Published in final edited form as:

J Mol Biol. 2011 October 14; 413(1): 138–149. doi:10.1016/j.jmb.2011.07.066.

Structure of the food-poisoning *Clostridium perfringens* enterotoxin reveals similarity to the aerolysin-like pore-forming toxins

David C. Briggs^{a,1,*}, Claire E. Naylor^{a,*}, James G. Smedley III^{b,2}, Natalya Lukoyanova^a, Susan Robertson^b, David S. Moss^a, Bruce A. McClane^b, and Ajit K. Basak^a

^aDept of Biological Sciences, Institute of Structural and Molecular Biology, Birkbeck College, Malet Street, London, UK, WC1E 7HX.

^bDepartment of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15219, USA.

Abstract

Clostridium perfringens enterotoxin (CPE) is a major cause of food poisoning and antibiotic-associated diarrhoea. Upon its release from *C. perfringens* spores, CPE binds to its receptor, claudin, at the tight junctions between the epithelial cells of the gut wall, and subsequently forms pores in the cell membranes. A number of different complexes between CPE and claudin have been observed and the process of pore-formation has not been fully elucidated. We have determined the 3D-structure of the soluble form of CPE in two crystal forms by X-ray crystallography, to a resolution of 2.7 and 4.0 Å respectively, and found that the N-terminal domain shows structural homology with the aerolysin-like β -pore-forming family of proteins. We show that CPE forms a trimer in both crystal forms and that this trimer is likely to be biologically relevant but is not the active pore form. We use this data to discuss models of pore formation.

Keywords

Antibiotic associated diarrhoea; spores; claudin; intestine; tight-junction; trimer

© 2011 Elsevier Ltd. All rights reserved.

Corresponding Author: Dr Ajit K. Basak, Tel: +44-20-7631-6823, Fax: +44-20-7631-6802, a.basak@mail.cryst.bbk.ac.uk.

*These authors contributed equally

¹Present Address: Faculty of Life Sciences, The University of Manchester, Simon Building, Oxford Road, Manchester, UK, M13 9PL

²Present Address: KBI Biopharma Inc, PO Box 15579, 1101 Hamlin Road, Durham, NC 27704, USA

David C. Briggs: David.C.Briggs@manchester.ac.uk

Claire E. Naylor: C.naylor@mail.cryst.bbk.ac.uk

Natalya Lukoyanova: n.lukoyanova@mail.cryst.bbk.ac.uk

James G. Smedley III: jsmedley@kbiopharma.com

Susan Robertson: sur18@pitt.edu

David S. Moss: D.moss@mail.cryst.bbk.ac.uk

Bruce A. McClane: bamcc@pitt.edu

Accession Numbers Atomic co-ordinates and structure factors have been submitted to the Protein Databank with accession numbers 2XH6 and 2YHJ.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Clostridium perfringens is a Gram-positive anaerobic spore forming bacterium that is ubiquitous in nature. It can secrete an array of toxins that vary from strain to strain. This toxin diversity allows *C. perfringens* to cause a range of diseases including gas gangrene, human enteritis necroticans, and enterotoxaemia in livestock.

CPE¹ is the major virulence determinant for *C. perfringens* type-A food poisoning^{1,2}. This food poisoning is a significant problem across the globe, ranking as the second most common bacterial food-borne illness in both the USA and UK. The Centers for Disease Control and Prevention (CDC) in the USA estimates that nearly 1 million cases of *C. perfringens* type A food poisoning occur annually in the USA³. This illness results from consumption of food contaminated with *C. perfringens* strains that express CPE in the intestines. Two observations identify CPE as the responsible toxin. Firstly, inactivation of the *cpe* gene eliminated the enteric pathogenicity of a food poisoning strain in animal models and this attenuation was reversible by complementation⁴. Secondly, ingestion of purified CPE by human volunteers resulted in the GI² symptoms of type-A food poisoning⁵.

CPE is also associated with hospital- and community-acquired AAD³ and SD⁴, which are more severe than typical cases of type-A food-borne illness⁶. While the media focuses on AAD and SD caused by strains of *Clostridium difficile*, CPE is also responsible for a large fraction of hospital-acquired enteric illnesses^{7,8}. CPE has also been linked with sudden infant death syndrome⁹ and some veterinary GI diseases^{10,11}.

CPE differs from other *C. perfringens* toxins in two ways, i) it is not a secreted toxin and ii) it is only produced during sporulation, at which time CPE can represent up to 15% of total cell protein. Rather than being secreted, CPE is released at the completion of sporulation upon lysis of the mother cell. This membrane-interacting toxin is a 319 residue polypeptide chain of 35 kDa molecular weight. It can bind to several mammalian cell types via certain claudins, e.g., Claudin-3 or -4², which belong to the large claudin family of 20-27 kDa tight junction proteins^{12,13}. Cytotoxicity and cell death, which occur within a very short period of time following CPE exposure, result from alterations in membrane permeability caused by pore formation.

