THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 1, pp. 268–274, January 4, 2008

Structure of the Claudin-binding Domain of *Clostridium perfringens* Enterotoxin*

Received for publication, September 27, 2007, and in revised form, October 16, 2007 Published, JBC Papers in Press, October 31, 2007, DOI 10.1074/jbc.M708066200

Christina M. Van Itallie[‡1], Laurie Betts[§1], James G. Smedley III[¶], Bruce A. McClane[¶], and James M. Anderson[‖2]

From the Departments of [‡]Medicine, [§]Pharmacology, and [‖]Cell and Molecular Physiology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, North Carolina 27599 and the [¶]Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261

Clostridium perfringens **enterotoxin is a common cause of food-borne and antibiotic-associated diarrhea. The toxin's receptors on intestinal epithelial cells include claudin-3 and -4, members of a large family of tight junction proteins. Toxin-induced cytolytic pore formation requires residues in the NH_2-terminal half, whereas residues near the COOH terminus are required for binding to claudins. The claudin-binding COOH-terminal domain is not toxic and is currently under investigation as a potential drug absorption enhancer. Because claudin-4 is overexpressed on some human cancers, the toxin is also being investigated for targeting chemotherapy. Our aim was to solve the structure of the claudin-binding domain to advance its therapeutic applications. The structure of a 14-kDa fragment containing residues 194 to the native COOH terminus at position 319 was solved by x-ray diffraction to a resolution of 1.75 Å. The structure is a nine-strand β sandwich with previously unappreciated similarity to the receptor-binding domains of several other toxins of spore-forming bacteria, including the collagen-binding domain of ColG from *Clostridium histolyticum* and the large Cry family of toxins (including Cry4Ba) of *Bacillus thuringiensis*. Correlations with previous studies suggest that the claudin-4 binding site is on a large surface loop between strands β8 and β9 or includes these strands. The sequence that was crystallized (residues 194–319) binds to purified human claudin-4 with a 1:1 stoichiometry and affinity in the submicromolar range similar to that observed for binding of native toxin to cells. Our results provide a structural framework to advance therapeutic applications of the toxin and suggest a common ancestor for several receptor-binding domains of bacterial toxins.**

Enterotoxin-positive *Clostridium perfringens* type A is one of the most common food poisoning agents in the United States; among the several toxins it produces is *C. perfringens* enterotoxin (CPE).[3] CPE has also been implicated in antibiotic-associated diarrhea in humans and is an important cause of gastrointestinal illness in domestic animals (1). CPE binds to human ileal epithelium, where it induces fluid and electrolyte loss along with epithelial cell necrosis (2). CPE cytotoxicity is a multistep process that initiates with CPE binding to an extracellular loop on specific members of the large claudin family of tight junction proteins (3, 4). This is followed by formation of SDS-resistant complexes of ~450 and ~600 kDa, which contains CPE and claudins; the ~600-kDa complex also contains a second tight junction transmembrane protein, occludin (5, 6). These large complexes are thought to represent the cytotoxic pores, which create a hole in the plasma membrane and lead to cell death (7). The structure of CPE remains unknown, and this has slowed a more detailed understanding of its complex pathogenic mechanism.

The 35-kDa CPE toxin lacks primary sequence homology to any identified bacterial proteins, except for limited similarity (of unknown functional significance) between the NH_2-terminal two-thirds of CPE and some *Clostridium botulinum* proteins. CPE has been functionally divided into two major domains, with the NH_2-terminal half responsible for toxicity and the COOH-terminal half containing the claudin-binding sequences (8) (Fig. 1*A*). Exposure to trypsin or chymotrypsin removes a putative prosequence resulting in a 2–3-fold increase in toxicity. The sequences sufficient for binding to cells and to claudins are located within the COOH-terminal 30 residues. Removal of these residues eliminates binding, and a fusion protein containing only these residues can compete for binding with full-length CPE to isolated brush border membranes (9). In other studies, removal of the last 5 amino acids was sufficient to completely abrogate binding (8) (Fig. 1*A*), although it remains unclear whether these residues contain the binding site or if their removal destabilizes the protein.

Although the COOH-terminal half of CPE, C-CPE, is not cytotoxic, it does have distinct effects on the tight junction barrier of epithelia. For example, exposure of cultured Madin-Darby canine kidney cell epithelia monolayers to C-CPE-(184–319) results in a reversible decrease in transepithelial electrical resistance and increase in paracellular permeability for uncharged solutes (4). This is accompanied by a selective loss of claudin-4, whereas the levels and distribution of claudins that do not bind CPE are unaffected. The selective removal of spe-

* This work was supported by grants from the National Institutes of Health, including a Pilot Feasibility project from P30 DK 034987 (to C. M. V. I.), R01 DK45134 (to J. M. A.), and R37 AI19844 (to B. A. M.).

The atomic coordinates and structure factors (code 2QUO) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: Cell and Molecular Physiology, University of North Carolina at Chapel Hill, 6314 MBRB, CB#7545, 103 Mason Farm Rd., Chapel Hill, NC 27599-7545. Fax: 919-966-6413; E-mail: jandersn@med.unc.edu.

