G 2W1, Canada; mlugoa@gmail.com

\* Correspondence: rmerrill@uoguelph.ca; Tel.: +1-519-824-4120 (ext. 53806)

Received: 30 May 2019; Accepted: 15 June 2019; Published: 21 June 2019



**Abstract:** The C3-like toxins are single-domain proteins that represent a minimal mono-ADP-ribosyl transferase (mART) enzyme with a simple model scaffold for the entire cholera toxin (CT)-group. These proteins possess a single (A-domain) that modifies Rho proteins. In contrast, C2-like toxins require a binding/translocation partner (B-component) for intoxication. These are A-only toxins that contain the E-x-E motif, modify G-actin, but are two-domains with a C-domain possessing enzymatic activity. The N-domain of the C2-like toxins is unstructured, and its function is currently unknown. A sequence-structure-function comparison was performed on the N-terminal region of the mART domain of the enzymatic component of the CT toxin group in the CATCH fold (3.90.210.10). Special consideration was given to the N-domain distal segment, the  $\alpha$ -lobe ( $\alpha_1$ – $\alpha_4$ ), and its different roles in these toxin sub-groups. These results show that the role of the N-terminal  $\alpha$ -lobe is to provide a suitable configuration (i) of the  $\alpha_2$ – $\alpha_3$  helices to feature the  $\alpha_3$ -motif that has a role in NAD<sup>+</sup> substrate binding and possibly in the interaction with the protein target; (ii) the  $\alpha_3$ – $\alpha_4$  helices to provide the  $\alpha_{3/4}$ -loop with protein-protein interaction capability; and (iii) the  $\alpha_1$ -N<sub>tail</sub> that features specialized motif(s) according to the toxin type (A-only or A-B toxins) exhibiting an effect on the catalytic activity via the ARTT-loop, with a role in the inter-domain stability, and with a function in the binding and/or translocation steps during the internalization process.

**Keywords:** mono-ADP-ribosylation toxins; C3-like toxins; C2-like toxins; ADP-ribosylation; CT-toxins; target substrate motifs; N-terminal  $\alpha$ -lobe

**Key Contribution:** The role of the  $\alpha$ -lobe in CT-group bacterial mono-ADP-ribosyltransferase toxins is to provide a scaffold for (i) the  $\alpha_2$ – $\alpha_3$  helices that feature the  $\alpha_3$ -motif involved in NAD<sup>+</sup> substrate binding and protein target interaction; (ii) the  $\alpha_3$ – $\alpha_4$  helices that provide the  $\alpha_{3/4}$ -loop with protein–protein interaction capability; and (iii) the  $\alpha_1$ -N<sub>tail</sub> in the T-segment that features specialized motif(s) according to the toxin type (A-only or A-B toxins). These elements participate in the catalytic activity via the ARTT-loop; inter-domain stability; and play a role in the binding and/or translocation steps during the host cell intoxication process.

## 1. Introduction

Bacterial mono ADP-ribosyl transferase toxins (mART toxins) belong to a family of toxins that catalyzes the covalent transfer of an ADP-ribose moiety from NAD<sup>+</sup> to a protein or DNA target in a host cell, changing target activity and impairing target cell function and survival [1]. This toxin family includes diverse members such as exotoxin A (ExoA) from *Pseudomonas aeruginosa*, pertussis toxin (PT) from *Bordetella pertussis*, and cholera toxin (CT) from *Vibrio cholerae*. Members of this family are

steadily increasing and are classified as either CT-group (after “Cholera toxin”) or DT-group (after “Diphtheria toxin”).

All bacterial mART toxins share a common structural fold composed of ~100 residues with low sequence homology formed by a core scaffold of two perpendicular  $\beta$ -sheets, flanked by variable helical sub-structures and exposed-loops that constitute a cleft for the binding of the NAD<sup>+</sup> substrate. The SCOP2 database [2] assigns mART toxins (SCOP2: 56400) to the  $\alpha + \beta$  class of proteins (SCOP2: 53931) and to the “unusual” ADP-ribosylation fold (SCOP2: 56398). This fold is shared with eukaryotic mART proteins, Ecto-ARTs (SCOP2: 82814), and with the C-terminal domain of poly (ADP-ribose) polymerases known as PARPs (SCOP2: 56398), among others ADP-ribosylating proteins (for a review see Fieldhouse, 2008 [3]).

The CT-group of mART toxins is a large group of A-only toxins (containing the catalytic component only, e.g., C3bot1, Vis, ExoS) [4–6], and binary A-B toxins (A: catalytic, B: cell binding/translocation components in distinct proteins, e.g., *iota*, and CT) [7,8] and A/B (A and B components in the same polypeptide, e.g., Certhrax [9–11] with low sequence identity but containing three conserved regions with the consensus (Y/F)R (Underline refers to conserved residues) (Region 1), (Y/F)xSTS (Region 2), and the catalytic (Q/E)xE (Region 3) motifs, with either NAD<sup>+</sup>-binding or catalytic function (for reviews see Fieldhouse, 2010 [12]). The CT-group of toxins is classically divided into ExoS-like, C2-like, C3-like, and CT/PT-like subgroups of toxins, based mainly on the catalytic motif (Region 3), subunit organization, and macromolecule targets.

Exoenzyme S (ExoS) (UP: Q51448) (UP: UniProtKB accession number) from *P. aeruginosa* is an A-only, four-domain, bi-functional toxin/effector with a C-terminal mART domain; it requires factor-activating ExoS (FAS) for activation [13]. Three additional ExoS-like toxins have been identified: ExoT toxin (UP: Q9I788) from *P. aeruginosa* [6], the single-domain VopT toxin (UP: Q87G19) from *V. parahaemolyticus* [14], and AexT toxin (UP: Q93Q17) from *Aeromonas* sp. [15]. ExoS-like toxins are secreted into target cells via the type III secretion system (T3SS), possess the ExE catalytic motif and target the RAS family of G-proteins (except for AexT which targets G-actin) (for a review see Barbieri, 2004 [6]).

C2 is a mART toxin (UP: D4N871) from *Clostridium botulinum* [16] and is an A-B multi-unit exotoxin whereas its C2I mART subunit (A-component) requires the C2II receptor-binding subunit (B-component) to facilitate cell entry of the catalytic A-subunit. The A-component is a bi-domain protein with both halves structurally, but not functionally equivalent. The C-terminal domain harbors the mART activity, while the N-terminal domain interacts with the B-component. The C2-like subgroup includes members such as *iota* toxin (UP: Q46220, Ia subunit) from *Clostridium perfringens* type E [17], CST toxin (UP: O06497, SA subunit) from *Clostridium spiroforme* [18], VIP toxin (UP: G8C882, VIP2 subunit) from *Bacillus cereus* [19], CDT toxin (UP: Q9KH42, CDTa subunit) from *Clostridium difficile* [20]; and the newly discovered CPILE toxin (UP: X512D7, CPILE-a subunit) from *Clostridium perfringens* W5052 strain [21]. C2-like toxins contain the catalytic ExE motif, form an AB<sub>7</sub> oligomer to gain cell access, and ADP-ribosylate G-actin (for reviews see Tsuge 2017 [22]).

The C3 exotoxin (UP: P15879) from *C. botulinum* [23] (C3bot1) is a single-domain (A-only) mART toxin. Eight other members of C3-like toxins have been described so far: C3bot2 (UP: Q00901) from *C. botulinum* [24], C3lim (UP: Q46134) from *C. limosum* [25], C3cer (UP: Q8KKNY0) from *B. cereus* [26], C3larvin (UP: W2E3J5) and Plx2A (UP: M9V3B7) from *Paenibacillus larvae* [27,28], and three isoforms C3stau1 (UP: P24121), C3stau2 (UP: Q8GAX6), and C3stau3 (UP: Q8VVU2) produced by *Staphylococcus aureus* [29,30]. C3-like toxins harbor the catalytic QxE signature and modifies RhoA, B, and C, among others GTPases (for reviews see Vogelsang, 2007 [31] and Tsuge, 2017 [22]).

The CT-like subgroup (after cholera/pertussis toxins) is a more heterogeneous group in terms of the unit organization, catalytic motif, molecular target, and modified residues. This group includes CT (UP: P01555, A-subunit) and Chelt toxins (UP: A2PU44, A-subunit) from *V. cholera* [12,32], PT toxin (UP: P04977, A-subunit) from *Bacillus. pertussis* [33], human LT (UP: P43530, A-subunit) from *Eshcherichia coli* [34], Scabin toxin (UP: C9Z6T8) from *S. scabies* [35], among others.

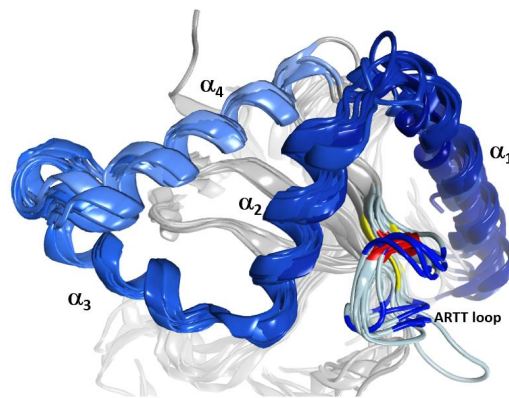
However, there are other mART toxins that do not clearly match with any of the previous classification schemes for the CT group. In these “ungrouped” toxins are the following: SpyA toxin (UP: Q1J858/ Q1JN57) from *Streptococcus pyogenes* [36], which is structurally similar to C3-like toxins (A-only) and with a secretion signal-sequence, but harbors the ExE motif and ADP-ribosylates vimentin among other substrates [37], and Plx2A toxin (UP: M9V3B7) from *P. larvae* [28]. These proteins are related to C3-like toxins (single-domain A-component with QxE catalytic motif that ADP-ribosylate Rho proteins) but also show features of C2-like toxins (require a binding/translocation partner, B-component) for intoxication. SpvB toxin (UP: P555220) from *Salmonella enterica* [38], VahC toxin (UP: Q49TP5) from *Aeromonas hydrophila* [39], and Photox toxin (UP: Q7N9B1) from *Photorhabdus luminescens* [40] possess the ExE motif, ADP-ribosylate G-actin, and are two-domains toxins (C-domain has mART activity). AexU toxin (UP: A0FKE5) from *A. hydrophila* [41] has a domain organization like AexT and is secreted into target cells by using the type III secretion system and harbors the catalytic QxE motif rather than the ExE motif characteristic of ExoS- like toxins. Certhrax toxin (UP: Q4MV79) from *B. cereus* strain G9241 [9] is a two-domain A/B toxin with its catalytically C-domain homologous to C3-like toxins (QxE motif), but its N-domain is homologous to the PA-binding domain of the anthrax lethal factor from *Bacillus anthracis*. Vis toxin (UP: A3UNN4) from *Vibrio splendidus* strain 12B01 [5] is a single-domain A-only toxin homologous to C3-toxins with an N-terminal secretion signal peptide. Vis harbors the ExE motif, although Vis does not covalently modify conventional C2-like nor ExoS-like substrates. Mav toxin (UP: AOQLI5) from *Mycobacterium avium* strain 104 [12] is a tetra-domain A/B toxin with the mART activity in the C-terminal domain that contains the ExE catalytic motif. EFV toxin (UP: Q838U8) from *Enterococcus faecalis* strain V8583 [12] is a bi-domain protein with the mART activity in the C-terminal domain and an ExE catalytic motif that targets actin (unpublished results).

The CATH database [42] assigns the closely related codes 3.90.176.10 for C2-, C3-, and ExoS- like subgroups, and 3.90.210.10 for the CT/PT-like subgroup of CT-group of mART toxins. The enzymatic component of the CT-group (CATCH fold, 3.90.176.10) at the N-terminal region, i.e., upstream of the strand  $\beta_1$  of the  $\beta$ -core scaffold, is well structured with a similar fold, albeit not much attention has been given to this section. In the present work, we performed a sequence-structure-function comparison of the N-terminal region of CT-group mART toxins in the CATCH 3.90.176.10 category (non-CT/PT-like toxins), with special emphasis on the N-most distal segment (a putative helix-coil structure) and its different roles in these toxins.

## 2. Results and Discussion

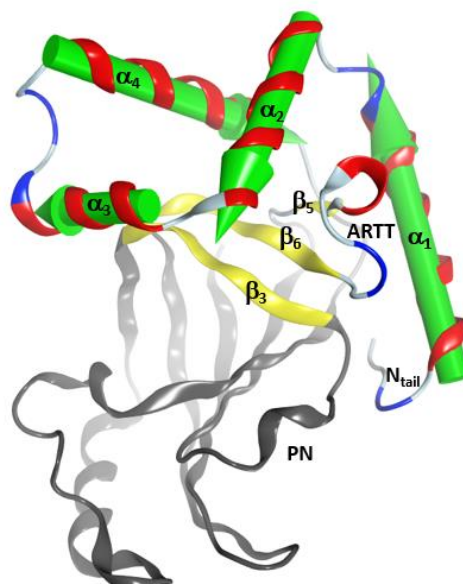
### 2.1. The Overall Structure of the N-Terminal $\alpha$ -Lobe

The C3-like toxins are single-domain proteins that represent a minimal mART enzyme with a simple scaffold that serves as the model for the entire CT-group. The N terminus is a helical region formed by four consecutive  $\alpha$ -helices,  $\alpha_1$ – $\alpha_4$  (Figure 1). This region shares a lower average similarity for those residues in structured elements (~68%, from 52% to 88% pairwise) than with the set of  $\beta$ -strand residues that form the  $\beta$ -core superstructure (~73%, from 59% to 99% pairwise). Structurally, this N-terminal region consists of a conserved fold configured within a compact lobe, the “ $\alpha$ -lobe”, with a high degree of overlap among the C3-like toxins (average  $\text{RMSD}_\alpha = 2.4 \text{ \AA}$ ; Figure 1).