In mammalian cells, CPE forms two large complexes named CH-1 and CH-2^{14,15}. Those two CPE complexes are both SDS-resistant and their stoichiometry and mass has been the subject of continued study, with recent estimates indicating sizes of approximately 425-500 kDa and 550-660 kDa for CH-1 and CH-2 respectively. These complexes consist of a CPE hexamer, various claudins and, for CH-2, another tight junction protein named occludin¹⁶.

CPE interacts with claudin receptors via its C-terminal domain (residues 200-319¹⁷). In particular, tyrosines 306, 310 and 312 and leucine 315 are important for claudin-binding^{18,19}. The structure of the C-terminal domain, in isolation, has been determined to 1.75 Å²⁰ and the residues associated with claudin-binding reside at the base of a cleft in the C-terminal protein surface (supplementary Fig. S1). However, by itself the C-terminal domain is insufficient for cytotoxicity. The N-terminal domain (residues 1-200) of CPE is necessary for toxin oligomerization into large complexes and for insertion of those complexes into membranes to form active pores. Alanine-scanning mutagenesis identified

¹*Clostridium perfringens* enterotoxin

²Gastrointestinal

³antibiotic-associate diarrhoea

⁴sporadic diarrhoea

aspartate 48 as a key residue for large complex formation²¹, while the deletion of residues 1-36 from native CPE results in an increase in cytotoxicity²².

Beta-pore-forming toxins are a group of cytotoxic proteins with divergent structures and sequences. They are characterized by their common ability to permeabilize cell membranes and ultimately cause cell-death²³. Despite their diverse sequences, these toxins have a common mode of action in that they each possess a stretch of amphipathic residues, which, on binding to the appropriate cell-surface receptors, form a β -hairpin that contributes to the membrane-spanning oligomeric β -barrel of the pore. At high concentrations, CPE is able to form cation-preferring pores in pure lipid bilayers in the absence of receptor-claudin or other proteins²⁴. An amphipathic stretch of residues (residues 81-106) has been identified in the N-terminus of CPE, and deletion of this stretch of residues produces a protein that can form large complexes but not pores¹⁵. On that basis, CPE has been identified as a β -pore-forming toxin. Electric current measurements across lipid bilayers containing CPE pores show that it preferentially transports small positive ions, with ions larger than approximately 6.0 Å unable to pass through the CPE pore²⁴.

The only protein with which CPE has significant sequence homology is the HA3 subcomponent of the botulinum progenitor toxin. The structure of HA3 is known²⁵ and consists of two chains, formed by proteolytic cleavage from a single polypeptide. HA3a (residues 8-184) is a single domain and shows 27% sequence identity with the first 200 residues of CPE. HA3b (residues 204-623) has three domains of which the first two share 25% sequence identity with all 319 residues of CPE (supplementary Fig. S2). The two polypeptides of HA3 form a dimer of trimers with an apparent central pore, though this is of insufficient depth to cross a membrane.

Although the 3D-structure of the C-terminal domain of CPE is known, this provides inadequate information about the oligomerization and pore-forming behaviour of CPE, as these activities are functions of its N-terminal domain. In this manuscript we present the X-ray crystallographic structure of full-length CPE in two crystal forms at 2.7 and 4.0 Å respectively. As we were preparing this manuscript for submission, Kitadokodu *et al*²⁶ published their independently determined structure of CPE in a different crystal form. In this paper we present details about the possible pore-forming mechanisms and on the oligomeric state of CPE that add to the results that provide additional insight into the action of this toxin.

Results

We determined the 3D-structure of CPE to 2.7 Å resolution in space group *C2* with cell dimensions $a=211.1$ Å, $b=119.5$ Å, $c=74.7$ Å, $\beta=110.6^\circ$ containing 3 molecules in the asymmetric unit (PDB ID 2XH6) and in space group *P2₁3* with cell dimensions $a=b=c=159.7$ Å and two molecules in the asymmetric unit (PDB ID 2YHJ). Unless otherwise stated, the results and discussion will describe the higher resolution, crystal form I structure. The 2.7 Å X-ray crystal structure shows that each of the three ncs⁵-related CPE molecules in the asymmetric unit consists of residues 35-319 of CPE ordered into two domains, both with mostly β -sheet-topology (Fig. 1). The three ncs-related copies of the CPE molecule in the asymmetric unit adopt the same conformation, with the C α atom rmsd⁶ between copies in the asymmetric unit of 0.66 Å (between molecules A and B), 0.58 Å (A and C) and 0.40 Å (B and C). These and all subsequent superpositions were carried out using SSM²⁷, unless otherwise stated. In further discussion it may be assumed that it is to

⁵noncrystallographic symmetry

⁶root mean square deviation

the A molecule that we refer. There is no interpretable electron density for the first 34 residues, though we have shown them to be present in the protein used for crystallization, and they are therefore likely to be disordered. It has been established that the deletion of the first 36 CPE residues results in a 2- to 3-fold increase in activity²², indicating these residues are not important for the correct folding of the active protein.