[3] The abbreviations used are: CPE, *C. perfringens* enterotoxin; C-CPE, COOH-terminal domain of CPE; N-CPE, NH_2-terminal domain of CPE.

cific claudins from the tight junction appears to be a different mechanism than CPE-induced toxicity, since the former is slow and reversible and does not affect plasma membrane integrity, whereas CPE-induced toxicity is associated with increased cell membrane permeability and is not reversible. This has led to investigation of C-CPE as an agent to enhance transepithelial drug absorption (10).

CPE is also being investigated in cancer diagnosis and as a potential cancer chemotherapeutic agent. A large number of recent studies document dramatic up-regulation of the CPE receptors, claudin-3 and/or claudin-4, in many pancreatic ovarian, breast, and uterine cancers (reviewed by Morin (11)). For example, injections of CPE into pancreatic tumors induced in nude mice resulted in tumor necrosis and significant reductions in tumor growth (12). Injection of CPE into the peritoneum of mice seeded with human ovarian cancer cells eliminated the malignant cells (13). In other studies, intracranial CPE administration inhibited tumor growth and increased survival in two murine models of breast cancer brain metastasis (14).

In the current study, we determined the structure of the COOH-terminal domain of CPE, residues 194–319, which approximates the C-CPE domain used for functional studies by others and includes the claudin-binding site. The structure of C-CPE-(194–319) is a 9-strand β-sandwich and allows us to rationalize results of the claudin-binding studies. Structural similarity algorithms reveal previously unappreciated structural and sequence homology to the receptor-binding domains located at the COOH termini of several other toxins from spore-forming bacteria, suggesting a common origin. Availability of the structure of C-CPE-(194–319) will accelerate its therapeutic applications in cancer chemotherapy and drug delivery.

EXPERIMENTAL PROCEDURES

Oligonucleotides were synthesized by Invitrogen, and DNA was sequenced at the Genome Analysis Core of the University of North Carolina Lineberger Comprehensive Cancer Center. Crystallography and structural studies were performed in the University of North Carolina Biomolecular X-ray and Structural Bioinformatics Cores.

Protein Expression and Purification

CPE Fragments—Sequences encoding C-CPE-(194–319) were amplified by PCR from C-CPE-(168–319) using the oligonucleotide primers 5′-GGGAATTCGGCgatatagaaaaagaaatccttgatttagctgctgc and 5′-GTGGAAATTACCCTTATTCAATATTATTTCAAAAATTTTAATAAGCTTG and cloned into the EcoRI and HindIII sites of pMAL K4 (15) and expressed in *Escherichia coli* strain DH5α cells. Following induction with 0.3 mM isopropyl 1-thio-β-D-galactopyranoside, bacteria were lysed by sonication in phosphate-buffered saline with protease inhibitor tablets (Complete Mini, EDTA-free; Roche Applied Science) and 1 mM phenylmethylsulfonyl fluoride, and insoluble material was removed by centrifugation at 10,000 $\times$ *g* for 15 min at 4 °C. MBP-C-CPE-(194–319) was affinity-purified from the soluble fraction with cobalt resin (Talon Superflow metal affinity resin; Clontech); purified protein was eluted with 150 mM imidazole in 10 mM Hepes, 120 mM NaCl, pH 7.4. C-CPE-(194–319) was cleaved from MBP by overnight digestion with thrombin (Sigma); after the addition of 1 mM phenylmethylsulfonyl fluoride, thrombin was removed by adsorption to *para*-aminobenzamidine-agarose (Sigma). C-CPE-(194–319) was separated from cleaved MBP by size exclusion chromatography on a Hi-Prep Sephacryl S-100 column (1.6 $\times$ 60 cm; GE Healthcare) in 10 mM Hepes, 120 mM NaCl; C-CPE-(194–319)-containing fractions were concentrated to 10 mg/ml using first a Centriplus then Centricon concentrator. Concentrated protein was clarified by centrifugation at 100,000 $\times$ *g* for 20 min and used for crystallization trials; purity was assessed by SDS-PAGE.

Claudin-4—Full-length human claudin-4 with an NH_2-terminal His_{10} tag was expressed in Sf9 baculovirus cells and purified as previously described (16) with the following modifications. Frozen cell pellets were thawed and washed once in phosphate-buffered saline and then resuspended in 10 mM Hepes, pH 7.4, 120 mM NaCl with protease inhibitor tablets (Complete Mini, EDTA-free, Roche Applied Science) and 1 mM phenylmethylsulfonyl fluoride, 20 units/ml DNase I. Cells were sonicated briefly, dodecylmaltoside (Anatrace, Manumee, OH) was added to 2%, and sample was incubated at 90 min at room temperature with gentle mixing. Insoluble material was removed by centrifugation at 30,000 $\times$ *g* for 40 min, and supernatant was incubated with Co^{2+} resin (Talon) overnight. Resin was washed with 10 mM imidazole, 10 mM Hepes, pH 7.4, 120 mM NaCl, 5 mM dodecylmaltoside, and eluted in the same buffer with 150 mM imidazole. Claudin-4 was further purified by size exclusion chromatography on a Superdex S200 column (1 $\times$ 30 cm; GE Healthcare), and peak fractions were ~70–90% pure (different preparations) based on SDS-PAGE and Coomassie Blue staining. Based on column calibration with molecular weight standards, recombinant claudin-4 elutes at a size consistent with a monomer (15.2 kDa) associated with a DDM micelle (average micelle mass of 50 kDa). For some binding assays, claudin-4 was further purified by rebinding to and re-elution from metal affinity resin after size exclusion chromatography.