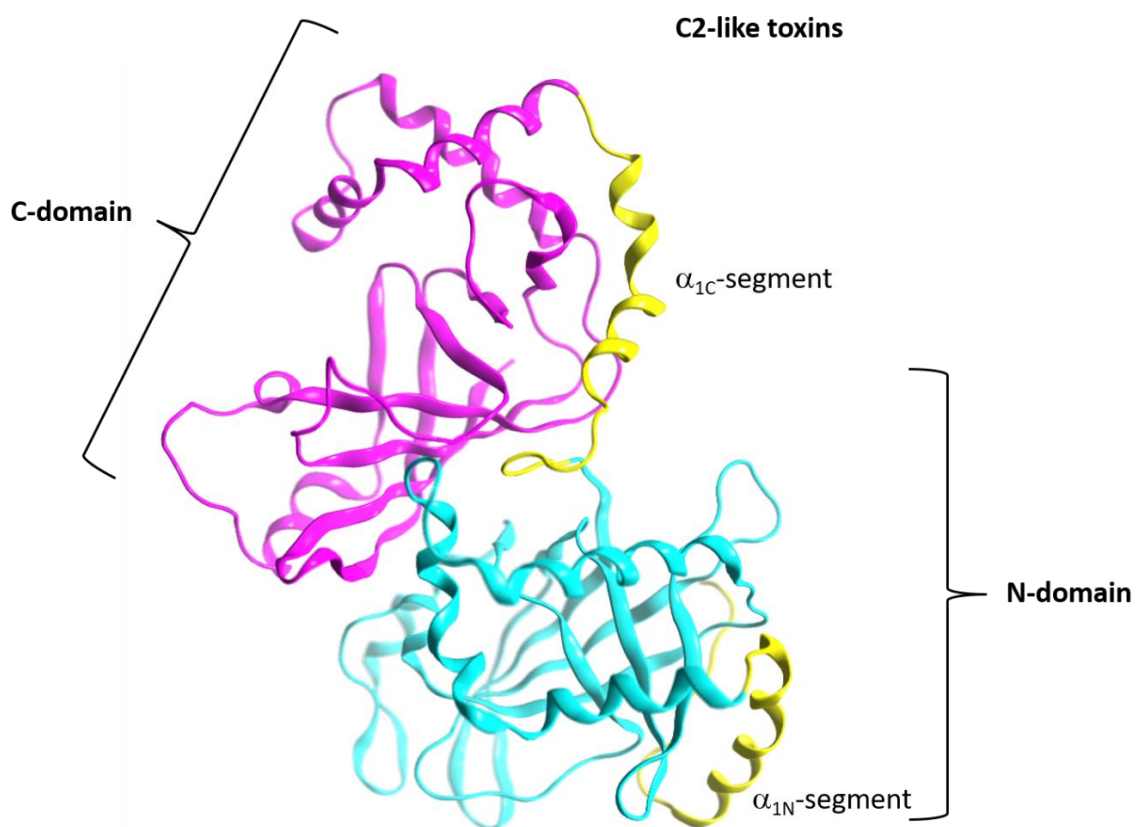
**Figure 1.** The  $\alpha$ -lobe in C3-like toxins. Superposition of six C3-like toxins showing a compact N-terminal substructure that encompasses four  $\alpha$ -helices. The ARTT loop (ADP-ribosyl-turn-turn) is a catalytic region that is responsible for the target substrate recognition and is shown at the bottom right of the overlaid structures. The superposed structures include C3 exotoxin (PDB:1G24) and C3bot2 (PDB:1R45) from *C. botulinum*, C3lim (PDB:3BW8) from *C. limosum*, C3cer (PDB:4XSG) from *B. cereus*, C3larvin (PDB:4TR5) from *P. larvae*, and C3stau2 (PDB:1OJQ) produced by *S. aureus*.

This  $\alpha$ -lobe is packed following a V–L– $\alpha\alpha$ -corner topology (Figure 2), which encloses the  $\beta_{II}$  sheet and its connecting loops/turns according to: (i) an open V-shaped  $\alpha_1$ – $\alpha_2$  superstructure that surrounds the catalytic ARTT loop followed by (ii) an L-shaped  $\alpha_2$ – $\alpha_3$  superstructure that exposes the  $\alpha_{2/3}$ -turn and the key  $\alpha_3$ -motif (defined later) to the  $\text{NAD}^+$ -binding pocket and to the protein target; and finally, (iii) an  $\alpha_3$ – $\alpha_4$ -corner superstructure. The loop that links helices  $\alpha_3$  and  $\alpha_4$ , the “ $\alpha_{3/4}$ -loop”, offers the flexibility required to form an  $\alpha\alpha$ -corner superstructure with a longer connection by turning the inertial axis of helix  $\alpha_4$  orthogonal to both the  $\alpha_3$  and the plane formed by the  $\alpha_2$  and  $\alpha_3$  axes. As a result,  $\alpha_4$  is in a transverse orientation in relation to the  $\beta_{II}$  sheet, increasing the contact surface between the  $\alpha$ -lobe and the  $\beta_{II}$  sheet.



**Figure 2.** The  $\alpha$ -lobe topology in C3-like toxins. Depiction of the V–L– $\alpha\alpha$ -corner topology of the  $\alpha$ -lobe. The green arrows correspond to the inertial axes of the  $\alpha$ -helices. The  $\beta_{II}$  sheet is shown in yellow ribbons with  $\beta_3$ ,  $\beta_5$ , and  $\beta_6$  strongly interacting with the elements of the  $\alpha\alpha$  corner. The PN (phosphate-nicotinamide) loop is an  $\text{NAD}^+$  substrate-binding loop and it is indicated along with the N-terminus (N-terminal tail) of the domain.

In addition, C3-like toxins are characterized by an unstructured N-terminal segment, the “N<sub>tail</sub>”, of variable length (Figure 2). Furthermore, the C3-like  $\alpha$ -lobe is a well-structured, globular motif and is topologically identical or similar to: (i) the N-terminal helical region of the single-domain enzymatic component of the A-B binary Plx2 and Larvin toxins (Plx2A, PDB:5URP); (ii) the N-terminal helical region of the catalytic C-domain, “C2<sub>C</sub>-domain” of the A-component of the binary C2 toxin (Figure 3), C2I (PDB: 2J3Z), and C2-like toxins such as Vip2 (PDB: 1QS1), *iota* Ia (PDB: 1GIQ), CdtA (PDB: 2WN4), and SA component of the CST toxin (HM<sup>2</sup>). In these toxins, the N-terminal region of the non-catalytic N-domain, “C2<sub>N</sub>-domain”, interacts with the binding/translocation component (see later), shows a similar topology and superposes well with the C3-like  $\alpha$ -lobes; it also possesses an additional well-defined helix ( $\alpha_{4a}$ ) connecting  $\alpha_3$  and  $\alpha_4$ ; (iii) the helical region of the Certhrax toxin C-terminal domain (PDB: 4GF1)—Certhrax toxin has a significantly longer  $\alpha_{3/4}$ -loop than most C3- and C2-like toxins; and (iv) the helical region of the VahC C-terminal domain (PDB: 4FML) and SpvB (PDB: 2GWM). The first two crystallographically ‘solved’ helices in these toxins superpose well with the  $\alpha_2$  and  $\alpha_3$  helices of the  $\alpha$ -lobe. There is also a 20-residue long helix-loop insertion between the second (equivalent to  $\alpha_3$ ) and third (equivalent to  $\alpha_4$ ) solved helices in these toxins. Additionally, the mART domain of Photox toxin reveals a high sequence homology with VahC and SpvB toxins; consequently, homology models of Photox show the canonical  $\alpha_2$ – $\alpha_4$  topology; and finally, (v) the N-terminal segment of Vis toxin (PDB: 4XZJ) which has a longer  $\alpha_1$  helix with a slightly different orientation.

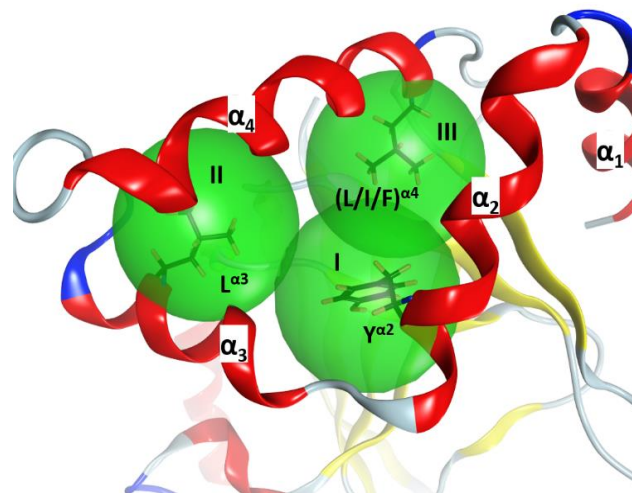


**Figure 3.** Topology of the A-component of the C2-like toxins. Structure of C2I toxin (PDB: 2J3Z) as a representative member of the C2-like class, showing the “adaptor” N domain (in cyan ribbons) and the “catalytic” C domain (in fuchsia ribbons). The T-segment of each domain is colored in yellow ribbons.

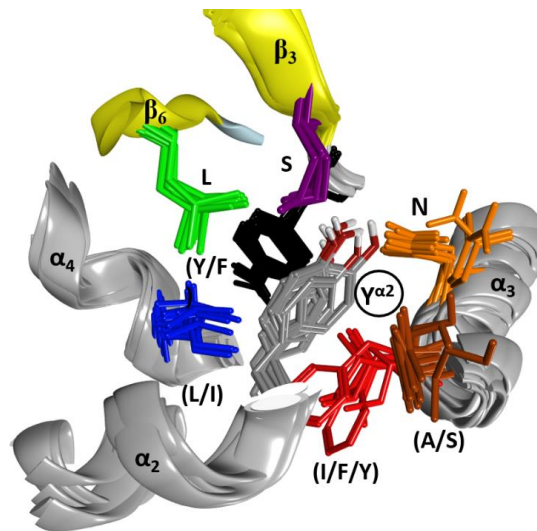
There is no X-ray structure for any member of the ExoS-like group. However, Sun et al. reported homology models of the N-terminal mART domains of ExoS and ExoT toxins, based on Ia toxin as a template; these models superpose well with the C3-like  $\alpha$ -lobe [43]. On the other hand, for the enzymatic component of the CT/PT-subgroup (e.g., CT, LT-A/IIB, PT, and Scabin toxins, among others) there is no structural equivalent to the C3-like  $\alpha$ -lobe. Instead, certain coils, helices, and strands (and their connecting loops) of the  $\beta_{II}$  sheet occupy the same location within the  $\alpha$ -lobe without significant overlap with other elements—only the backbone structure of the  $\alpha_3$ -motif is roughly traced by an active-site helix or loop.

## 2.2. Stability of the $\alpha_2$ – $\alpha_4$ Superstructure

In C3-like toxins, the  $\alpha_2$ – $\alpha_4$  superstructure of the  $\alpha$ -lobe is clustered and is in contact with the rest of the protein by a network of hydrophobic interactions centered in three key residues: Tyr in  $\alpha_2$  (Tyr $^{\alpha_2}$ ), Leu in  $\alpha_3$  (Leu $^{\alpha_3}$ ), and Leu/Ile/Phe in  $\alpha_4$ —one in each helix (Figure 4). In addition, an aromatic residue is present in the  $\beta_{2/3}$ -turn, (Tyr/Phe) $^{\beta_{2/3}}$  (not shown). Effectively, Tyr $^{\alpha_2}$  is the center of a cluster of interactions, “cluster<sub>I</sub>”, that cements the whole  $\alpha_2$ – $\alpha_4$  segment with the rest of the structure (Figure 5). The structural relevance of this hydrophobic and polar cluster of residues is evidenced by the large number of conserved and similar residues (including an invariant Leu in  $\beta_6$ , see Figure 5) and the significant overlap of their side-chains. The reduced mobility of these residues as reported by their crystallographic *B*-factors reveals the mutually imposed constraint within the compact structure of the  $\alpha$ -lobe. Tyr $^{\alpha_2}$  is conserved in Plx2A in the C2<sub>C</sub>-domains, in most of the C2<sub>N</sub>-domains, and even in most of the toxins of the CT-group with known  $\alpha$ -lobe topology. Compatible with its structural role, Tyr $^{\alpha_2}$  is conservatively replaced with a Phe in Vis toxin, and with Leu in C2I<sub>N</sub> and Vip2<sub>N</sub> domains. However, Tyr $^{\alpha_2}$  might also play an “active-site” role as its side-chain hydroxyl bridges the conserved Ser $^{\beta_3}$  (part of the STS motif) with Asn $^{\alpha_3}$  (part of the NLR or  $\alpha_3$ -motif, see later); these are two critical residues involved in NAD<sup>+</sup> substrate binding. The substitution of Tyr $^{\alpha_2}$  with His $^{\alpha_2}$  in HopU1 toxin is still compatible with the suggested role, particularly considering that HopU1 has a longer  $\alpha_{2/3}$ -link that might assist in the binding of NAD<sup>+</sup> [44].