The C-terminal domain (residues 198-319) is a nine-stranded mostly antiparallel β -sandwich with a single short helix between the first two strands (residues 211-217; Fig. 1). C-terminal domain has a C α atom root-mean-square deviation (rmsd) of 0.70 Å when compared to its structure in isolation as determined by van Itallie *et al*²⁰, reflecting the fact that its conformation is unchanged. The N-terminal domain shares the same topology as HA3a and HA3b domain I to which it is distantly related²⁵. This domain has a dog-leg shape and can be divided into two halves (Fig. 1). The first half consists of strands β 1, β 4, β 6', β 7 and β 8 arranged in an antiparallel β -grasp-like topology around the single helix, α A. The second half consists of β 2, β 3, β 5, β 6 and β 7' in a 5-stranded antiparallel β -sandwich. Two long, kinked β -strands, β 6 and β 7 stretch the length of the whole domain.

The HA3b chain domains I and II (residues 204-373) have homology to the N- and C-terminal domains of CPE, with the only difference being that the single helix between strands 2 and 3 of HA3b domain II has been lost in the CPE C-terminal domain, whose single helix is instead between strands β 9 and β 10. However, the two proteins superpose quite poorly, with a C α atom rmsd of 3.86 Å over 202 equivalent atoms, while the expected rmsd for proteins with around 25% sequence homology is approximately 1.1 Å²⁸. This poor superposition is the result of a change in relative domain orientation, with the C-terminal domain rotated 34.2° around a hinge at residue 198 relative to the N-terminal β -grasp and the β -sandwich rotated 25.5° relative to the β -grasp around a hinge passing through residues 49,74,109,135,156 and 185. Once these two rigid body motions are accounted for the rmsd falls to 1.8 Å

The aerolysin-like β -pore-forming toxins are a family of proteins that have structurally related C-terminal domains that are associated with oligomerisation and pore-formation²⁹. The N-terminal domains of these proteins are associated with receptor and/or ligand binding domains and have disparate, unrelated folds. The sequence identity between family members is low to undetectable (no higher than 25%). The structure of the N-terminal domain of CPE reveals that it is related to the aerolysin-like C-terminal domain, as illustrated in Fig. 2, despite having no detectable sequence homology with them. The proteins share a 5-stranded β -sandwich (pink in Fig. 2) and a 4-stranded β -sheet (green in Fig. 2) and have a pair of long twisted β -strands that run the length of both subdomains. CPE and HA3 differ from the aerolysin-like group in that they have a helix lying on this sheet (red in Fig. 2), while the aerolysin-like proteins have a β -hairpin (also red). However, there is evidence to show that both this helix, together the strand preceding it (β 4), in CPE¹⁵ and the β -hairpins in the aerolysin-like family³⁰⁻³³ are amphipathic and likely to be the membrane-spanning residues, in a manner analogous to the β -hairpin of *S. aureus* α -haemolysin³⁴. In support of this data, the helix and strand β 4 (residues 81-106) in CPE form a subdomain that could undergo a conformational change without affecting the overall fold of the rest of the molecule. In addition, the C α atoms of residues 79 and 110 are just 5.8 Å apart and residue 109 is a glycine, providing the required conformational flexibility. The peptide is therefore ideally situated to form a membrane spanning β -hairpin. A surface patch of serine and threonine residues has been observed in the aerolysin-like family, and it has been hypothesized to have a function either in oligomerisation³⁵ or in orientation at the membrane prior to pore insertion³⁶. Interestingly, this serine/threonine track is also present in CPE (Fig. 3) but not in HA3 for which there is no direct evidence for pore-formation.

The MATRAS³⁷ server was used to perform a structure-based sequence alignment of the oligomerisation (pink and green) domains of CPE and HA3 with the equivalent domains in aerolysin-like family members (see Supplementary Fig. S3). The alignment gives pairwise C α RMSDs from 1.46 Å between *C. perfringens* ϵ -toxin (PDB ID 1UYJ) and a non-toxic crystal protein from *B. thuringiensis* (PDB ID 2D42) to 3.36 Å for CPE and aerolysin (PDB ID 1PRE). We used this alignment to produce a HMM⁷ profile using HMMER3³⁸, and a sequence database search using this profile identifies not only the input protein sequences but also a number of other proteins such as *Lysinibacillus sphaericus* Mtx2/3 toxin-like protein (Uniprot ID B1HQ61) and a *B. thuringiensis* crystal protein (Uniprot ID Q45729) that are expected to have pore-forming activities.