Binding Assays

To measure the interaction between bacterially produced C-CPE-(194–319) and claudin-4 purified from insect cells, His-tagged claudin-4 was immobilized on cobalt affinity resin (Talon), and aliquots of 50 μl of beads (5 μM claudin-4) were incubated with 0.5–20 μM purified C-CPE-(194–319) overnight at 4 °C in 10 mM Hepes, pH 7.4, 120 mM NaCl, 5 mM dodecylmaltoside. Beads were washed four times in binding buffer and eluted with binding buffer containing 150 mM imidazole. Eluted material was diluted in SDS-sample buffer, resolved by SDS-PAGE, and detected by Coomassie Brilliant Blue staining. Relative protein levels were quantified using a Licor Odyssey infrared imaging system (Licor Biosciences, Lincoln, NE); binding assays were repeated twice. Binding curves were analyzed by nonlinear regression assuming a single binding site using Prizm software (GraphPad, San Diego, CA).

Crystallization

Crystallization conditions were screened at 4 °C and room temperature using hanging drop vapor diffusion (17) and a commercially available screen (Salt Rx; Hampton Research, Aliso Viejo, CA); small crystals were used to make microseed stocks. Crystals used in analysis were obtained by mixing 2 μl of protein solution at 10 mg/ml, 2 μl of reservoir solution containing 0.1 M Tris, pH 8.3, 0.8 M lithium sulfate, 20% glycerol, and 0.5 μl of microseed solution and allowing the drop to equilibrate against 0.7 ml of reservoir solution using the sitting drop method. Crystals were flash-cooled in liquid nitrogen with cryo-oil (Hampton Research) as cryoprotectant, and all data sets were collected at 100 K (nitrogen stream).

Data Collection, Phasing, and Refinement

The structure was solved by single wavelength anomalous dispersion, using the short halide soak derivatization method of Dauter *et al.* (18). The crystal used for the NaI soak was harvested from its growth drop and soaked for 60 s in a 2.0-μl drop of its mother liquor to which NaI was added to 0.8 M. Data were collected at CuKα at the University of North Carolina Biomolecular X-ray Crystallography Facility on an R-Axis IV++ Image plate. The one sodium iodide-soaked crystal had a different unit cell (same space group, $P2_12_12_1$), whereby the *a* axis was reduced by about 8 Å, the *c* axis was increased by about 12 Å, and the *b* axis was increased by about 2 Å.

The change in unit cell in the NaI-soaked crystal results in completely different packing interfaces and a slight shrinkage (3.6%) in unit cell volume. The NH_2 terminus points in a different direction with respect to the main C-CPE fold in the two crystal forms; however, this segment does not have iodide ions bound in the NaI soak, which might have explained the slightly different directions. As expected, there are iodide ions bound to the surface of the protein and participating in packing interactions in the soaked crystal, which could explain the unit cell rearrangement. It is also possible that the two unit cell packing arrangements are present in native crystals, and we happened to select the different form for the NaI quick soak, but it seems more likely that the rearrangement was caused by the soak. We did not collect data from any other native crystals, but since the structure of the protein is the same (except for the first 7 amino acids, which are not part of the COOH-terminal β sandwich) in the two forms, this observation is not biologically relevant.

Iodide ion sites were found by SHELXD (19) using data to 2.2 Å (three major sites, SHELXD CC = 40.96 all/22 weak). Solvent-flattened maps created by SHELXE were of excellent quality, such that the ArpWarp (20) automatic chain tracing program could place 112 of 126 of the amino acids of C-CPE-(194–319) correctly. Manual model rebuilding was carried out with COOT (21), and all but the first four amino acids are well ordered (these are at the NH_2 terminus of the construct). The model from the iodide soak unit cell was placed into the native unit cell using the molecular replacement program PHASER (22); waters were located, and the model was adjusted with COOT, and then refinement was carried out using Refmac from the CCP4 (23). Final refinement used the native crystal data, because the signal-to-noise ratio was higher in all resolution shells than the iodide-soaked crystal. The final model has good stereochemistry and R_{free} of 22.7%.