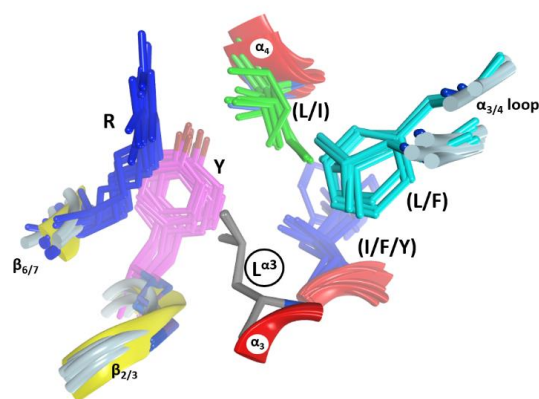


**Figure 4.** The  $\alpha$ -lobe stability.  $\alpha_2$ – $\alpha_4$  superstructure of C3larvin (PDB: 4TR5), showing three hydrophobic clusters (depicted as green spheres) centered at three key residues. The structural “clusters” for the  $\alpha_2$ – $\alpha_4$  superstructure are designated as I, II, and III and the superstructure is framed by three  $\alpha$  helices ( $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ ).



**Figure 5.** Interactions in Cluster<sub>I</sub> of the  $\alpha_2$ – $\alpha_4$  superstructure. The cluster of residues around the conserved Tyr $^{\alpha 2}$  (circled label; grey residue) is shown in this set of six overlaid C3 structures. The side-chains are shown for Tyr/Phe $^{\beta 2,3}$  (black), Ser $^{\beta 3}$  (mauve), Leu $^{\beta 6}$  (green), Leu/Ile $^{\alpha 4}$  (blue), Ile/Phe/Tyr $^{\alpha 3}$  (red), Ala/Ser $^{\alpha 3}$  (brown), and Asn $^{\alpha 3}$  (orange).

In C3-like toxins, the conserved Leu $^{\alpha 3}$  is part of the (NLR) $^{\alpha 3}$  motif and is the center of a second cluster of hydrophobic interactions, “cluster<sub>II</sub>” that links  $\alpha_3$  and  $\alpha_4$  and connects them to other elements (Figure 6). In this cluster, Leu $^{\alpha 3}$  interacts with (Leu/Ile) $^{\alpha 4}$ , with some residues in the  $\alpha_{3/4}$ -loop, and with two other regions that act as hinges between both  $\beta$ -sheets—the conserved Tyr $^{\beta 2/3}$  and the conserved Arg $^{\beta 6/7}$  (part of the (LPR) $^{\beta 6/7}$  motif). This cluster is structurally well defined in C3-like toxins as evidenced by superposing the structures on Leu $^{\alpha 3}$  (grey residue in Figure 6). It is remarkable how identical this configuration (i.e., same side-chain torsion angles) is for these clustered residues. Leu $^{\alpha 3}$  is conserved in Plx2A, the C2-like toxins (C2<sub>C</sub>-domains, but not the C2<sub>N</sub>-domains) and in most of the non-PT-like toxins (except, for example, the TccC3 and TccC5 toxins) [45]. This observation suggests an active role for Leu $^{\alpha 3}$  in the stability of the  $\alpha_3$ -motif, which is relevant for the binding of the NAD $^+$  substrate, rather than serving a pure structural role in the lobe stability.



**Figure 6.** Interactions in Cluster<sub>II</sub>. The cluster of residues around the conserved Leu $^{\alpha 3}$  (circled label) in Cluster<sub>II</sub> is shown in this set of six overlaid C3 structures. The side-chains are shown for Arg $^{\beta 6/7}$  (dark blue), Tyr $^{\beta 2/3}$  (mauve), Leu $^{\alpha 3}$  (grey), Ile/Phe/Tyr $^{\alpha 3}$  (light blue), Leu/Phe $^{\alpha 3,4}$  (cyan) and Ile/Val $^{\alpha 4}$  (green). The residue nomenclature is as follows: the superscript refers to the alpha or beta secondary structure followed by the strand or helix number in the C3bot1 structure (PDB:1G24).

A hydrophobic residue at the N terminus of  $\alpha_4$ , either Leu or Ile, is the center of the “cluster<sub>III</sub>” in C3-toxins and most non-PT-like toxins (not shown). The central (Leu/Ile) $^{\alpha 4}$  contacts the conserved (Tyr/Phe) $^{\beta 2/3}$  (Figure 6) and the conserved Leu $^{\beta 6}$  (Figure 5) in most toxins of the CT-group (an exception is HopU1 toxin) [46]. It is important to highlight (Tyr/Phe) $^{\beta 2/3}$  (Figure 6); this residue welds (coordinates) the three hydrophobic clusters in all the catalytic domains with the  $\alpha$ -lobe configuration, even in the C<sub>2N</sub>-domains, which reaffirms the important structural role of this aromatic motif. Moreover, the presence of (Tyr/Phe) $^{\beta 2/3}$  as part of the (Y/F) $\alpha$ STS motif in the toxins of the CT-group [3,12] reveals a structural role in preserving the N-terminal configuration, regardless of whether this region has the  $\alpha$ -lobe topology in contact with the  $\beta$ -scaffold.

### 2.3. The $\alpha_3$ -Helix and the $\alpha_3$ -Motif

The  $\alpha_3$ -helix is defined in all the C2/C3-like toxins and most of the CT-like toxins with the  $\alpha$ -lobe topology. In C2/C3-like toxins, the  $\alpha_3$  helix harbors the  $\alpha_3$ -motif,  $\underline{Y}^{\alpha 2}-(\underline{IN}-\underline{LR})^{\alpha 3}$  (Figure 5), which includes residues from both the  $\alpha_2$  and  $\alpha_3$  helices. The spatial orientation of the  $\alpha_3$ -motif along the  $\alpha_3$  inertial axis offers a recognition surface aligned with the long axis of the bound pose of the NAD<sup>+</sup> substrate; hence, functionally, the  $\alpha_3$ -motif binds the NAD<sup>+</sup> substrate and contacts the target protein substrate. Thus, the  $\alpha_{2/3}$ -loop (N-terminal end), the conserved Asn $^{\alpha 3}$  (at the center), and the semi-conserved Arg $^{\alpha 3}$  (at the C-end) of the  $\alpha_3$ -motif, all point their side-chains towards the binding cavity and contact the bound NAD<sup>+</sup> and/or the target protein substrate. In effect, in the complexes of C3bot1 (PDB: 2C8F) and C3stau2 (PDB: 1OJZ) with NAD<sup>+</sup>, the NH of Asn $^{\alpha 3}$  H-bonds the NAD<sup>+</sup> A- phosphate, and Arg $^{\alpha 3}$  stacks with the NAD<sup>+</sup> adenine ring. The relevance of these two residues in NAD<sup>+</sup> binding is evidenced by their absence in the non-catalytic C<sub>2N</sub>-domains (see later). However, Asn $^{\alpha 3}$  is absolutely conserved in all the non-PT-like toxins, while Arg $^{\alpha 3}$  is less conserved (e.g., an Ile residue in certain C2-toxins). The role of Tyr $^{\alpha 2}$  and of Leu $^{\alpha 3}$  in this motif was already discussed above.

### 2.4. Role of the N<sub>tail</sub>- $\alpha_1$ Segment in C3- and C2-Like Toxins

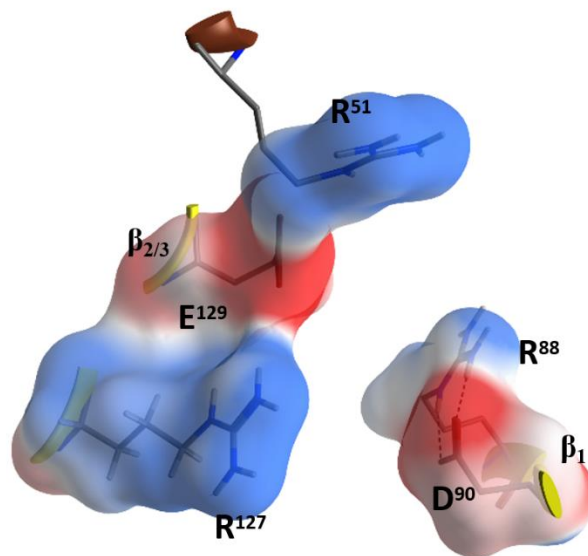
Contrary to the conservation seen in the  $\alpha_2$ - $\alpha_4$  region, the N<sub>tail</sub>- $\alpha_1$  of C3-like toxins is a variable segment. The length of the unstructured N<sub>tail</sub> is variable, being non-existent in C3larvin to 20 residues in C3stau(s) and C3cer. Likewise,  $\alpha_1$  is 15 residues in C3lim, but is only eight residues in both C3cer (which superposes with the N-terminal half of C3lim  $\alpha_1$ ) and C3larvin (which superposes with the C-terminal half of C3lim  $\alpha_1$ ).

The eight known C3-like toxins do not possess a specialized cell binding/translocation component or domain for access/entry into their target host cells, and the details on the mechanism and molecular determinants involved in the toxin internalization await further characterization (for a recent review see Rohrbeck, 2016 [47]). This is relevant for all C3-like toxins except C3stau toxins (C3stau1, C3stau2 and C3stau3), since *S. aureus* infects the host cell and releases the toxins into the cell cytoplasm [29]. Thus, C3bot1, C3bot2, C3lim, C3cer, and C3larvin toxins will be referred to collectively as “C3-etoxins” (with “e” after extracellular).

Due to the relatively high concentrations and long incubation times required for C3-etoxins to enter the target host cells, it was previously suggested that these toxins gain host access by non-specific pinocytosis [25]. Also, C3-etoxins have a basic  $\alpha_1$  helix; basic peptides have been shown to interact with charged phospholipids on the outer layer of the cell membrane of host cells, causing destabilization of the lipid bilayer [48]. In fact, a short “transport” peptide fused to the C terminus of C3bot1 enabled chimeric toxin entry into the cytoplasm by a receptor-independent mechanism [49]. However, C3bot1 and C3lim may be selectively internalized into the cytoplasm of macrophage-like murine cells likely by a specific endocytotic mechanism [50]. Indeed, Rohrbeck and colleagues [51] identified a membrane partner that binds C3bot1, and vimentin (rod domain) was established as the cellular receptor in neuronal and macrophage cell lines. A more recent study showed the 88-RGD-90 sequence in C3bot1 functions as a vimentin binding-motif in neuronal cells [52]. Effectively the RGD motif is unique in C3-etoxins; however, it should be noted that this Arg residue (at  $\beta_1$ ) is an invariant residue in all



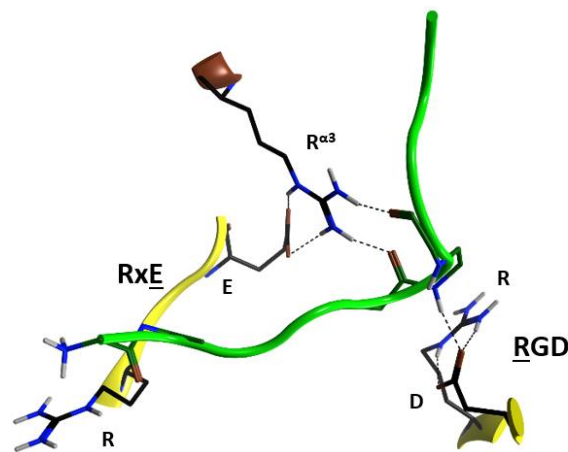
CT-like toxins (the signature “R”) involved in the binding of the  $\text{NAD}^+$  substrate. The conserved Arg residue is buried in the  $\text{NAD}^+$ -binding pocket and is stabilized by strong salt-bridges with the Asp in this motif in the apo structures. The fact that C3stau harbors  $\text{RLL}$ , instead of the  $\text{RGD}^{\beta 1}$  motif, and only weakly intoxicates HT22 cells [53], would suggest that there is a concerted participation of the residues involved. In addition, because of the highly buried nature of the  $\text{RGD}^{\beta 1}$  motif (e.g., Arg88 is only partially exposed to solvent), we envisage that other residue(s) that form the  $\text{NAD}^+$ -binding interacting surface might participate in the recognition motif for vimentin. In this sense, the  $\text{RGD}^{\beta 1}$  motif is close to the conserved  $\text{R}^{\alpha 3}$  of the signature  $\alpha 3$ -motif (Arg51 in C3bot1), and to the  $\text{RxE}$  motif located in the  $\beta_2$  strand (residues 127–129 in C3bot1) of the C3 group, (Figure 7). The  $\text{RGD}^{\beta 1}$  motif along with the  $\text{RxE}^{\beta 2}$  motif and  $\text{R}^{\alpha 3}$  form an ‘electrostatic clamp’ with complementary charges and H-bond capabilities in C3-etoxins (Figure 7).



**Figure 7.** Putative vimentin binding motifs in C3bot1. Depiction of Arg88 and Asp90 of the RGD motif, the Arg127 and Glu129 of the RxE motif, and the conserved Arg51 in the signature  $\alpha 3$  motif.

Interestingly, monomeric (soluble) vimentin is reported to be ADP-ribosylated by SpyA in the head domain of the protein [36,37]. This implies that vimentin must not bind into the  $\text{NAD}^+$ -binding pocket of SpyA in order to behave as a protein substrate for the transfer reaction. Incidentally, SpyA lacks the  $\text{RGD}^{\beta 1}$  motif ( $\text{RYV}^{\beta 1}$  instead), lacks  $\text{R}^{\alpha 3}$  ( $\text{D}^{\alpha 3}$  instead), and lacks the  $\text{RxE}^{\beta 2}$  motif ( $\text{YxK}^{\beta 2}$ ). Additionally, no C2-like toxin exhibits these motifs without also binding vimentin. Plx2A harbors  $\text{RGT}^{\beta 1}$ ,  $\text{R}^{\alpha 3}$  and  $\text{LxE}^{\beta 2}$  and enters mouse macrophages in the absence of the Plx2B protein [28]. Therefore, the  $\text{R}^{\alpha 3}$  and  $\text{E}^{\beta 2}$  residues may play a key role in the binding of the vimentin rod domain.