The asymmetric unit of the 2.7 Å, C2 structure contains a trimer (Fig. 4) while the asymmetric unit of the 4.0 Å $P2_13$ structure contains two monomers, each of which is part of a trimer that is completed by the crystallographic three-fold axis. The interface is identical in all the three trimers we have observed crystallographically, with 12.2% of each monomer's surface area buried. The interface area is large with 5106 Å² buried surface area for the whole trimer. On average, across two interfaces, each monomer has 859 Å² buried surface area. There are 27 residues (9.5% of the entire sequence) per monomer forming each of the three distinct interfaces that each include 11 hydrogen bonds and a salt bridge. Using PISA³⁹ we calculated ΔG on complexation as -17.2 kcal mol⁻¹ and the trimer "Complexation Significance Score" of 1.00, indicates that it is highly likely to be of biological significance.

Trimers similar to those seen in our crystallographic experiments were observed by negative-stain electron microscopy (EM) of CPE (Fig. 5) in the presence of divalent cations (Mg²⁺, Zn²⁺). However, the toxin appears to be monomeric in the absence of these cations. At the closest approach of the monomers in the crystallographic trimer, Glu 94 is separated from its symmetry equivalents in the trimer by just 4.48 Å (O ϵ 2-O ϵ 1) and Glu 110 (O ϵ 1-O ϵ 1) from its symmetry equivalent by 3.46 Å. The 4.0 Å $P2_13$ structure is from was determined using crystals grown in the presence of 10 mM ZnCl₂, and there is strong (> 3.0 rmsd) difference density in the area between those sidechains that can be attributed to Zn²⁺ (Fig. 6).

The claudin-binding pockets present on the surface of the C-terminal domain of CPE have already been described²⁰. In the trimeric CPE seen in our full-length 3D-structure, these binding pockets are all in a plane on the same side of the molecule (Fig. 4). This alignment would allow interaction with several claudin molecules (claudin is known to be multimeric⁴⁰) in a co-ordinated manner at tight junctions on epithelial cell surfaces. However, a surface representation of the CPE trimer (Fig. 4) shows that the current conformation of CPE is not the active pore form of the protein, as it has no central channel for ions to pass through. In addition, the trimer does not have any large hydrophobic surface patches and, at a maximum width of 50 Å is not wide enough to span the membrane (approx 70 Å).

Discussion

The mechanism of pore formation is of obvious interest to the CPE research community. The residues likely to participate in forming the b-hairpins of the membrane-spanning channel have been identified as residues 81-106, both from this structure and its similarity to aerolysin, and in earlier experiments¹⁵. However, the number of molecules required to form a single pore has not yet been unequivocally determined.

⁷Hidden Markov Model

The trimer seen in these crystal structures, by EM and by others²⁶ must be considered first as it is the only oligomeric form so far observed, and it has now been observed in a wide variety of conditions and so is evidently stable across a range of environments. The bottom surface (the opposite face to the claudin-binding pocket-containing surface) of the crystallographic trimer is planar and made up of the hydrophobic residues and the Ser/Thr track that, in the aerolysin-like β -pore-forming family, has been proposed by some authors to be involved in correctly orientating the molecule on the membrane surface³⁶. Residues 45-53²², and in particular Aspartate 48²¹, have been shown to be essential for CPE's large complex and active pore formation. These residues are located between the two halves of the N-terminal domain. The sidechain of Asp 48 is surface exposed and pointing towards the centre of the trimer (Fig. 4): at its closest approach it is 10 Å from the predicted membrane-spanning residues (81-106) (Asp 48 O δ 1-Ile 106 C δ 1). However, residues 81-106 physically block what would otherwise be a central pore (Fig. 4 and S4). From these observations it is possible to envisage that pore formation might occur by unfolding of residues 79-110 from the centre of the molecule to form 3 β -hairpins, which would each contribute 2 strands to a 6-stranded β -barrel that would form the channel of the pore. Binding of the CPE trimer to a membrane might be signaled to the pore-forming residues via Asp 48, either by interaction of this residue with the membrane surface, or by a conformation change, and this then would trigger insertion of the β -hairpin residues into the bilayer.

We have shown the trimer forms under a number of conditions, and has the attributes of a biologically relevant complex. This trimeric model requires a minimal conformation change of the trimer revealed by crystallography and EM. In addition, the pore is cation-selective and cations are required to observe the trimer by EM. The trimer has a number of electronegative residues at its centre and a likely Zn²⁺ ion has been observed near these residues in the 4.0 Å $P2_13$ structure.