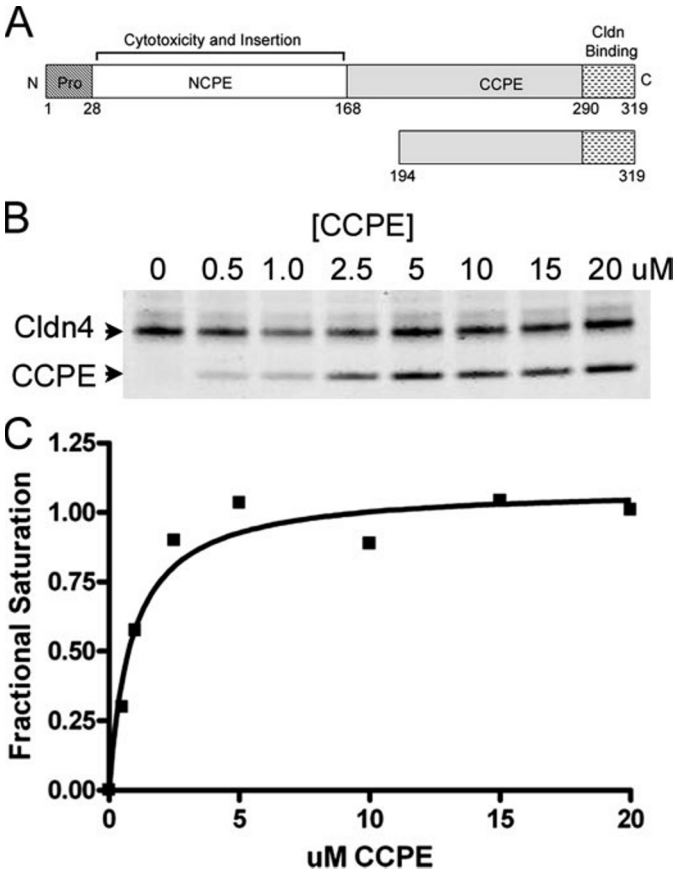

FIGURE 1. **Functional organization of the *C. perfringens* enterotoxin and binding of C-CPE-(194–319) to human claudin-4.** *A*, residues 1–168 have been referred to as N-CPE (NH_2-terminal), and 168–319 have been referred to as C-CPE. Residues within the range 47–51 are required for cytotoxicity via induction of a complex of tight junction proteins to form a transmembrane pore. Removal of residues 290–319 destroys claudin-binding, and a fragment 290–319 can itself compete with CPE in cell surface binding assays. *B*, C-CPE-(194–319) binds with 1:1 stoichiometry and submicromolar affinity to pure claudin-4. A fixed concentration of His_{10}-claudin-4 (5 μM) was immobilized on Co^{2+} beads and mixed with increasing concentrations of purified C-CPE-(194–319) (0–20 μM). The bound C-CPE was resolved by SDS-PAGE and stained with Coomassie Brilliant Blue. There was no detectable binding of C-CPE-(194–319) to Co^{2+} beads alone (not shown). Saturation binding curves of C-CPE-(194–319) to His_{10}-claudin-4. Known concentrations of C-CPE and His_{10}-claudin-4 were loaded on the same gel to confirm the linearity of Coomassie binding. *C*, binding curves were plotted using nonlinear regression analysis of the scanned Coomassie gel and are representative of two similar experiments; $K_d \approx 0.60$ and 0.65 μM in independent experiments. Saturation occurred at 1:1 stoichiometry.

RESULTS AND DISCUSSION

C-CPE-(194–319) Binds to Claudin-4 with High Affinity and a 1:1 Stoichiometry—Our long term goal is to understand how CPE and C-CPE interact with their cellular receptors, which are a subset of the large claudin family. The full-length toxin tends to aggregate after purification; consequently, we initially expressed, purified, and attempted to crystallize the previously characterized C-CPE-(168–319) fragment (Fig. 1*A*). This fragment is known to contain the claudin-binding region within residues 290–319. We found that this fragment is also unstable and oligomerizes upon storage at high concentration. By progressive deletion, we identified a fragment containing residues

TABLE 1
Summary of data collection, phasing, and refinement statistics

Data collection	C-CPE	
	Native	NaI fast soak
Wavelength (Å)	1.5418	1.5418
Resolution range (Å)	40–1.75	40–1.75
No. of measured reflections	111,494	248,626
No. of unique reflections	13,415	13,002
Redundancy (last shell)	3.8 (3.6)	7.5 (5.3)
Completeness (last shell) (%)	99.7 (98.2)	99.6 (97.7)
I/σ (last shell)	33 (7)	23 (2.4)
R_{sym} (last shell) (%)	4.0 (18.5)	9.3 (56.9)
Space group	$P2_12_12_1$	$P2_12_12_1$
a (Å)	37.6	29.6
b (Å)	49.0	50.8
c (Å)	69.1	82.0
No. of molecules in asymmetric unit	1	1
Phasing and refinement		
Figure of merit from experimental phasing (iodide derivative)	0.65	
Resolution range (Å)	40–1.75	
R_{work}/R_{free} (%)	19.7 (22.9)	
No. of reflections for refinement	12,749	
No. of reflections for test set (%)	663 (5%)	
No. of protein residues/non-hydrogen atoms	126 (1218)	
No. of water molecules	213	
Bond lengths (root mean square deviations in Å relative to ideal values)	0.007	
Bond angles (root mean square deviations in degrees relative to ideal values)	1.296	
Residues in favored regions (%)	94	
Residues in additionally allowed regions (%)	6	
Residues in generously allowed/disallowed regions (%)	0	
Average atomic B-factor protein atoms (water atoms)	11.7 (30)	
Estimated average coordinate error from Luzzati plot (Å)	0.2	