Notably, the  $\text{RxE}$  motif ( $\text{TxE}$  in C3cer) is only present in the C3-etoxins. Thus, active-site residues,  $\text{RGD}$  (and likely also  $\text{R}^{\alpha 3}$  and  $\text{RxE}^{\beta 2}$ ), while important for interaction with  $\text{NAD}^+$  for catalysis, have the role of binding vimentin on the membrane surface of the host cell. This dual function may be rationalized because these two activities are manifested in different compartments (i.e., the cytoplasm for enzymatic activity and the extracellular space for the vimentin recognition) and in different stages of the intoxication process. Therefore, considering that C3-like toxins correspond to a “minimal” toxin, then Asp90 and Arg127 (in C3bot1) may have evolved to use the invariant Arg51, Glu129, and Arg88 residues to bind a required membrane component for toxin internalization. Validation of the role of these residues in protein–protein interactions lies in the crystal contacts of a C3larvin fragment—of a symmetrically related molecule—docked into the  $\text{NAD}^+$ -binding pocket in the C3larvin crystal structure (PDB: 4TR5) (Figure 8).

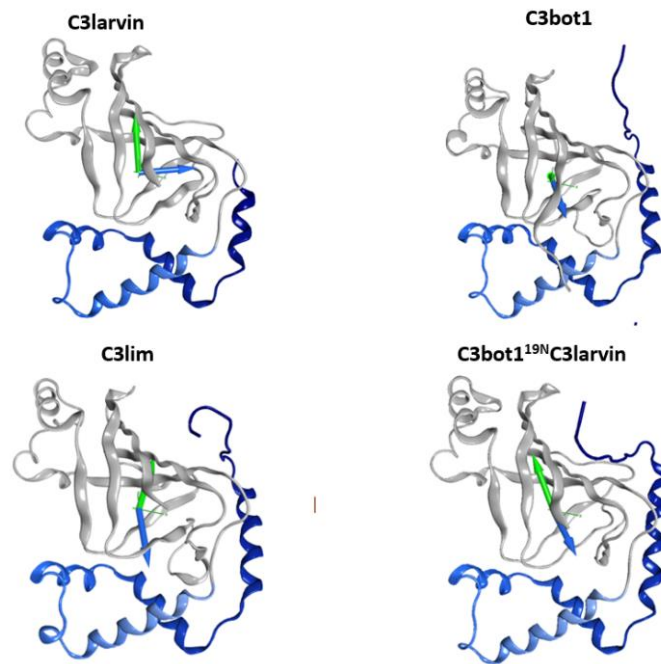


**Figure 8.** Protein–protein interactions in C3larvin. Depiction of the docking of a segment of C3larvin (green backbone) into the NAD<sup>+</sup>-binding pocket of another C3larvin molecule in the crystal structure (PDB: 4TR5). The residues correspond to the motifs presented in Figure 7.

There is additional evidence to suggest that the helical  $\alpha$ -lobe might be part of the binding machinery with specific membrane component(s). In principle, the  $R^{\alpha 3}$  signature fulfills this proposal since  $R^{\alpha 3}$  is in the  $\alpha$ -lobe. However, C3larvin fails to enter vimentin-expressing mouse macrophage cells [27] despite harboring RGD <sup>$\beta 1$</sup> ,  $R^{\alpha 3}$ , and RTE <sup>$\beta 2$</sup> . Notably, C3larvin possesses a truncated  $\alpha_1$ ; a chimeric construct formed by adding 18 N-terminal residues (Ala1-Trp18) from the  $\alpha_1$  helix of C3bot1 achieved cell penetration [27].

Inspection of the short  $\alpha_1$  in C3larvin reveals a negative residue at the N-terminus, Glu2, which is unique for this position (consensus Asn in C3-like toxins and Lys in C2-like toxins) and remarkably represents the only difference with the entire sequence of the Plx2A toxin, which shows excellent cell binding and penetration [28]. This Glu residue at one end and a cluster of basic residues at the other end of  $\alpha_1$  confer a significant dipole-moment to the helix and suggest an electrostatic mode of interaction for this element. Notably, the calculated protein dipole moment for C3larvin points in a direction perpendicular to C3bot1 and C3lim dipole vectors (Figure 9); however, the protein dipole moment for the *in silico* C3bot1<sup>18N</sup>C3larvin chimera is practically aligned to the dipole vectors of C3bot1 and C3lim toxins. Thus, it seems that not only the net charge, but also the charge distribution may be important determinants for cell internalization.

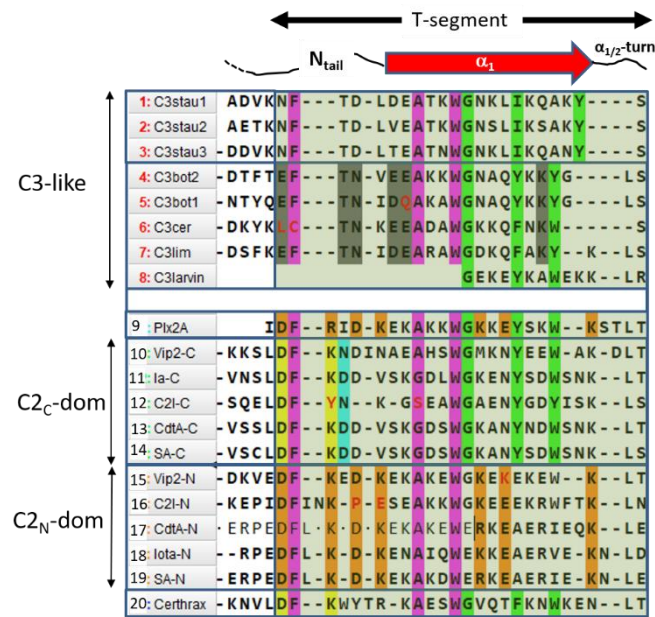
The apparent role of the  $N_{\text{tail}}-\alpha_1$  in mediating toxin cell entry may not be exclusive to C3-etoxins. In the enzymatic component of binary C2-like toxins, the N- (adaptor) and C- (catalytic) domains face one another, so there is an obstruction in the central cleft (NAD<sup>+</sup>-binding pocket) of the N-domain, rendering it inactive (Figure 3). The N-domain of C2I toxin is considered a specialized structure that interacts with the binding/translocation protein partner (B-component) [54], and studies of chimeric fusion proteins with C2 toxin have revealed the  $\alpha_N$ -lobe of this toxin mediated cell entry. In effect, a construct formed by fusing the C2I<sub>N</sub>-domain with the C3lim toxin (C2I<sub>N</sub>C3lim construct) was able to intoxicate mammalian cells in the presence of the activated C2II protein (B-component) [55]. Moreover, the N-terminal  $\alpha$ -lobe of the C2I<sub>N</sub>-domain (residues 1–87) alone was enough to facilitate uptake of the C2I<sub>N</sub><sup>87N</sup>C3bot1 fusion protein [56]. Furthermore, a construct of C2I<sub>N</sub> without the first 29 residues and the C3bot1 toxin, C2I<sub>N</sub> <sup>$\Delta 29N$</sup> C3bot1, failed to enter into HeLa cells [57]. Notably, the C3-like  $N_{\text{tail}}-\alpha_1$  element is preserved in the two domains of all C2-like toxins, and it is comparable in structure and relative location in each of them.



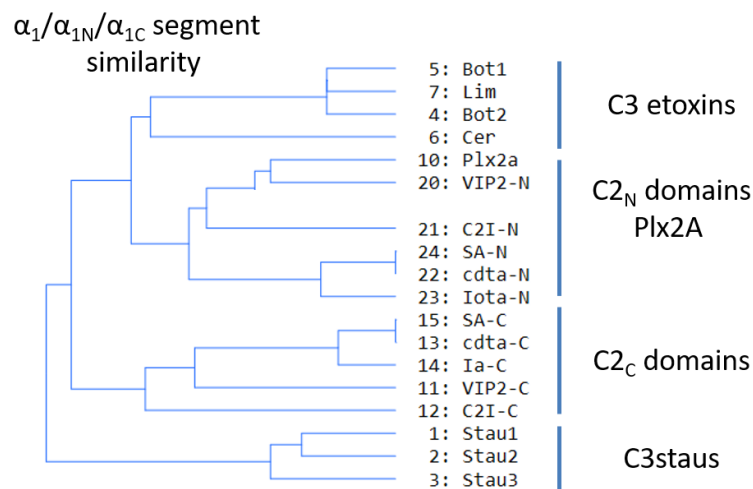
**Figure 9.** Dipole and hydrophobic moments for C3 toxins. The arrows represent the vectors of the electric (blue) and hydrophobic (green) dipole moments for the wild-type (WT) and chimeric (bottom right) proteins.

Plx2A toxin resembles C3-like toxins in the sense that its A-component is a single domain mART enzyme. Accordingly, the  $\alpha$ -lobe of Plx2A is highly like the  $\alpha$ -lobe of C3-like toxins, e.g., it possesses a small  $\alpha_{3/4}$ -loop (or  $\alpha_{4a}$  helix), and the  $N_{tail}-\alpha_1$  is a terminal structure. However, at the same time, Plx2A resembles C2-like toxins because it requires a B-component, Plx2B, as a binding/translocation partner for the cell internalization [28,58]. In this sense, the advent of Plx2A offers an invaluable opportunity to assess the “properties” encoded in the terminal  $N_{tail}-\alpha_1$  segment of the C2- and C3-like toxins.

Sequence-function analysis of the T-segment—a structure-based, multiple-sequence-alignment (MSA) enhanced with the pattern of molecular interactions—was performed for the  $N_{tail}-\alpha_1$  segment of the C3-like toxins, Plx2A, and the  $C2_N$ - and  $C2_C$ -domains (Figure 10). With the previous alignment and taking Plx2A toxin as the “master” sequence because it has the shortest  $N_{tail}$  among the aforementioned toxins, the “T-segment” (“T” after terminal) was identified that comprises five residues of the  $N_{tail}$ , the  $\alpha_1$  helix, and the  $\alpha_{1/2}$ -turn (Figure 10). Subsequently, a cluster analysis based on the similarity of the T-segment residues was performed on this toxin group, excluding C3larvin because it lacks the  $N_{tail}$  and has a truncated  $\alpha_1$  (Figure 11).



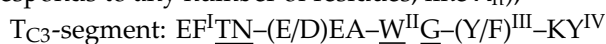
**Figure 10.** The T-segment motif. Multiple sequence alignment of the N<sub>tail</sub>-α<sub>1</sub> segment of C3-like toxins (#1–#8), Plx2A toxin (#9), and C2-like toxins (#10–#19), and definition of the T-segment (translucent green box). The functional motifs are shown in color: purple, S-motif; green, C-motif; yellow and cyan, I-motif; light brown, B-motif; and gray, T-motif. Non-synonymous substitutions are highlighted in red text. Certhrax toxin (#20) is included for comparison.



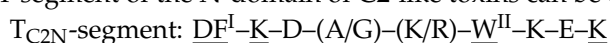
**Figure 11.** T-segment similarity. Cluster analysis of C3-like toxins (except C3larvin), C2-like toxins (N- and C-domains), and Plx2A, based on T-segment similarity.

Three noteworthy features of the T-segment include: (i) C3staus are unique from other domains/toxins; (ii) single-domain C3-etoxins cluster closer with C2<sub>N</sub>-domains than with the catalytic C2<sub>C</sub>-domains of C2-like toxins; and (iii) single-domain Plx2A toxin clusters with C2<sub>N</sub>-domains, rather than with the single-domain C3-etoxins or with the catalytic C2<sub>C</sub>-domains.

The T-segment in C3-etoxins can be defined by the consensus sequence (The long hyphen “–” corresponds to any number of residues; like X<sub>n</sub>),



which comprises four (numbered as I to IV) highly conserved aromatic residues. Likewise, the T-segment of the N-domain of C2-like toxins can be defined by the



and the T-segment of the C-domain of C2-like toxins can be defined by

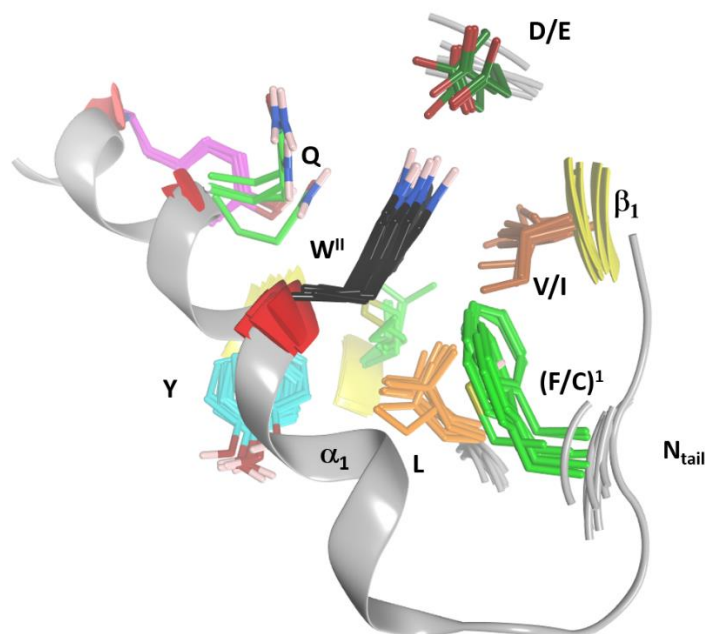
$T_{C2C}\text{-segment: } \underline{D}F^I\text{-K(N/D)D-A-}\underline{W}^{II}\underline{G}\text{-}\underline{Y}^{III}\text{-(Y/W)}^{IV}\text{-}\underline{K}$

Thus, combining the three previous definitions, along with the T-segments of Plx2A, the following consensus sequences arise:

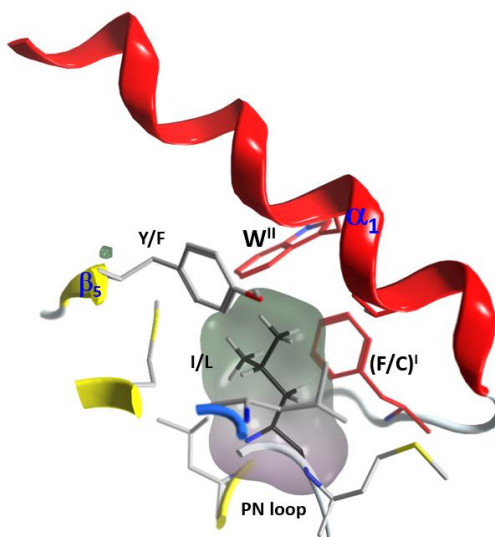
- (S1)  $F^I\text{-}\underline{A}\text{-}\underline{W}^{II}\underline{G}\text{-(Y/F)}^{III}\text{-(Y/W)}^{IV}$ , in Plx2A, and C3-etoxins,  
 (S2)  $\underline{D}F^I\text{-D-(A/G)-}\underline{W}^{II}\underline{G}\text{-}\underline{Y}^{III}\text{-W}^{IV}\text{-}\underline{K}$ , in Plx2A, and C2<sub>C</sub> domains, and  
 (S3)  $\underline{D}F^I\text{-K-D-(K/R)-}\underline{A}\text{-}\underline{W}^{II}\text{-(K/R)-E-}\underline{K}$ , in Plx2A, and C2<sub>N</sub> domains.