A number of different sized CPE-claudin-other-protein complexes have been purified from brush-border membranes and their properties have been investigated^{2, 15}. Claudin-4 has a molecular weight of 22 kDa, and has been observed forming oligomers of different sizes up to and including hexamers⁴⁰. The pre-pore complex's mass was initially estimated at approximately 155 kDa¹⁵ which would correspond to a trimer of CPE complexed with a trimer of Claudin-4. The mass of the active pore has been estimated to be 200 kDa¹⁴, and a complex formed of a trimer of CPE, a trimer of claudin and a single occludin molecule, would be just a little heavier than this at 210 kDa.

However, this trimeric pore model has a number of features that rule it out as the most likely candidate. The smallest β -barrel pores observed to date have had 8 strands, and while the CPE pore is small²⁴, and a 6-stranded pore cannot be theoretically ruled out⁴¹, it would be highly unusual. All the aerolysin-like β -pore-forming family for which the pore-size is known are either heptamers or hexamers⁴²⁻⁴⁵. In addition, the molecular mass of the pre-pore and large complexes have been revised by Roberson *et al*¹⁶, who reported a pre-pore complex mass in the region of 425-500 kDa, containing 6 CPE molecules and some receptor and nonreceptor claudins. Finally it should be noted that the suggested membrane interaction regions and claudin-binding pockets are on opposite sides of a large and apparently fairly rigid structure, which seems intuitively unlikely.

We therefore conclude that while the observation of the trimer across a range of conditions suggests that, while it has a role in the biology of the CPE, it seems that this role is unlikely to be the pore itself. The trimer may perhaps exist in the spore prior to lysis or as an intermediate in oligomerisation, though the true oligomer may also form from recruitment of monomers.

If the trimer is unlikely to represent the oligomeric state of the CPE pore, alternative models must be considered. Other members of the aerolysin-like β -pore-forming family for which data are available are all hexa- or heptameric, and HA3 forms a dimer of trimers that contains six aerolysin-like oligomerisation and membrane-binding domains. Because of the homology between the CPE N-terminal domain and both HA3a and HA3b domain II, this dimer of trimers would correspond to a hexamer of CPE. By superposing the CPE C-terminal domain, N-terminal β -grasp and β -sandwich domains in turn on each of the equivalent HA3 domains in its oligomer (to allow for the difference in relative orientation discussed earlier), a CPE hexamer may be produced (Fig. 7). This hexamer has a central pore, and its size is in agreement with the latest, larger estimates of CPE complex size. It utilizes the same surfaces for oligomerisation as the heptamer seen in EM reconstructions of the non-toxic aerolysin mutant Y221G⁴⁶. This interface utilizes the Ser/Thr tracks that have been proposed to be important for oligomerisation by others³⁵. Finally, Aspartate 48, which is known to be essential for toxicity, forms part of the oligomeric interface and residues 81-106 are on the same side of the oligomer as the claudin-binding pockets, and are thus well positioned for insertion into the cell's lipid bilayer (Fig. 7).

Although we have not been able to identify a hexameric complex in native gels, by EM or crystallography to date, most aerolysin-like family members only oligomerise at the cell surface following receptor binding, and to date only one oligomer for this family has been observed at atomic resolutions (by EM)⁴⁶. A hexamer may form at the cell surface by recruitment of monomers of CPE via interaction with the receptor, claudin, which is known to form oligomers upto and including hexamers⁴⁰. We also note that the interaction surface of the model CPE hexamer, at 900 Å², whilst still large, is much smaller than the 2550 Å² surface buried in the HA3 dimer of trimers upon which it is based, primarily because of the loss in CPE of large loops in the oligomerisation domains of HA3 that interact with the adjacent monomer. However, this may reflect the dynamic nature of a surface-assembling CPE hexamer.

We therefore conclude that CPE is a new member of the aerolysin-like family of β -pore-forming toxins and, by comparison with these proteins, residues 81-106 while folded into an ordered helix in this soluble form, probably extend to form a membrane-spanning β -hairpin in the pore-form of the protein. Furthermore, we conclude that the trimer seen in both crystal forms described here has biological relevance. However the trimer does not represent the active pore form of the protein: it may be the form adopted in the spore or as an intermediate in pore formation. We also present a hexamer model for the active pore, which could form from the recruitment of the monomers to the cell surface via claudin interaction.

Materials and Methods

Protein production, purification and crystallization were carried out as described previously⁴⁷. Briefly, protein was produced from *C. perfringens* strain NCTC8239 that had been induced to sporulate and purified by ammonium sulphate precipitation followed by size exclusion chromatography. This purified CPE was shipped from the USA to the UK under export licenses D388080 and D446774 from the United States Department of Commerce, Bureau of Industry and Security. The protein was concentrated using a gyrovac and holed eppendorf⁴⁷. Protein ran as a single species on native electrophoresis and gave a mass of 33539 $\pm$ 15 Da by electrospray mass spectrometry, consistent with the theoretical mass of 33530 Da.