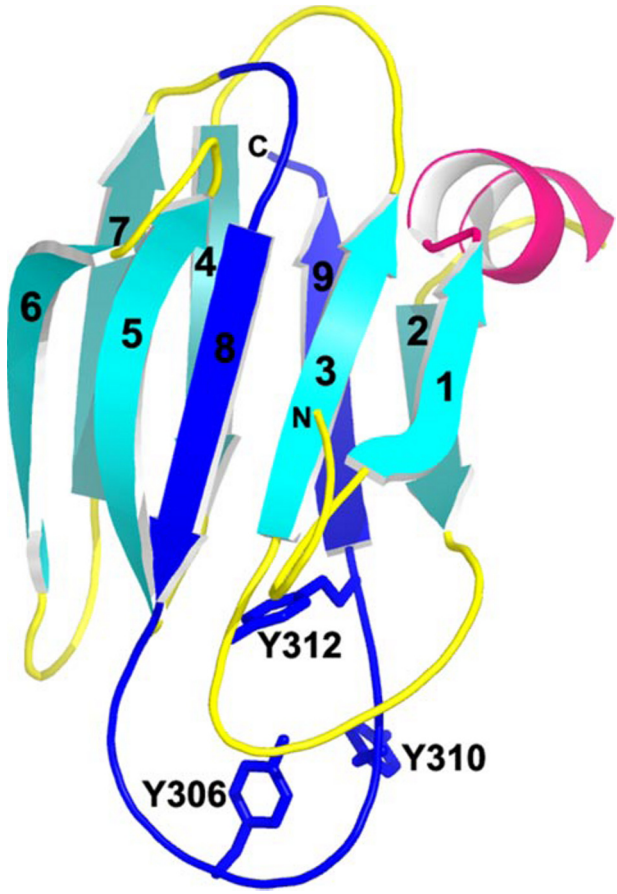

FIGURE 2. **Structure of C-CPE-(194–319).** The structure is a β sandwich with antiparallel orientations except for the uncommon parallel orientation of adjacent strands β1 and β3. The binding site for claudin is within the COOH-terminal 30 residues (*dark blue*), including strands β8 and β9 and the intervening surface loop. Mutagenesis of Tyr^{306}, Tyr^{310}, and Tyr^{316} was previously shown (26) to interfere with binding to claudin.

194–319, which is highly stable and a monomer as determined by gel exclusion chromatography. In order to determine whether this fragment retained the ability to bind to claudin-4, we performed binding studies using purified His_{10}-tagged human claudin-4 immobilized on Co^{2+} resin. Binding saturated with a 1:1 stoichiometry and K_d in the submicromolar range (Fig. 1, *B* and *C*). This high affinity is similar to that previously reported for C-CPE-(168–319) binding to claudin-expressing cultured cells (24), and we conclude that C-CPE-(194–319) is an appropriate fragment of CPE for structural analysis.

Structure of C-CPE-(194–319)—The structure of C-CPE-(194–319) was solved to 1.75 Å resolution using single wavelength anomalous dispersion and the short halide soak derivatization method of Dauter *et al.* (17). The final model has good stereochemistry and R_{free} of 22.7%. A full summary of data collection, phasing, and refinement statistics is included in Table 1.

C-CPE-(194–319) forms a nine-strand β sandwich with a short helical element between strands β1 and β2 (Fig. 2). Adjacent strands within each sheet have antiparallel orientations except for the uncommon parallel alignment of strands β1 and β3 (Figs. 2 and 3). The NH_2-terminal seven residues, which are structured, interact with water molecules and not the main β sandwich, suggesting that the minimal C-CPE domain is formed by residues Ala^{205} through the COOH terminus at Phe^{319}. This should assist future efforts aimed at producing a more solute and stable C-CPE domain for therapeutic application.

The claudin-binding region is known to be within the COOH-terminal 30 residues (25), which corresponds to strands β8 and β9, each positioned in the center of the opposing β sheets, and the intervening surface loop spanning residues Lys^{304}–Tyr^{312}. This region is highlighted in *dark blue* in Fig. 2. Deletion of the COOH-terminal 30 residues eliminates claudin binding (8). However, given the central position of strands β8 and β9, this manipulation may severely destabilize the overall structure, thus raising concerns about the interpretation of these binding studies. In contrast, a fusion protein encoding the same sequence can compete with native CPE for binding to cells, implicating residues in this region as the claudin-binding site. Furthermore, mutagenesis studies more specifically implicate the large surface loop between β8 and β9 in claudin binding (26). Single mutations of Tyr^{306}, Tyr^{310}, or Tyr^{312} to Ala all reduce claudin binding, whereas double mutants of Tyr^{306} with either of the other tyrosines eliminates binding (Fig. 2). Tyr^{306} and Tyr^{312} are buried under the loop; we speculate that their mutation will distort presentation of the loop surface but probably not disrupt the overall β sandwich fold or surface residues on strands β8 and β9. Mutation of Tyr^{306} to phenylalanine, also large and more hydrophobic, did not interfere with claudin binding, whereas mutation to the