In turn, four sequence motifs emerge:

(1) The “S-motif”,  $F^I\text{-(A/G)-}\underline{W}^{II}$ , which is contained in the three S1–S3 consensus sequences motifs (fuchsia residues in Figure 10). The S-motif appears in all T-segments (including C3staus) and therefore likely participates in stabilizing the T-segment and neighboring regions (“S” after structure). In effect, Phe<sup>I</sup> (Cys in C3cer) stabilizes the coiled N<sub>tail</sub> segment with β<sub>1</sub> in a β-like configuration. In turn, (Phe/Cys)<sup>I</sup> along with Ala/Gly and the conserved Trp<sup>II</sup> orient the α<sub>1</sub> helix of the α-lobe by means of a network of hydrophobic and H-bond interactions (Figure 12). In particular, the interaction of the S-motif residues with both semi-conserved (Iso/Leu)<sup>PN</sup> in the PN-loop, and with (Tyr/Phe)<sup>β5</sup> in the ARTT-loop (Figure 13) is relevant and will be discussed later.

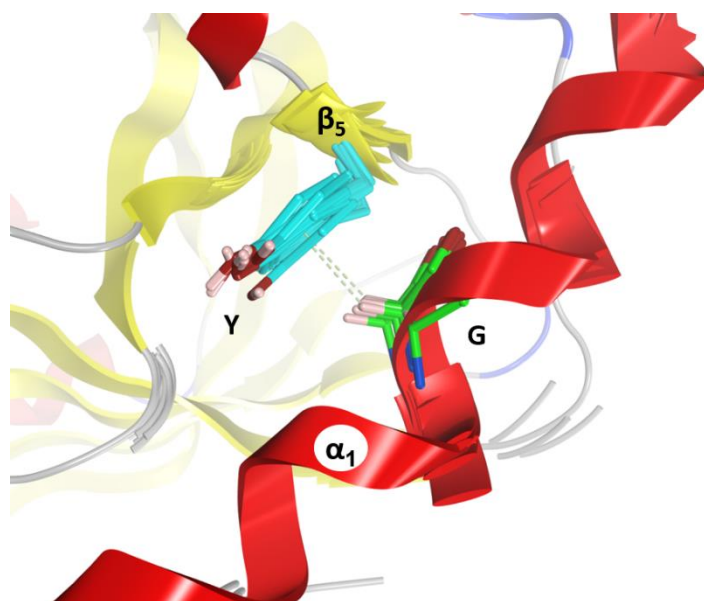


**Figure 12.** S-motif residues. Important residues that cluster around the two aromatic cornerstone residues of the S-motif. The cornerstone residues include Phe (mid-green) and Trp (black) and are surrounded by a cluster of residues, Gln (light green), Tyr (cyan), Leu (orange), Val/Ile (brown) and Asp/Glu (dark green).



**Figure 13.** S-motif key interactions with PN- and ARTT-loops. Depiction of the interactions of the two aromatic residues ( $F^I$  and  $W^{II}$ ) of the S-motif with PN- and ARTT-loop residues (Y/F and I/L).

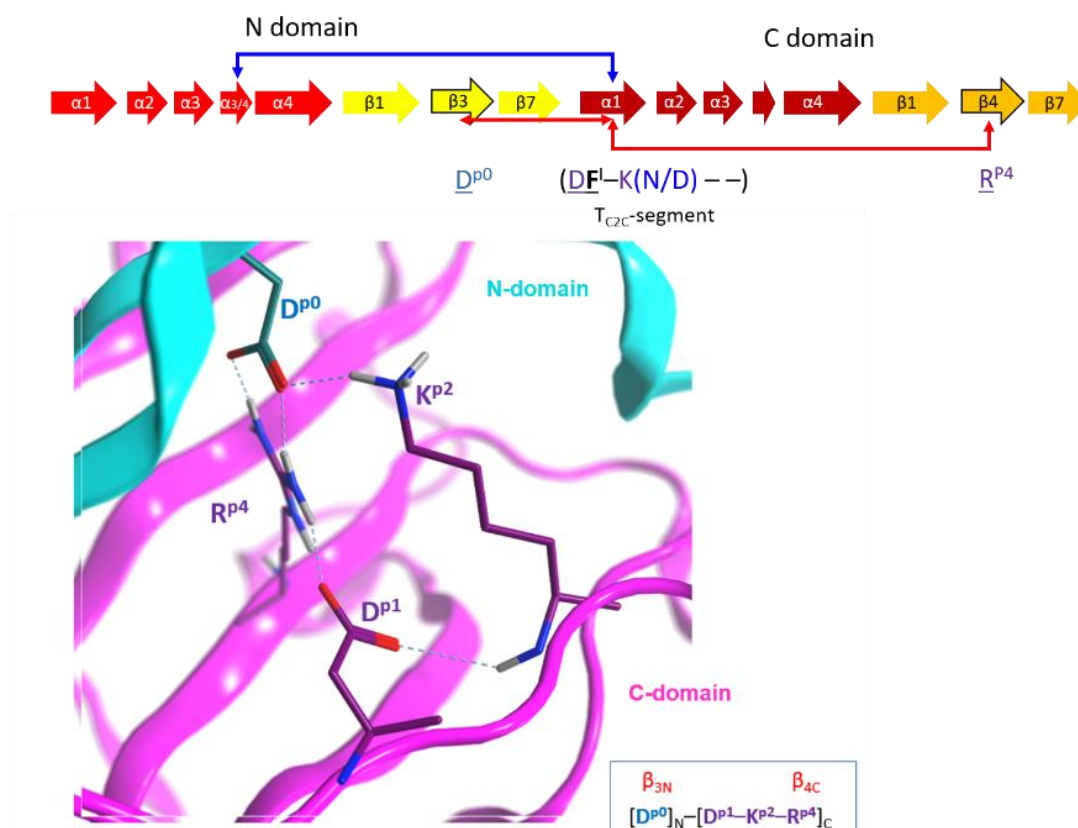
(2) The “C-motif”,  $G-(Y/F)^{III}-(Y/W)^{IV}$  is found in the S1 and S2 consensus sequences and features residues not found in the S-motif (green residues in Figure 10). This motif only appears in catalytic domains/toxins, including C3stau’s (with Ile instead of  $(Y/F)^{II}$ ) and C3larvin, and therefore might be related to the stabilization of catalytic residues (“C” after catalytic). In effect, the conserved Gly residue fits into a small space and interacts/stabilizes the conserved Tyr $^{\beta 5}$  (Phe in C2I<sub>C</sub>) adjacent to the ARTT-loop by a H-Pi-type H-bond (Figure 14). The last two aromatic residues in the motif,  $(Tyr/Phe)^{III}$  and  $(Tyr/Trp)^{IV}$ , in addition to stabilizing the  $\alpha_1$ - $\alpha_2$  superstructure, interact with conserved residues at or adjacent to the ARTT-loop (not shown). Thus, the C-motif might have evolved along with the ARTT-loop and any key determinant might have been lost in the co-evolution of the C2<sub>N</sub>-domains with the “atrophy” of catalytic signatures in the ARTT-loop.



**Figure 14.** Conserved  $\alpha_1$ -ARTT interactions. Depiction of the H Pi-type H-bond between Gly (light green) of the C-motif and the conserved Tyr/Phe residue (cyan) in the ARTT-loop in C3-like toxins.

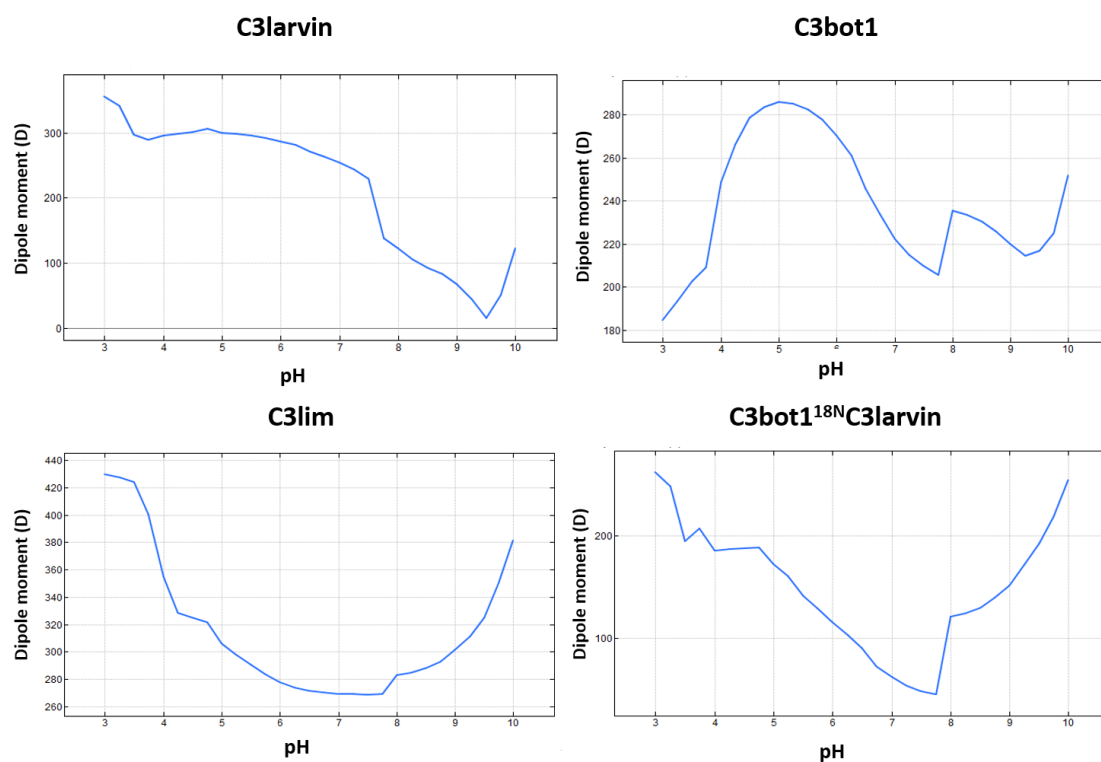
(3) The “B-motif”,  $\underline{D}-(K/R)-D-(K/R)-(K/R)-E-\underline{K}$ , appears only in the S3 consensus sequence and includes residues not present in the S- or C- motifs (orange residues in Figure 10). This unique motif appears in the “terminal” T-segment of the binary toxins A-component—Plx2A and  $C2_N$ -domains. Incidentally, no other basic or acidic residue located in any helix displays this degree of conservation in these toxins/domains ( $C2_N$ -domains, Plx2A) but shows no consensus in the rest of toxins/domains ( $C3$ -like toxins and  $C2_C$ -domains). The length/flexibility, exposure, electric charge, and H-bond capability of the B-motif residues qualify it as a candidate for the binding of the B-component and/or to mediate the translocation of the complex into the cytoplasm (“B” after binding). The B-motif is remarkable, since it confirms that Plx2A shares elements with binary  $C2_N$ -domains that might be implicated in the host cell internalization process.

(4) The “I-motif”,  $D-K(N/D)$ , has residues in the  $T_{C2C}$ -segments that are not included in the previous motifs (yellow and cyan residues in Figure 10) and this motif participates in the inter-domain stabilization (“I” after interdomain). This motif can be defined by three positions ( $p_1 \dots p_3$ ) according to  $\underline{D}^{p1}-K^{p2}(N/D)^{p3}$ . The  $p_1$  and  $p_2$  positions of the I-motif might have evolved along with a residue in the  $\beta_{3N}$ -strand (additional position  $p_0$ ) and a residue in the  $\beta_{4C}$ -strand (additional position  $p_4$ ), to form a quaternary cluster,  $[D^{p0}]_N-[D^{p1}-K^{p2}-R^{p4}]_C$  that stabilized both domains (Figure 15). In agreement with the previous assertion,  $Asp^{p0}$  is conserved in  $\beta_{3N}$  and does not appear in  $\beta_{3C}$ ;  $Arg^{p4}$  is invariant in  $\beta_{4C}$  and does not appear in  $\beta_{4N}$ . Incidentally,  $Asp^{p0}$  and  $Arg^{p4}$  are absent in equivalent elements (strands  $\beta_3$  and  $\beta_4$ , respectively) of single-domain toxins. In addition, the 3rd position of the I-motif,  $(N/D)^{p3}$ , makes effective contacts with the  $\alpha_{3/4}$ -loop in the  $C2_N$ -domains ( $\alpha_{3/4N}$ -loop) to stabilize both domains (see the next section).