Two crystal forms were obtained. Crystal form I grew in sitting drop vapour diffusion experiments, with a 500 μ l reservoir containing 32-40% dioxane in milliQ water and drops containing 1 μ l of 14 mg/ml CPE in milliQ water with 0.1% (v/v) β -octylglucoside mixed

with an equal volume of reservoir solution. Crystals grew at 277K after approximately one week. Crystal form II grew at 277 K after 16 weeks from microbatch trials in which 1 μ l of 10 mg/ml protein in milliQ water were mixed with 1 μ l crystallization buffer (50 mM Citrate buffer, pH 4.3, 10 mM ZnCl₂ and 1.4 M hexane-1,6-diol under 4 ml paraffin oil).

Form I crystals were cryoprotected by soaking in a solution of mother liquor supplemented with 28% glycerol and form II crystals were cryoprotected in 50 mM NaCitrate, pH 4.3, 10 mM ZnCl₂ and 2.0 M hexane-1,6-diol. All crystals were then flash-frozen by immersion in liquid nitrogen and X-ray diffraction data were collected at 100K as has been described previously⁴⁷. Details are provided in Table 1. Indexing, integration and scaling of all data were carried out with Mosflm⁴⁸ and Scala⁴⁹. All other data manipulation was carried out using programs from the CCP4 package⁵⁰.

Data for crystal form I were collected at ESRF station ID14-EH2, to a resolution of 2.7 Å, and were shown to belong to space group *C2* with cell dimensions $a=211.1$ Å $b=119.5$ Å $c=74.7$ Å, $\beta=110.6^\circ$. These dimensions suggested between 3 (Matthews coefficient 4.20 Å³/Da, solvent content 71%) and 6 (Matthews coefficient 2.1 Å³/Da, solvent content 41%) molecules in the asymmetric unit. Data for crystal form II were collected at ESRF station ID14-EH4, to a resolution of 4.0 Å. These data were in space group *P2₁3* and twinned, with a twin fraction estimated to be close to 50%, the Matthew's coefficient suggested between 2 ($V_M=4.73$, 74% solvent) and 4 ($V_M=2.36$, 48% solvent) molecules in the asymmetric unit.

Crystal Form I

The structure of crystal form I was determined by the molecular replacement method using Phaser⁵¹. Three copies of the C-terminal domain (residues 194-319) of CPE (PDB ID 2QUO)²⁰ could be placed with a final translation function Z-score of 19.7. However, this did not result in an interpretable electron density map. A chimeric model was then constructed of domain I in HA3b (PDB ID 2ZS6) and the CPE C-terminal domain located at the same relative position to HA3b domain I as domain II of HA3b. All non-conserved sidechains in the model were mutated to alanine. This model produced a solution containing three molecules in the asymmetric unit with a Z-score of 20.0 in Phaser. Ncs averaging and solvent-flattening were then carried out using DM⁵² resulting in an interpretable electron density map. Initial *R*- and *R_{free}*-factors following rigid-body refinement were 47.4 and 47.1%. Refinement rounds in Phenix⁵³ utilizing ncs restraints, simulated annealing and TLS-refinement were alternated with manual rebuilding using Coot⁵⁴. The final model contained three molecules of CPE (residues 37-319), two β -octylglucoside, three dioxane and 179 water molecules. Three residues are outliers in the Ramachandran plot, as is not uncommon in structures to this resolution, they are not at locations important for the conclusions described here. The final *R*- and *R_{free}*-factors were 21.1 and 24.9%. Structure refinement details are provided in Table 1.

Crystal Form II

This form was solved by molecular replacement with Phaser, using the refined co-ordinates from crystal form I as the model. Molecular replacement revealed the space group to be *P2₁3*, with 2 molecules in the asymmetric unit, giving a Z-score of 44.8 and an initial *R*-factor of 45.3%. Initial maps showed no significant difference in conformation between the CPE molecules in this crystal form compared to those in crystal form I, and due to the low resolution no attempt was made to manually rebuild the structure. A single round of refinement with Buster⁵⁵ restrained to the higher-resolution crystal form I conformation produced *R* and *R_{free}*-factors of 24.9 and 29.9% respectively. Details of the refinement are provided in Table 1.

For electron microscopy, purified CPE (3 μ L; 0.02–0.05 mg/ml) in 100 mM KCl, 5 mM MgCl₂, 20 mM Hepes pH 8.0 was applied to a freshly glow-discharged carbon-coated copper grid (400 mesh, Agar) and stained with a few drops of 1% w/v uranyl acetate. Low dose micrographs of negatively stained samples were recorded on a Gatan 4k $\times$ 4k CCD camera (15 μ m/pixel) using a Tecnai F20 microscope (FEI) at 200 keV and 67,000 $\times$ magnification.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Claire Naylor was supported by the MRC (grant G0700051) and Natalya Lukyanova by the BBSRC (BB/D00873/1) during this work. We would like to acknowledge the Wellcome Trust (grant 079605/2/06/2) for their support for the Electron Microscopy facilities. B. A. McClane and Susan Robertson are supported by the National Institute of Allergy and Infectious Diseases (reference R37AI019844-29).