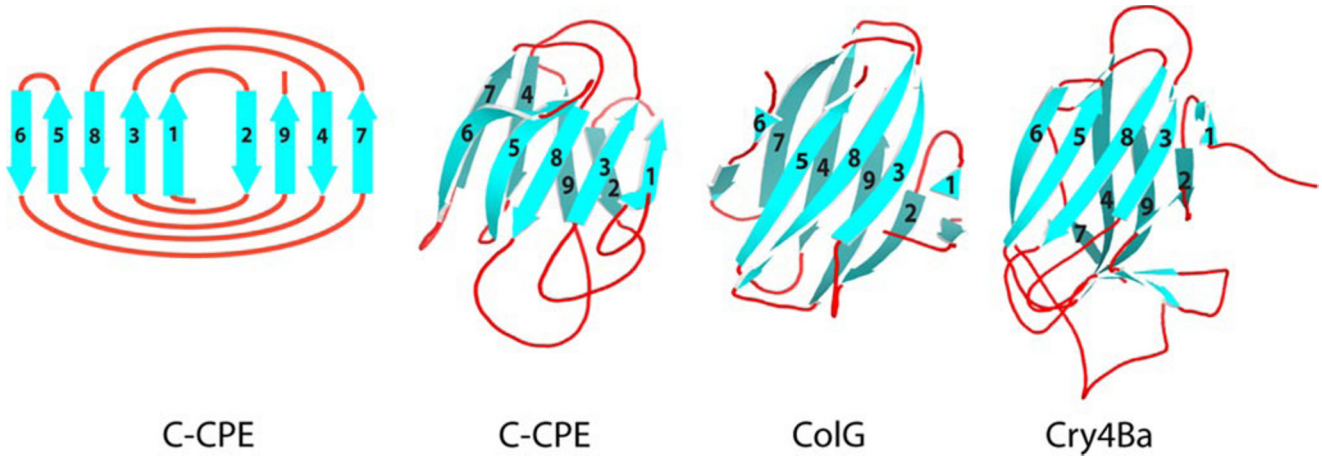

FIGURE 3. **The topology of the ligand-binding domains of C-CPE-(194–319), ColG, and Cry4Ba are similar.** The nine β strands (*blue*) of C-CPE-(194–319) are numbered from the amino end with strands 6-5-8-3-1 and 2-9-4-7 positioned on opposing sheets. Adjacent strands are antiparallel except for the uncommon parallel orientation of strands β3 and β1. This signature topology is shared by the receptor-binding domains of ColG and Cry4Ba. Surface loops are colored *red*. The short helix between β1 and β2 in C-CPE-(194–319) and three short β strands in Cry4Ba between β2 and β3, which are not part of the β sandwich, are removed in this presentation. Structures are oriented to optimize alignment of secondary structure elements using PRGOGRAM.

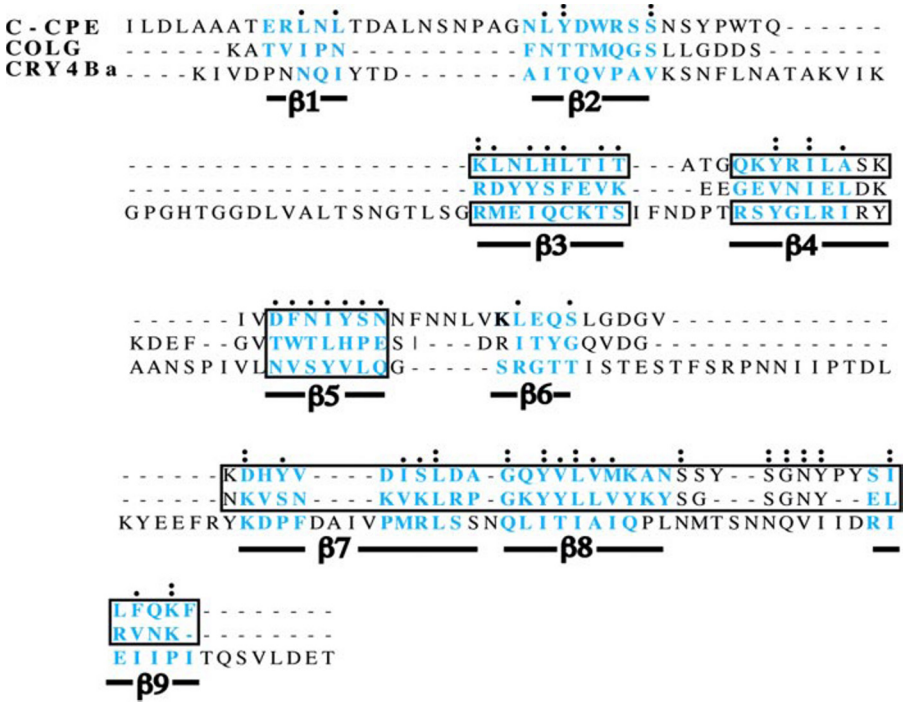

FIGURE 4. **Sequence homology is detected at analogous structural positions along the corresponding strands of C-CPE-(194–319), ColG, and Cry4Ba.** The three sequences are aligned at residues that share analogous three-dimensional structural positions along each strand (*blue*), as detected by the Insight II algorithm. Amino acid similarities (*dots*) and identities (*double dots*) are observed at analogous structural positions within most strands. Searching for sequence homology with ClustalW results in alignment of some (*boxed*) but not all structurally analogous positions. CPE contains an α-helical segment between β1 and β2; Cry4Ba contains three β strands between strands β2 and β3, which are not part of the nine-strand β sandwich.

charged residue lysine eliminated binding (27). Knowledge of the structure suggests very direct mutagenesis studies to define the claudin-binding residues without disrupting the overall domain fold; further identification of these residues and their folding context could lead to development of peptidomimetics with claudin-binding activity. Strands β8 and β9 and especially the large intervening surface loop are attractive sequences for future investigation of claudin-targeting strategies.