**Figure 15.** Inter-domain interactions. Depiction of the inter-domain interactions between the N-domain ( $D^{p0}$ ) and the C-domain ( $D^{p1}K^{p2}-R^{p4}$ ) residues of the I-motif that stabilize the structure of  $C2$ -like toxins.

(5) The “T-motif” E–TN–(E/D)E–K contains residues that are unique in the T<sub>C3</sub>-segments and are not found in the previous motifs (gray residues in Figure 10). The role of this motif might be related to the translocation (“T” after translocation), consistent with the previous section. Accordingly, the increased toxicity observed for C3bot1 and C3lim when the pH is reduced from 7.4 to 5.5 correlates with the increase in both the protein net charge and dipole moment (Figure 16). Furthermore, assessing the pH-dependence of the cellular toxicity, the calculated pKas for Asp90 and Glu129 (the two acidic residues in the proposed binding-motifs of C3bot1) are 2.5 and 2.8, respectively—these are too low to be responsible for the pH response. However, Glu8 and Asp13 (both in the T-motif of C3bot1) have calculated pKa values of 4.6 and 3.8, respectively. C3lim exhibits a higher pH-dependence than C3bot1 in intoxicating macrophage cells [59]. Incidentally, C3lim has four acidic residues in its T-segment, with pKa values that range from 3.7 to 4.0, while the acidic residues of the interacting motifs possess lower pKas of 2.6 and 1.4, respectively. Obviously, the alkaline nature of the T-segment is enhanced by the protonation (neutralization) of the acidic residues.



**Figure 16.** pH-dependence of the protein dipole-moment of C3 toxins. The pH profile of the electrical dipole-moment is shown for three C3-like toxins and for the chimeric protein consisting of the fusion of the N-terminus of C3bot1 with C3larvin toxin.

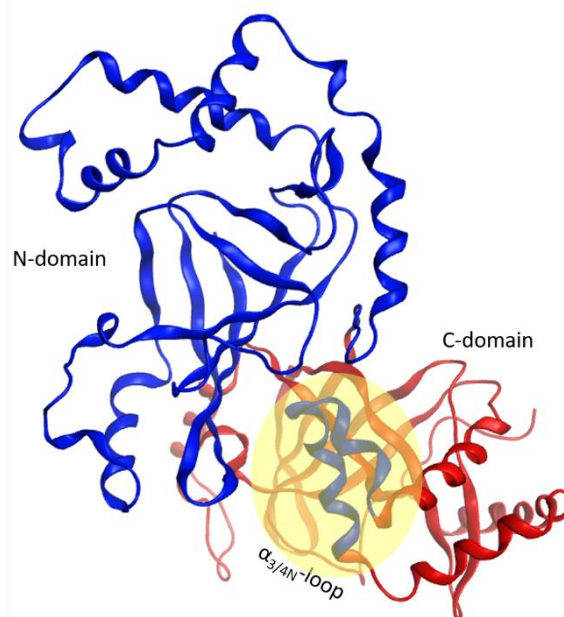
### 2.5. The Putative Role of the $\alpha_{3/4}$ -Loop

In the single-domain Plx2A, and C3-like toxins, the segment that connects helices  $\alpha_3$  and  $\alpha_4$ , the “ $\alpha_{3/4}$ -loop”, is either short and unstructured (e.g., C3bot1) or contains a small  $3_{10}$ -helix configuring a loop-helix-loop motif (e.g., C3lim), called the LHL-motif that protrudes from the  $\alpha$ -lobe [60]. In the C<sub>2C</sub>-domains, the  $\alpha_{3/4C}$ -loop is also short and unstructured (e.g., Vip2) or harbors a small  $\alpha$ -helix (e.g., CdtA). In most of these toxins/domains, there is an abundance of Asn, Gly, and Pro residues, which imply a structural role in allowing the  $\alpha\alpha$ -corner that connects  $\alpha_3$  with  $\alpha_4$  (Gly and Pro allow the turns, while the uncharged, but polar Asn, is either exposed or buried). However, the presence of both charged and hydrophobic residues in the turn of the  $\alpha_{3/4}$ -loops of some C3-like toxins (e.g., C3bot1 and C3bot2), along with their solvent exposure, makes the  $\alpha_{3/4}$ -loop a probable interacting motif in



these toxins. Indeed, even the short LHL-motif of C3bot1 mediates the non-enzymatic interaction of the toxin with the RalA GTPase [60].

On the other hand, the  $\alpha_{3/4N}$ -loop is longer than its counterpart in C3-like toxins and C2<sub>C</sub>-domains and harbors an invariant Phe as well as other polar and hydrophobic residues. Interestingly, in Ia, CdtA, and SA toxins, the  $\alpha_{3/4N}$ -loop and  $\alpha_{4N}$  are highly conserved with an average identity of 80.3% (90.2% similarity), and the Ia  $\alpha_{3N}$ - $\alpha_{4N}$  superstructure may be involved in the interaction with Ib (*iota* B-component) [61]. In effect, the segment responsible for the binding/internalization in Ib of *iota* toxin is not located at the N-terminus (T<sub>C2N</sub>-segment), likewise in C2I toxin (see previous section). Alternatively, in Ia toxin, this segment is more centrally located (residues 62–257), arising from  $\alpha_{3/4N}$  (in the  $\alpha_N$ -lobe) to  $\alpha_{3C}$  (in the  $\alpha_C$ -lobe) including both elements [61]. This postulate is reasonable if the  $\alpha_{3/4N}$ -loop can be considered part of the C2<sub>C</sub>-domain according to distance and packing criteria (Figure 17), such that the spatial proximity between helix  $\alpha_{1C}$  and the  $\alpha_{3/4N}$ -loop may have a functional role. Notably, short segments defined by residues 42–177 (which includes  $\alpha_{3N}$ - and  $\alpha_{4N}$ -helices) and 222–257 ( $\alpha_{1C}$ - $\alpha_{2C}$  of the  $\alpha_C$ -lobe) may control binding to the Ib protein with a C3-like toxin chimera [61]. In addition, the (N/D)<sup>P3</sup> residue of the I-motif interacts with the enlarged  $\alpha_{3/4N}$ -loop to stabilize the inter-domain architecture (see the previous section) with an invariant Ile residue in  $\alpha_{4aN}$  (in  $\alpha_{4N}$  for Vip2<sub>N</sub>).



**Figure 17.** Inter-domain motif proximity. Identification of domains based on geometric criteria by the MOE 2018 algorithm, where the  $\alpha_3$  and  $\alpha_{4a}$  helices of the N-domain (in red ribbons) are part of the C-domain (in blue ribbons) in C2-like toxins. The yellow oval highlights the special proximity of the  $\alpha_1$  of the C-domain with the  $\alpha_{3/4}$ -loop of the N-domain.

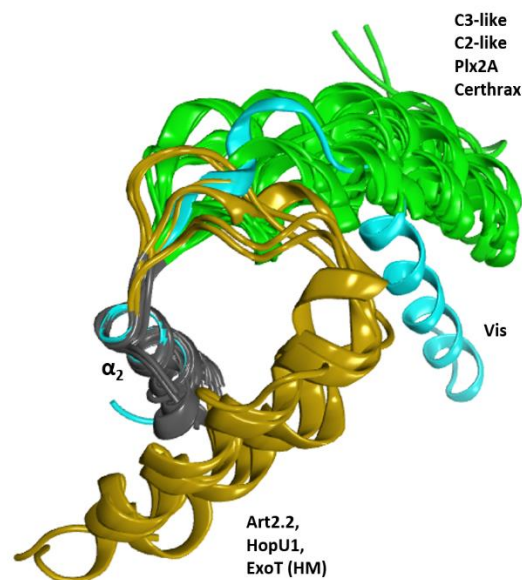
## 2.6. The T-Segment in Other CT-Toxins

Certhrax toxin from *B. cereus* (PDB: 4GF1) is a bi-domain A-component toxin with the catalytically C-domain homologous to the C3-like toxin or the C2<sub>C</sub>-domain, but with the N-domain homologous to the PA-binding domain of anthrax lethal factor from *B. anthracis* [9]. Accordingly, Certhrax T-segment fulfills the S- and C-motifs of the canonical  $\alpha_1$  configuration and catalytic ability (Figure 10). Incidentally, Certhrax T-segment lacks the T-(C3-like translocation) and B-(C2<sub>N</sub>-like binding/translocation) motifs; it exhibits inter-domain interactions via a modified I-motif. Thus, Certhrax clusters with the C2<sub>C</sub>-subgroup based on T-segment similarity (not shown), which is compatible with its catalytic non-terminal T-segment.

The X-ray structure of the mART domain of SpvB (PDB: 2GWM), a single-component actin-ADP-ribosylating toxin from *Salmonella spp*, lacks atomic coordinates of the region corresponding to the T-segment (i.e., upstream in the first “solved” helical structure,  $\alpha_2$ ), and lacks the S- and C-motifs. Accordingly, an estimate of the secondary structure of this segment by the PSIPRED server predicted a coiled structure for this region, and several in-house homology models of the full-length SpvB from *S. typhimurium* report an extended coil for the T-segment.

Another mART actin-modifying toxin with known structure is VahC from *A. hydrophila* [39]. Unfortunately, the N-terminal truncated structures (PDBs: 4FML and 3NTS) lack the coordinates corresponding to the T-segment (i.e., upstream  $\alpha_2$ ). The sequence of the VahC T-segment is highly like the corresponding segment in SpvB, and both segments harbor a poly-proline sequence. Likewise, Photox toxin from *P. luminescens* [40], has a T-segment that is highly similar and unstructured as found in both VahC and SpvB toxins. Interestingly, SpvB, VahC, and Photox lack the conserved Tyr <sup>$\beta^5$</sup> , which is part of the canonical  $\alpha_1$  configuration with catalytic ability (i.e., C3-like and C2<sub>C</sub>-like toxins, including Plx2A, and Certhrax). The phenol side chain of Tyr <sup>$\beta^5$</sup>  is the H-acceptor from the conserved Gly (backbone) in the  $\alpha_1$ -motif (Figure 14). Thus, the G  $\alpha^1$ –Tyr <sup>$\beta^5$</sup>  pair is involved in the mutual stability of  $\alpha_1$  and the ARTT-loop. Notably, VahC, SpvB, and Photox do not show GH activity [7,39,40,43] which might be related to an unstructured T-segment that does not enclose the ARTT-loop (see later).

Vis toxin from *V. splendidus* (PDB: 4XZJ) is a special case. The elongated  $\alpha_1$  of Vis toxin follows a V-shaped configuration of the  $\alpha_1$ – $\alpha_2$  superstructure like C2/C3-toxins, although in a slightly ‘altered’ orientation (Figure 18). Accordingly, Vis has the aromatic W<sup>II</sup> in its T-segment and does not fully qualify as an S-motif that would stabilize helix  $\alpha_1$  in the canonical  $\alpha$ -lobe configuration. Rather, the stabilization of  $\alpha_1$  is achieved by an Arg residue in the N<sub>tail</sub> that forms a salt-bridge with a Glu in the PN-loop and an Asp in the ARTT-loop (see later). Although it possesses some aromatic residues in the T-segment, Vis does not possess a C-motif and is unique in having the small, polar Ser residue replacing the conserved Gly. Notably, this Ser residue interacts in a similar manner with a residue (Thr <sup>$\beta^5$</sup> , also a unique substitution in Vis) proximal to the ARTT-loop as observed for Gly in the C-motif.



**Figure 18.** Conformations of the  $\alpha_1$  helix. Superposition of several toxins of the CT-group for the  $\alpha_2$  helix (gray ribbons) showing three conformation clusters for the  $\alpha_1$  helix. These clusters are the canonical conformation in green, the altered conformation of Vis in cyan, and the folded conformation of HopU1, Art2.2, and ExoT (HM) in ochre.

In contrast, the HopU1 toxin from *Pseudomonas syringae* (PDB: 3UOJ) has  $\alpha_1$  with an orientation diametrically opposed to the canonical conformation: an  $\alpha_1$ – $\alpha_2$  forming an  $\alpha\alpha$ -corner rather than the V-shaped configuration (Figure 18). The  $\alpha\alpha$ -corner configuration between  $\alpha_1$  and  $\alpha_2$  is facilitated by a long  $\alpha_{1,2}$ -loop (termed L1) [44] and is stabilized by a hydrophobic cluster of aromatic residues in the C-terminus of the T-segment (not shown). Although four aromatic residues are found in its T-segment, all the motifs in the canonical T-segment are absent in this toxin; notably, HopU1 lacks the S-motif that anchors  $\alpha_1$  to the  $\beta_1$  strand and lacks the catalytic C-motif. Nevertheless, in HopU1 the N-end of the ARTT-loop (the end linked to strand  $\beta_5$ ) is stabilized by the  $\alpha\alpha$ -corner configuration, while the C-end of the ARTT-loop (the end linked to strand  $\beta_6$ ) is left partially exposed.

### 3. Conclusions

In summary, the role of the  $\alpha$ -lobe is to provide a suitable configuration (location and orientation) of (i) the  $\alpha_2$ – $\alpha_3$  helices to feature the  $\alpha_3$ -motif that has a role in  $\text{NAD}^+$  substrate binding and possibly in the interaction with the protein target; (ii) the  $\alpha_3$ – $\alpha_4$  helices to provide the  $\alpha_{3/4}$ -loop with protein–protein interaction capability; and (iii) the  $\alpha_1$ - $\text{N}_{\text{tail}}$ , defined in the T-segment, that features specialized motif(s) according to the toxin type (A-only or A-B toxins) exhibiting an effect on the catalytic activity via the ARTT-loop, with a role in the inter-domain stability, and with a role in the binding and/or translocation steps during the internalization process.