References

- Adak GK, Long SM, O'Brien SJ. Trends in indigenous foodborne disease and deaths, England and Wales: 1992 to 2000. *Gut*. 2002; 51:832–41. [PubMed: 12427786]
- McClane BA. The complex interactions between *Clostridium perfringens* enterotoxin and epithelial tight junctions. *Toxicon*. 2001; 39:1781–91. [PubMed: 11595640]
- Prevention, C.f.D.C.a. [cited 2011] CDC Estimates of Foodborne Illness in the United States. 2011. major causes of food-poisoning]. Available from: <http://www.cdc.gov/foodborneburden/2011-foodborne-estimates.html>
- Sarker MR, Carman RJ, McClane BA. Inactivation of the gene (cpe) encoding *Clostridium perfringens* enterotoxin eliminates the ability of two cpe-positive *C. perfringens* type A human gastrointestinal disease isolates to affect rabbit ileal loops. *Mol Microbiol*. 1999; 33:946–58. [PubMed: 10476029]
- Skjelkvale R, Uemura T. Experimental Diarrhoea in human volunteers following oral administration of *Clostridium perfringens* enterotoxin. *J Appl Bacteriol*. 1977; 43:281–6. [PubMed: 201601]
- Carman RJ. *Clostridium perfringens* in spontaneous and antibiotic-associated diarrhoea of man and other animals. *Rev. Med. Microbiol*. 1997; 8:S46–S48.
- Brett MM, et al. Detection of *Clostridium perfringens* and its enterotoxin in cases of sporadic diarrhoea. *J Clin Pathol*. 1992; 45:609–11. [PubMed: 1517462]
- Mpamugo O, Donovan T, Brett MM. Enterotoxigenic *Clostridium perfringens* as a cause of sporadic cases of diarrhoea. *J Med Microbiol*. 1995; 43:442–5. [PubMed: 7473678]
- Songer JG. Clostridial enteric diseases of domestic animals. *Clin Microbiol Rev*. 1996; 9:216–34. [PubMed: 8964036]
- Collie RE, McClane BA. Evidence that the enterotoxin gene can be episomal in *Clostridium perfringens* isolates associated with non-food-borne human gastrointestinal diseases. *J Clin Microbiol*. 1998; 36:30–6. [PubMed: 9431915]
- Fisher DJ, et al. Association of beta2 toxin production with *Clostridium perfringens* type A human gastrointestinal disease isolates carrying a plasmid enterotoxin gene. *Mol Microbiol*. 2005; 56:747–62. [PubMed: 15819629]
- Tsukita S, Furuse M, Itoh M. Multifunctional strands in tight junctions. *Nat Rev Mol Cell Biol*. 2001; 2:285–93. [PubMed: 11283726]
- Van Itallie CM, Anderson JM. Claudins and epithelial paracellular transport. *Annu Rev Physiol*. 2006; 68:403–29. [PubMed: 16460278]
- Singh U, et al. CaCo-2 cells treated with *Clostridium perfringens* enterotoxin form multiple large complex species, one of which contains the tight junction protein occludin. *J Biol Chem*. 2000; 275:18407–17. [PubMed: 10749869]