The C-CPE-(194–319) Structure Reveals Unexpected Similarity to the Receptor-binding Domains of Other Bacterial Toxins—C-CPE-(194–319) lacks convincing primary sequence homology to domains of any other bacterial toxin when subjected to standard BLAST searching and is thought to be biologically unique. However, a search for similar three-dimensional structures using the DALI (Z-score) and root mean square (D_{RMS}) similarity algorithms (28) revealed structural similarity with domains of more than 10 bacterial toxins. We highlight two examples that have striking biological analogy to C-CPE-(194–319) despite the initial absence of obvious sequence homology, namely domains within the ColG collagenase of *Clostridium histolyticum* (29) (Z-score = 9.1, D_{RMS} = 2.9) and the Cry4Ba toxin of *Bacillus thuringiensis* (30) (Z-score = 6.5, D_{RMS} = 3.8). In both cases, the similarity is with the ~14-kDa COOH-terminal domain of the toxin/enzyme, which, like C-CPE-(194–319), functions in receptor binding by these toxins. Both are β sandwiches that share a similar secondary structure topology with C-CPE-(194–319). This includes the same relationship among β strands in each sheet, 6-5-8-3-1 on one side and 2-9-4-7 on the other, and the unusual parallel orientation of strands β1 and β3, noted in Fig. 3. *Ribbon models* of the three domains are oriented in Fig. 3 to optimize alignment of the secondary structure elements.

Primary sequence homology can be identified within the corresponding β strands after first aligning residues that are

located at analogous three-dimensional positions, using the Insight II algorithm (Accelrys Software Inc.). Alignment of homologous positions within the three sequences is presented in Fig. 4 with residues within each strand *colored blue*. Interestingly, when ClustalW is used to search for sequence homology among the three sequences, it aligns (*boxed regions* of Fig. 4) some but not all of the analogous three-dimensional positions. For example, ClustalW will align β4 of C-CPE and Cry4Ba but not ColG, despite ColG having chemically homologous residues at the analogous three-dimensional positions along the strand. Sequence homology is found at analogous positions within strands 3, 4, 5, 7, 8, and 9. Strands 1, 2, and 6, which are on the periphery of the sheets, are not as clearly related, perhaps suggesting conservation of interactions within but not at the edges of the sandwich. We speculate that these receptor-binding domains share an evolutionary origin because of their similar folds, sequence homology at corresponding structural positions, and biological activities. Others have noted that β sandwich domains are common in receptor-binding domains for both microbial and eukaryotic proteins (31, 32).

C. histolyticum is an anaerobic spore-forming bacterium that causes gas gangrene and produces collagenases that lead to extensive tissue destruction. The mature enzyme class I collagenase from *C. histolyticum* is composed of an NH_2-terminal metalloprotease domain, a spacer domain, and two collagen binding domains. The COOH-terminal binding domain not only shows structural similarity to C-CPE-(194–319), but mutational analysis of selected amino acids demonstrates that, like CPE, several tyrosines near the extreme COOH terminus are critical for collagen binding (29). However, these are not analogous residues, since they are exposed to the solvent and present on strand β8 and not buried on the loop between β8 and β9 like in CPE.

B. thuringiensis is also a spore-forming Gram-positive bacterium. It produces a large number of insecticidal toxins stored as crystalline (Cry) protoxins (33). In an analogous fashion to CPE, following ingestion and activation by gut proteases, the Cry4Ba toxin binds to specific receptors on insect gut epithelial cells. Binding is followed by a conformational change that leads to the formation of membrane pores in the epithelial cells and cell lysis (33). The Cry toxins have three domains, the first of which is the toxic/pore-forming domain, and the second two domains, II and to a lesser extent domain III, are involved in receptor binding (30, 34). Domain III, which has structural similarity to C-CPE, resembles a number of carbohydrate binding domains that are found in microbial hydrolases and esterases (35), but specific residues known to be critical for carbohydrate binding are not generally conserved in the domain III of Cry toxins. Thus, the receptor binding surfaces on Cry proteins and CPE do not appear to be analogous.

C-CPE has been reported to interact with the second extracellular domain of claudins (24). However, although we could demonstrate high affinity interactions with full-length claudin-3 (not shown) and claudin-4, we were unable to demonstrate binding of purified C-CPE to either synthetic peptides or fusion proteins comprising the second extracellular domain of claudin-3 or claudin-4. Future investigations will focus on the region of claudin-4 that interacts with C-CPE. In summary, the availability of a structure for C-CPE-(194–319) suggests an unsuspected shared origin for the receptor-binding domains of several bacterial toxins and should allow rational modifications of the protein for future therapeutic applications.

Acknowledgments—We thank Dr. Brenda Temple (University of North Carolina) for assistance with bioinformatics methods and Drs. Alan Fanning (University of North Carolina) and Arnon Lavie (University of Illinois, Chicago) for helpful suggestions.