The canonical (C3-like)  $\alpha$ -lobe configuration has the  $\alpha_1$  helix forming a V-shape with  $\alpha_2$  that surrounds the ARTT-loop. This configuration is stabilized by the ubiquitous S-motif in the T-segment of C3-like, C2-like (N- and C-domains), Plx2A and Certhrax toxin. Accordingly, other non-PT-like toxins (e.g., Vis and HopU1 toxins) with known structures show the canonical  $\alpha_2$ – $\alpha_4$  configuration. However, the  $\alpha_1$  helix in an alternative configuration consistent with the lack of the S-motif in its T-segment.

The presence of the catalytic signature on the ARTT-loop is not enough to guarantee GH-activity—the contrary is true—no catalytic residues in the ARTT-loop means no GH-activity. In non-PT-like toxins, GH-activity requires a stabilized ARTT-loop conformation, and this is achieved in toxins with the canonical  $\alpha$ -lobe in the V-shape configuration of the  $\alpha_1$ – $\alpha_2$  helices, and specifically by the C-motif of the T-segments. In particular, the  $\text{G}^{\alpha_1}$ – $\text{Tyr}^{\beta_5}$  pair is involved in the mutual stability of  $\alpha_1$  and the ARTT-loop. Incidentally, C3larvin, which has a truncated  $\alpha_1$ , still has the C-motif and  $\text{Tyr}^{\beta_5}$  and exhibits GH activity [27]; also, Certhrax has the canonical  $\alpha_1$  configuration,  $\text{Tyr}^{\beta_5}$ , and C-motif and shows GH activity [9]. Accordingly, the alternative  $\alpha_1$  conformation of Vis toxin encloses the ARTT-loop and shows GH activity [5]. On the contrary, VahC, SpvB, and Photox lack both the  $\text{Tyr}^{\beta_5}$  and the C-motif (their T-segments are likely unstructured) and do not show GH activity [7,39,40,43]. HopU1 has an alternative  $\alpha_1$  conformation and shows GH activity (unpublished data).

The specialization of the A-component of binary toxin classes (i.e., C2-like toxins) likely involves selective forces related to the intoxication mechanism that may dictate the composition of the T-segment. The bi-domain constitution might have arisen by gene duplication of an ancestral ADP-ribosyl transferase [62]; consequently, the  $\text{T}_{\text{C}2\text{N}}$ - and  $\text{T}_{\text{C}2\text{C}}$ -segments evolved to harbor residues appropriate for their location and roles. The  $\text{T}_{\text{C}2\text{N}}$ -segment features the S- and B-motifs. In effect, the inability of the  $\text{C}2_{\text{N}}$  domain to bind  $\text{NAD}^+$  must have forced the atrophy of the C-motif and the ARTT-loop (a required motif in the single-domain catalytic precursor), and instead, the  $\text{T}_{\text{C}2\text{N}}$ -segment evolved to interact with the B-component. On the other hand, the  $\text{T}_{\text{C}2\text{C}}$ -segment features the S-, C-, and I- motifs. In effect, the  $\text{T}_{\text{C}2\text{C}}$ -segment might have evolved according to (a) a higher specialization of its ancestral kinetic role. Accordingly, the C-motif is better defined in the  $\text{C}2_{\text{C}}$ -domains ( $\text{G}$ – $\text{Y}^{\text{III}}$ – $\text{W}^{\text{IV}}$ , with only the  $\text{Y}^{\text{IV}}$  variation in  $\text{C}2_{\text{I}}$ -domain) than in the single-domain toxins with the general definition; and (b) possesses an emerging structural role associated with the bi-domain topology—the I-motif. The I-motif, along with specific substitutions in the  $\beta_{3\text{N}}$  and  $\beta_{4\text{C}}$  strand, and with the  $\alpha_{3/4\text{N}}$ -loop, participates in inter-domain stabilization.

It has been reported that the  $\alpha$ -lobe of the  $\text{C}2_{\text{I}}$  domain for C2 toxin and the  $\text{RxG}$  motif (possibly along with  $\text{R}^{\alpha_3}$  and the  $\text{RxE}$  motif) for C3bot1 may be the minimum sub-structure/motif

necessary to stably bind to the membrane component—C2II and vimentin, respectively. In addition, the N-terminal T-segment may also participate in the process of internalization of both toxin groups into the host cell; it be involved in the translocation step from the early endosome to the cytoplasm. Accordingly, C3-like toxins feature the T-motif with acidic residues that might trigger the conformational changes required for membrane translocation; while C2-like toxins feature the B- motif. In this sense, the B-motif may bind to the B-component and/or to mediate the translocation into the cytoplasm. Intoxication experiments that monitor phenotypic alterations of the host cell do not distinguish which event is abolished when working with a toxin. Thus, in agreement with the translocation role of the B-motif, a construct of C2I<sub>N</sub> without the first 29 residues and the C3bot1 toxin, C2I<sub>N</sub><sup>Δ29N</sup>C3bot1, failed to be transported into HeLa cells, although binding of the construct to C2II on cell membranes was still observed [57]. Also, it is feasible that the α<sub>3/4N</sub>-loop is the sub-structure needed for binding (as has been observed in Ia toxin), while the B-motif is needed for the translocation step.

## 4. Materials and Methods

### 4.1. Ensemble of X-Ray Protein Structures

X-ray structures entries were downloaded from the Protein Data Bank (PDB). The datasets include high resolution (1.57 to 2.70 Å) X-ray structures mainly of WT proteins in *apo* forms. However, in some cases proteins in different liganded states and diverse catalytically altered variants were also included for comparative purposes. When multiple molecules were presented in the asymmetric unit of some crystal forms, the most complete molecule was selected.

### 4.2. Force-Field Settings and Structure Preparation

Protein preparation and molecular mechanics (MM) calculations were performed using the computational suite Molecular Operative Environment (MOE) release 2018.10 (Chemical Computing Group Inc, Montreal, CA, USA). The force field employed was the MOE Amber12: EHT, with AMBER12 parameters set (ff12) for protein, and parameters calculated from the Extended Hückley Theory for the NAD<sup>+</sup> molecule and co-solvents. For the implicit solvent model, the Generalized Born-Volume Integral (GB/VI) formalism was employed, with dielectrics  $\epsilon_{\text{pro}} = 2$  for the interior of the protein.

When short sections of X-ray structures were missing in the PDB data files, the peptide segments were crafted by using built-in homology model procedures in MOE. Then, the full X-ray structures were protonated using the MOE Protonate3D module to assign the ionization states and tautomers of protein side-chains and to orient crystallographic water molecules (CWMs) at T = 300 K, pH 7.4 and 0.1 M of ionic strength, along with the GB-VI (Generalized Born-Volume Integral) solvation model and MMFF94 partial charges. The protonated structures were initially geometry-optimized by keeping backbone coordinates fixed, tethering all other heavy atoms with a 100 kcal/mol force constant (0.25 Å buffer), and then energy-minimized until an RMS gradient  $\leq 0.001$  kcal/mol/Å<sup>2</sup>.

### 4.3. Others

The multiple sequence alignments (MSA) were also performed using MOE 2018 software and based on the overall matching of secondary and tertiary structural elements. For a higher resolution alignment of short segments, the procedure was enhanced from the pattern of common molecular interactions. All protein structures were rendered in MOE 2018 software.

**Author Contributions:** Conceptualization, A.R.M.; Data curation, M.R.L.; Formal analysis, M.R.L. and A.R.M.; Funding acquisition, A.R.M.; Project administration, A.R.M.; Resources, A.R.M.; Software, M.R.L.; Writing—original draft, M.R.L.; Writing—review & editing, A.R.M.

**Funding:** This research was funded by the Natural Sciences and Engineering Research Council of Canada, Discovery grant (#105440) and Strategic grant (#494213) to A. Rod Merrill.

**Acknowledgments:** We thank K. Heney for excellent technical support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Simon, N.C.; Aktories, K.; Barbieri, J.T. Novel bacterial ADP-ribosylating toxins: Structure and function. *Nat. Rev. Microbiol.* **2014**, *12*, 599–611. [[CrossRef](#)] [[PubMed](#)]
2. Andreeva, A.; Howorth, D.; Chothia, C.; Kulesha, E.; Murzin, A.G. Investigating protein structure and evolution with SCOP2. *Curr. Protoc. Bioinform.* **2015**, *49*, 1.26.1–1.26.21.
3. Fieldhouse, R.J.; Merrill, A.R. Needle in the haystack: Structure-based toxin discovery. *Trends Biochem. Sci.* **2008**, *33*, 546–556. [[CrossRef](#)] [[PubMed](#)]
4. Barth, H.; Fischer, S.; Moglich, A.; Fortsch, C. Clostridial C3 toxins target monocytes/macrophages and modulate their functions. *Front. Immunol.* **2015**, *6*, 339. [[CrossRef](#)]
5. Ravulapalli, R.; Lugo, M.R.; Pfoh, R.; Visschedyk, D.; Poole, A.; Fieldhouse, R.J.; Pai, E.F.; Merrill, A.R. Characterization of vis toxin, a novel ADP-ribosyltransferase from *Vibrio splendidus*. *Biochemistry* **2015**, *54*, 5920–5936. [[CrossRef](#)] [[PubMed](#)]
6. Barbieri, J.T.; Sun, J. Pseudomonas aeruginosa ExoS and ExoT. In *Reviews of Physiology, Biochemistry and Pharmacology*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 152, pp. 79–92.
7. Sakurai, J.; Nagahama, M.; Hisatsune, J.; Katunuma, N.; Tsuge, H. Clostridium perfringens iota-toxin, ADP-ribosyltransferase: Structure and mechanism of action. *Adv. Enzyme Regul.* **2003**, *43*, 361–377. [[CrossRef](#)]
8. Sanchez, J.; Holmgren, J. Cholera toxin structure, gene regulation and pathophysiological and immunological aspects. *Cell. Mol. Life Sci.* **2008**, *65*, 1347–1360. [[CrossRef](#)]
9. Visschedyk, D.; Rochon, A.; Tempel, W.; Dimov, S.; Park, H.W.; Merrill, A.R. Certhrax toxin, an anthrax-related ADP-ribosyltransferase from *Bacillus cereus*. *J. Biol. Chem.* **2012**, *287*, 41089–41102. [[CrossRef](#)]
10. Simon, N.C.; Barbieri, J.T. Bacillus cereus Certhrax ADP-ribosylates vinculin to disrupt focal adhesion complexes and cell adhesion. *J. Biol. Chem.* **2014**, *289*, 10650–10659. [[CrossRef](#)]
11. Simon, N.C.; Vergis, J.M.; Ebrahimi, A.V.; Ventura, C.L.; O'Brien, A.D.; Barbieri, J.T. Host cell cytotoxicity and cytoskeleton disruption by CerADPr, an ADP-ribosyltransferase of *Bacillus cereus* G9241. *Biochemistry* **2013**, *52*, 2309–2318. [[CrossRef](#)]
12. Fieldhouse, R.J.; Turgeon, Z.; White, D.; Merrill, A.R. Cholera- and anthrax-like toxins are among several new ADP-ribosyltransferases. *PLoS Comput. Biol.* **2010**, *6*, e1001029. [[CrossRef](#)] [[PubMed](#)]
13. Iglewski, B.H.; Sadoff, J.; Bjorn, M.J.; Maxwell, E.S. Pseudomonas aeruginosa exoenzyme S: An adenosine diphosphate ribosyltransferase distinct from toxin A. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 3211–3215. [[CrossRef](#)] [[PubMed](#)]
14. Kodama, T.; Rokuda, M.; Park, K.S.; Cantarelli, V.V.; Matsuda, S.; Iida, T.; Honda, T. Identification and characterization of VopT, a novel ADP-ribosyltransferase effector protein secreted via the *Vibrio parahaemolyticus* type III secretion system 2. *Cell. Microbiol.* **2007**, *9*, 2598–2609. [[CrossRef](#)] [[PubMed](#)]
15. Litvak, Y.; Selinger, Z. Aeromonas salmonicida toxin AexT has a Rho family GTPase-activating protein domain. *J. Bacteriol.* **2007**, *189*, 2558–2560. [[CrossRef](#)] [[PubMed](#)]
16. Aktories, K.; Barmann, M.; Ohishi, I.; Tsuyama, S.; Jakobs, K.H.; Habermann, E. Botulinum C2 toxin ADP-ribosylates actin. *Nature* **1986**, *322*, 390–392. [[CrossRef](#)] [[PubMed](#)]
17. Vandekerckhove, J.; Schering, B.; Barmann, M.; Aktories, K. Clostridium perfringens iota toxin ADP-ribosylates skeletal muscle actin in Arg-177. *FEBS Lett.* **1987**, *225*, 48–52. [[CrossRef](#)]
18. Popoff, M.R.; Milward, F.W.; Bancillon, B.; Boquet, P. Purification of the Clostridium spiroforme binary toxin and activity of the toxin on HEp-2 cells. *Infect. Immun.* **1989**, *57*, 2462–2469. [[PubMed](#)]
19. Chakroun, M.; Banyuls, N.; Bel, Y.; Escriche, B.; Ferre, J. Bacterial vegetative insecticidal proteins (vip) from entomopathogenic bacteria. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 329–350. [[CrossRef](#)]
20. Popoff, M.R.; Rubin, E.J.; Gill, D.M.; Boquet, P. Actin-specific ADP-ribosyltransferase produced by a Clostridium difficile strain. *Infect. Immun.* **1988**, *56*, 2299–2306.
21. Yonogi, S.; Matsuda, S.; Kawai, T.; Yoda, T.; Harada, T.; Kumeda, Y.; Gotoh, K.; Hiyoshi, H.; Nakamura, S.; Kodama, T.; et al. BEC, a novel enterotoxin of Clostridium perfringens found in human clinical isolates from acute gastroenteritis outbreaks. *Infect. Immun.* **2014**, *82*, 2390–2399. [[CrossRef](#)]
22. Tsuge, H.; Tsurumura, T.; Toda, A.; Murata, H.; Toniti, W.; Yoshida, T. Comparative studies of actin- and Rho-specific ADP-ribosylating toxins: Insight from structural biology. *Curr. Top. Microbiol. Immunol.* **2017**, *399*, 69–86. [[PubMed](#)]

23. Aktories, K.; Rosener, S.; Blaschke, U.; Chhatwal, G.S. Botulinum ADP-ribosyltransferase C3. Purification of the enzyme and characterization of the ADP-ribosylation reaction in platelet membranes. *Eur. J. Biochem.* **1988**, *172*, 445–450. [[CrossRef](#)] [[PubMed](#)]
24. Nemoto, Y.; Namba, T.; Kozaki, S.; Narumiya, S. *Clostridium botulinum* C3 ADP-ribosyltransferase gene. Cloning, sequencing, and expression of a functional protein in *Escherichia coli*. *J. Biol. Chem.* **1991**, *266*, 19312–19319. [[PubMed](#)]
25. Just, I.; Mohr, C.; Schallehn, G.; Menard, L.; Didsbury, J.R.; Vandekerckhove, J.; Van Damme, J.; Aktories, K. Purification and characterization of an ADP-ribosyltransferase produced by *Clostridium limosum*. *J. Biol. Chem.* **1992**, *267*, 10274–10280. [[PubMed](#)]
26. Wilde, C.; Vogelsgesang, M.; Aktories, K. Rho-specific *Bacillus cereus* ADP-ribosyltransferase C3cer cloning and characterization. *Biochemistry* **2003**, *42*, 9694–9702. [[CrossRef](#)] [[PubMed](#)]
27. Krska, D.; Ravulapalli, R.; Fieldhouse, R.J.; Lugo, M.R.; Merrill, A.R. C3larvin toxin, an ADP-ribosyltransferase from *Paenibacillus larvae*. *J. Biol. Chem.* **2015**, *290*, 1639–1653. [[CrossRef](#)] [[PubMed](#)]
28. Ebeling, J.; Funfhaus, A.; Knispel, H.; Krska, D.; Ravulapalli, R.; Heney, K.A.; Lugo, M.R.; Merrill, A.R.; Genersch, E. Characterization of the toxin Plx2A, a RhoA-targeting ADP-ribosyltransferase produced by *Paenibacillus larvae*. *Environ. Microbiol.* **2017**, *19*, 5100–5116. [[CrossRef](#)]
29. Wilde, C.; Just, I.; Aktories, K. Structure-function analysis of the Rho-ADP-ribosylating exoenzyme C3stau2 from *Staphylococcus aureus*. *Biochemistry* **2002**, *41*, 1539–1544. [[CrossRef](#)]
30. Evans, H.R.; Sutton, J.M.; Holloway, D.E.; Ayriss, J.; Shone, C.C.; Acharya, K.R. The crystal structure of C3stau2 from *Staphylococcus aureus* and its complex with NAD. *J. Biol. Chem.* **2003**, *278*, 45924–45930. [[CrossRef](#)]
31. Vogelsgesang, M.; Pautsch, A.; Aktories, K. C3 exoenzymes, novel insights into structure and action of Rho-ADP-ribosylating toxins. *Naunyn Schmiedebergs Arch. Pharmacol.* **2007**, *374*, 347–360. [[CrossRef](#)]
32. Gill, D.M.; Meren, R. Adp-ribosylation of membrane proteins catalyzed by cholera toxin: Basis of the activation of adenylate cyclase. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 3050–3054. [[CrossRef](#)] [[PubMed](#)]
33. Farizo, K.M.; Fiddner, S.; Cheung, A.M.; Burns, D.L. Membrane localization of the S1 subunit of pertussis toxin in *Bordetella pertussis* and implications for pertussis toxin secretion. *Infect. Immun.* **2002**, *70*, 1193–1201. [[CrossRef](#)] [[PubMed](#)]
34. Taxt, A.; Aasland, R.; Sommerfelt, H.; Nataro, J.; Puntervoll, P. Heat-stable enterotoxin of enterotoxigenic *Escherichia coli* as a vaccine target. *Infect. Immun.* **2010**, *78*, 1824–1831. [[CrossRef](#)] [[PubMed](#)]
35. Lyons, B.; Ravulapalli, R.; Lanoue, J.; Lugo, M.R.; Dutta, D.; Carlin, S.; Merrill, A.R. Scabin, a novel DNA-acting ADP-ribosyltransferase from *Streptomyces scabies*. *J. Biol. Chem.* **2016**, *291*, 11198–11215. [[CrossRef](#)]
36. Coye, L.H.; Collins, C.M. Identification of SpyA, a novel ADP-ribosyltransferase of *Streptococcus pyogenes*. *Mol. Microbiol.* **2004**, *54*, 89–98. [[CrossRef](#)] [[PubMed](#)]
37. Icenogle, L.M.; Hengel, S.M.; Coye, L.H.; Streifel, A.; Collins, C.M.; Goodlett, D.R.; Moseley, S.L. Molecular and biological characterization of Streptococcal SpyA-mediated ADP-ribosylation of intermediate filament protein vimentin. *J. Biol. Chem.* **2012**, *287*, 21481–21491. [[CrossRef](#)] [[PubMed](#)]
38. Edgar, R.J.; Chen, J.; Kant, S.; Rechkina, E.; Rush, J.S.; Forsberg, L.S.; Jaehrig, B.; Azadi, P.; Tchesnokova, V.; Sokurenko, E.V.; et al. SpyB, a small heme-binding protein, affects the composition of the cell wall in *Streptococcus pyogenes*. *Front. Cell. Infect. Microbiol.* **2016**, *6*, 126. [[CrossRef](#)]
39. Shniffer, A.; Visschedyk, D.D.; Ravulapalli, R.; Suarez, G.; Turgeon, Z.J.; Petrie, A.A.; Chopra, A.K.; Merrill, A.R. Characterization of an actin-targeting ADP-ribosyltransferase from *Aeromonas hydrophila*. *J. Biol. Chem.* **2012**, *287*, 37030–37041. [[CrossRef](#)]
40. Visschedyk, D.D.; Perieteanu, A.A.; Turgeon, Z.J.; Fieldhouse, R.J.; Dawson, J.F.; Merrill, A.R. Photox, a novel actin-targeting mono-ADP-ribosyltransferase from *Photorhabdus luminescens*. *J. Biol. Chem.* **2010**, *285*, 13525–13534. [[CrossRef](#)]
41. Sha, J.; Wang, S.F.; Suarez, G.; Sierra, J.C.; Fadl, A.A.; Erova, T.E.; Foltz, S.M.; Khajanchi, B.K.; Silver, A.; Graf, J.; et al. Further characterization of a type III secretion system (T3SS) and of a new effector protein from a clinical isolate of *Aeromonas hydrophila*—Part I. *Microb. Pathog.* **2007**, *43*, 127–146. [[CrossRef](#)]
42. Greene, L.H.; Lewis, T.E.; Addou, S.; Cuff, A.; Dallman, T.; Dibley, M.; Redfern, O.; Pearl, F.; Nambudiry, R.; Reid, A.; et al. The CATH domain structure database: New protocols and classification levels give a more comprehensive resource for exploring evolution. *Nucleic Acids Res.* **2007**, *35*, D291–D297. [[CrossRef](#)] [[PubMed](#)]

43. Sun, J.; Maresso, A.W.; Kim, J.J.; Barbieri, J.T. How bacterial ADP-ribosylating toxins recognize substrates. *Nat. Struct. Mol. Biol.* **2004**, *11*, 868–876. [[CrossRef](#)] [[PubMed](#)]
44. Jeong, B.R.; Lin, Y.; Joe, A.; Guo, M.; Korneli, C.; Yang, H.; Wang, P.; Yu, M.; Cerny, R.L.; Staiger, D.; et al. Structure function analysis of an ADP-ribosyltransferase type III effector and its RNA-binding target in plant immunity. *J. Biol. Chem.* **2011**, *286*, 43272–43281. [[CrossRef](#)]
45. Lang, A.E.; Schmidt, G.; Schlosser, A.; Hey, T.D.; Larrinua, I.M.; Sheets, J.J.; Mannherz, H.G.; Aktories, K. Photorhabdus luminescens toxins ADP-ribosylate actin and RhoA to force actin clustering. *Science* **2010**, *327*, 1139–1142. [[CrossRef](#)] [[PubMed](#)]
46. Lin, Y.; Wang, P.; Yang, H.; Xu, Y. Crystallization and preliminary crystallographic analysis of the ADP-ribosyltransferase HopU1. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2010**, *66*, 932–934. [[CrossRef](#)] [[PubMed](#)]
47. Von Elsner, L.; Hagemann, S.; Just, I.; Rohrbeck, A. C3 exoenzyme impairs cell proliferation and apoptosis by altering the activity of transcription factors. *Naunyn Schmiedebergs Arch. Pharmacol.* **2016**, *389*, 1021–1031. [[CrossRef](#)]
48. Derossi, D.; Joliot, A.H.; Chassaing, G.; Prochiantz, A. The third helix of the antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.* **1994**, *269*, 10444–10450. [[PubMed](#)]
49. Winton, M.J.; Dubreuil, C.I.; Lasko, D.; Leclerc, N.; McKerracher, L. Characterization of new cell permeable C3-like proteins that inactivate Rho and stimulate neurite outgrowth on inhibitory substrates. *J. Biol. Chem.* **2002**, *277*, 32820–32829. [[CrossRef](#)] [[PubMed](#)]
50. Fahrer, J.; Kuban, J.; Heine, K.; Rupps, G.; Kaiser, E.; Felder, E.; Benz, R.; Barth, H. Selective and specific internalization of clostridial C3 ADP-ribosyltransferases into macrophages and monocytes. *Cell. Microbiol.* **2010**, *12*, 233–247. [[CrossRef](#)] [[PubMed](#)]
51. Rohrbeck, A.; Schroder, A.; Hagemann, S.; Pich, A.; Holtje, M.; Ahnert-Hilger, G.; Just, I. Vimentin mediates uptake of C3 exoenzyme. *PLoS ONE* **2014**, *9*, e101071. [[CrossRef](#)] [[PubMed](#)]
52. Rohrbeck, A.; Just, I. Cell entry of C3 exoenzyme from *Clostridium botulinum*. *Curr. Top. Microbiol. Immunol.* **2017**, *406*, 97–118. [[PubMed](#)]
53. Rohrbeck, A.; Holtje, M.; Adolf, A.; Oms, E.; Hagemann, S.; Ahnert-Hilger, G.; Just, I. The Rho ADP-ribosylating C3 exoenzyme binds cells via an Arg-Gly-Asp motif. *J. Biol. Chem.* **2017**, *292*, 17668–17680. [[CrossRef](#)] [[PubMed](#)]
54. Ohishi, I.; Yanagimoto, A. Visualizations of binding and internalization of two nonlinked protein components of botulinum C2 toxin in tissue culture cells. *Infect. Immun.* **1992**, *60*, 4648–4655. [[PubMed](#)]
55. Barth, H.; Preiss, J.C.; Hofmann, F.; Aktories, K. Characterization of the catalytic site of the ADP-ribosyltransferase *Clostridium botulinum* C2 toxin by site-directed mutagenesis. *J. Biol. Chem.* **1998**, *273*, 29506–29511. [[CrossRef](#)] [[PubMed](#)]
56. Stiles, B.G.; Blocker, D.; Hale, M.L.; Guethhoff, M.A.; Barth, H. *Clostridium botulinum* C2 toxin: Binding studies with fluorescence-activated cytometry. *Toxicon* **2002**, *40*, 1135–1140. [[CrossRef](#)]
57. Barth, H.; Roebeling, R.; Fritz, M.; Aktories, K. The binary *Clostridium botulinum* C2 toxin as a protein delivery system: Identification of the minimal protein region necessary for interaction of toxin components. *J. Biol. Chem.* **2002**, *277*, 5074–5081. [[CrossRef](#)] [[PubMed](#)]
58. Funfhaus, A.; Poppinga, L.; Genersch, E. Identification and characterization of two novel toxins expressed by the lethal honey bee pathogen *Paenibacillus larvae*, the causative agent of American foulbrood. *Environ. Microbiol.* **2013**, *15*, 2951–2965. [[PubMed](#)]
59. Vogelsgesang, M.; Aktories, K. Exchange of glutamine-217 to glutamate of *Clostridium limosum* exoenzyme C3 turns the asparagine-specific ADP-ribosyltransferase into an arginine-modifying enzyme. *Biochemistry* **2006**, *45*, 1017–1025. [[CrossRef](#)] [[PubMed](#)]
60. Pautsch, A.; Vogelsgesang, M.; Trankle, J.; Herrmann, C.; Aktories, K. Crystal structure of the C3bot-RalA complex reveals a novel type of action of a bacterial exoenzyme. *EMBO J.* **2005**, *24*, 3670–3680. [[CrossRef](#)] [[PubMed](#)]

61. Ammam, F.; Marvaud, J.C.; Lambert, T. Distribution of the vanG-like gene cluster in *Clostridium difficile* clinical isolates. *Can. J. Microbiol.* **2012**, *58*, 547–551. [[CrossRef](#)] [[PubMed](#)]
62. Han, S.; Craig, J.A.; Putnam, C.D.; Carozzi, N.B.; Tainer, J.A. Evolution and mechanism from structures of an ADP-ribosylating toxin and NAD complex. *Nat. Struct. Biol.* **1999**, *6*, 932–936. [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).