REFERENCES

1. Smedley, J. G., III, Fisher, D. J., Sayeed, S., Chakrabarti, G., and McClane, B. A. (2004) *Rev. Physiol Biochem. Pharmacol.* **152,** 183–204
2. Fernandez Miyakawa, M. E., Pistone, C. V., Uzal, F. A., McClane, B. A., and Ibarra, C. (2005) *Infect. Immun.* **73,** 8407–8410
3. Katahira, J., Sugiyama, H., Inoue, N., Horiguchi, Y., Matsuda, M., and Sugimoto, N. (1997) *J. Biol. Chem.* **272,** 26652–26658
4. Sonoda, N., Furuse, M., Sasaki, H., +Yonemura, S., Katahira, J., Horiguchi, Y., and Tsukita, S. (1999) *J. Cell Biol.* **147,** 195–204
5. Singh, U., Van Itallie, C. M., Mitic, L. L., Anderson, J. M., and McClane, B. A. (2000) *J. Biol. Chem.* **275,** 18407–18417
6. Robertson, S. L., Smedley, J. G., III, Singh, U., Chakrabarti, G., Van Itallie, C. M., Anderson, J. M., and McClane, B. A. (2007) *Cell Microbiol.* **9,** 2734–2755
7. Chakrabarti, G., and McClane, B. A. (2005) *Cell Microbiol.* **7,** 129–146
8. Kokai-Kun, J. F., and McClane, B. A. (1997) *Clin. Infect. Dis.* **25,** Suppl. 2, 165–167
9. Hanna, P. C., Mietzner, T. A., Schoolnik, G. K., and McClane, B. A. (1991) *J. Biol. Chem.* **266,** 11037–11043
10. Kondoh, M., Masuyama, A., Takahashi, A., Asano, N., Mizuguchi, H., Koizumi, N., Fujii, M., Hayakawa, T., Horiguchi, Y., and Watanbe, Y. (2005) *Mol. Pharmacol.* **67,** 749–756
11. Morin, P. J. (2005) *Cancer Res.* **65,** 9603–9606
12. Michl, P., Buchholz, M., Rolke, M., Kunsch, S., Lohr, M., McClane, B., Tsukita, S., Leder, G., Adler, G., and Gress, T. M. (2001) *Gastroenterology* **121,** 678–684
13. Santin, A. D., Cane, S., Bellone, S., Palmieri, M., Siegel, E. R., Thomas, M., Roman, J. J., Burnett, A., Cannon, M. J., and Pecorelli, S. (2005) *Cancer Res.* **65,** 4334–4342
14. Kominsky, S. L., Tyler, B., Sosnowski, J., Brady, K., Doucet, M., Nell, D., Smedley, J. G., III, McClane, B., Brem, H., and Sukumar, S. (2007) *Cancer Res.* **67,** 7977–7982
15. Gardberg, A., Shuvalova, L., Monnerjahn, C., Konrad, M., and Lavie, A. (2003) *Structure* **11,** 1265–1277
16. Mitic, L. L., Unger, V. M., and Anderson, J. M. (2003) *Protein Sci.* **12,** 218–227
17. McPherson, A., Hankins, C. N., and Shannon, L. (1987) *J. Biol. Chem.* **262,** 1791–1794
18. Dauter, Z., Dauter, M., and Rajashankar, K. R. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** 232–237
19. Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58,** 1772–1779
20. Morris, R. J., Perrakis, A., and Lamzin, V. S. (2003) *Methods Enzymol.* **374,** 229–244
21. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132
22. Read, R. J. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57,** 1373–1382
23. Collaborative Computational Project 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
24. Fujita, K., Katahira, J., Horiguchi, Y., Sonoda, N., Furuse, M., and Tsukita, S. (2000) *FEBS Lett.* **476,** 258–261
25. Takahashi, A., Kondoh, M., Masuyama, A., Fujii, M., Mizuguchi, H., Horiguchi, Y., and Watanabe, Y. (2005) *J. Control Release* **108,** 56–62
26. Harada, M., Kondoh, M., Ebihara, C., Takahashi, A., Komiya, E., Fujii, M., Mizuguchi, H., Tsunoda, S., Horiguchi, Y., Yagi, K., and Watanabe, Y. (2007) *Biochem. Pharmacol.* **73,** 206–214
27. Ebihara, C., Kondoh, M., Hasuike, N., Harada, M., Mizuguchi, H., Horiguchi,

Y., Fujii, M., and Watanabe, Y. (2006) *J. Pharmacol. Exp. Ther.* **316,** 255–260
28. Guerra, C., and Istrail, S. (2003) *Mathematical Methods for Protein Structure Analysis and Design*, pp. 109–115, Springer-Verlag, New York
29. Wilson, J. J., Matsushita, O., Okabe, A., and Sakon, J. (2003) *EMBO J.* **22,** 1743–1752
30. Boonserm, P., Davis, P., Ellar, D. J., and Li, J. (2005) *J. Mol. Biol.* **348,** 363–382
31. Wendt, K. S., Vodermaier, H. C., Jacob, U., Gieffers, C., Gmachl, M., Peters, J. M., Huber, R., and Sondermann, P. (2001) *Nat. Struct. Biol.* **8,** 784–788
32. Stocker, A., and Azzi, A. (2000) *Antioxid. Redox Signal.* **2,** 397–404
33. Knowles, B. H. (1994) *Adv. Insect Physiol.* **24,** 275–308
34. Burton, S. L., Ellar, D. J., Li, J., and Derbyshire, D. J. (1999) *J. Mol. Biol.* **287,** 1011–1022
35. Pigott, C. R., and Ellar, D. J. (2007) *Microbiol. Mol. Biol. Rev.* **71,** 255–281