15. Smedley JG 3rd, Uzal FA, McClane BA. Identification of a prepore large-complex stage in the mechanism of action of *Clostridium perfringens* enterotoxin. *Infect Immun*. 2007; 75:2381–90. [PubMed: 17307943]
16. Robertson SL, et al. Compositional and stoichiometric analysis of *Clostridium perfringens* enterotoxin complexes in Caco-2 cells and claudin 4 fibroblast transfectants. *Cell Microbiol*. 2007; 9:2734–55. [PubMed: 17587331]
17. Kokai-Kun JF, McClane BA. Determination of functional regions of *Clostridium perfringens* enterotoxin through deletion analysis. *Clin Infect Dis*. 1997; 25(Suppl 2):S165–7. [PubMed: 9310665]
18. Harada M, et al. Role of tyrosine residues in modulation of claudin-4 by the C-terminal fragment of *Clostridium perfringens* enterotoxin. *Biochem Pharmacol*. 2007; 73:206–14. [PubMed: 17097620]
19. Takahashi A, et al. Domain mapping of a claudin-4 modulator, the C-terminal region of C-terminal fragment of *Clostridium perfringens* enterotoxin, by site-directed mutagenesis. *Biochem Pharmacol*. 2008; 75:1639–48. [PubMed: 18342294]
20. Van Itallie CM, et al. Structure of the claudin-binding domain of *Clostridium perfringens* enterotoxin. *J Biol Chem*. 2008; 283:268–74. [PubMed: 17977833]
21. Smedley JG 3rd, McClane BA. Fine mapping of the N-terminal cytotoxicity region of *Clostridium perfringens* enterotoxin by site-directed mutagenesis. *Infect Immun*. 2004; 72:6914–23. [PubMed: 15557612]
22. Kokai-Kun JF, McClane BA. Deletion analysis of the *Clostridium perfringens* enterotoxin. *Infect Immun*. 1997; 65:1014–22. [PubMed: 9038311]
23. Anderlueh G, Lakey JH. Disparate proteins use similar architectures to damage membranes. *Trends Biochem Sci*. 2008; 33:482–90. [PubMed: 18778941]
24. Hardy SP, et al. *Clostridium perfringens* type A enterotoxin forms mepacrine-sensitive pores in pure phospholipid bilayers in the absence of putative receptor proteins. *Biochim Biophys Acta*. 2001; 1515:38–43. [PubMed: 11597350]
25. Nakamura T, et al. Crystal structure of the HA3 subcomponent of *Clostridium botulinum* type C progenitor toxin. *J Mol Biol*. 2009; 385:1193–206. [PubMed: 19071137]
26. Kitadokoro K, et al. Crystal structure of *Clostridium perfringens* enterotoxin displays features of {beta}-pore-forming toxins. *J Biol Chem*. 2011; 286:19549–55. [PubMed: 21489981]
27. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:2256–68. [PubMed: 15572779]
28. Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. *EMBO J*. 1986; 5:823–6. [PubMed: 3709526]
29. Mancheno JM, et al. *Laetiporus sulphureus* lectin and aerolysin protein family. *Adv Exp Med Biol*. 2010; 677:67–80. [PubMed: 20687481]
30. Iacovache I, et al. A rivet model for channel formation by aerolysin-like pore-forming toxins. *EMBO J*. 2006; 25:457–66. [PubMed: 16424900]
31. Knapp O, et al. Identification of the channel-forming domain of *Clostridium perfringens* Epsilon-toxin (ETX). *Biochim Biophys Acta*. 2009; 1788:2584–93. [PubMed: 19835840]
32. Melton JA, et al. The identification and structure of the membrane-spanning domain of the *Clostridium septicum* alpha toxin. *J Biol Chem*. 2004; 279:14315–22. [PubMed: 14715670]
33. Pelish TM, McClain MS. Dominant-negative inhibitors of the *Clostridium perfringens* epsilon-toxin. *J Biol Chem*. 2009; 284:29446–53. [PubMed: 19720828]
34. Song L, et al. Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science*. 1996; 274:1859–66. [PubMed: 8943190]
35. Cole AR, et al. *Clostridium perfringens* epsilon-toxin shows structural similarity to the pore-forming toxin aerolysin. *Nat Struct Mol Biol*. 2004; 11:797–8. [PubMed: 15258571]
36. Akiba T, et al. Crystal structure of the parasporin-2 *Bacillus thuringiensis* toxin that recognizes cancer cells. *J Mol Biol*. 2009; 386:121–33. [PubMed: 19094993]

37. Kawabata T. MATRAS: A program for protein 3D structure comparison. *Nucleic Acids Res.* 2003; 31:3367–9. [PubMed: 12824329]
38. Durbin, R. *Biological sequence analysis : probabalistic models of proteins and nucleic acids.* Cambridge University Press; Cambridge, UK New York: 1998. p. xip. 356
39. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol.* 2007; 372:774–97. [PubMed: 17681537]
40. Mitic LL, Unger VM, Anderson JM. Expression, solubilization, and biochemical characterization of the tight junction transmembrane protein claudin-4. *Protein Sci.* 2003; 12:218–27. [PubMed: 12538885]
41. Sansom MS, Kerr ID. Transbilayer pores formed by beta-barrels: molecular modeling of pore structures and properties. *Biophys J.* 1995; 69:1334–43. [PubMed: 8534803]
42. Abe Y, Shimada H, Kitada S. Raft-targeting and oligomerization of Parasporin-2, a bacillus thuringiensis crystal protein with anti-tumour activity. *J Biochem.* 2008; 143:269–75. [PubMed: 18006515]
43. Ballard J, et al. Activation and mechanism of Clostridium septicum alpha toxin. *Mol Microbiol.* 1993; 10:627–34. [PubMed: 7968539]
44. Miyata S, et al. Clostridium perfringens epsilon-toxin forms a heptameric pore within the detergent-insoluble microdomains of Madin-Darby canine kidney cells and rat synaptosomes. *J Biol Chem.* 2002; 277:39463–8. [PubMed: 12177068]
45. Wilmsen HU, et al. The aerolysin membrane channel is formed by heptamerization of the monomer. *EMBO J.* 1992; 11:2457–63. [PubMed: 1378390]
46. Tsitrin Y, et al. Conversion of a transmembrane to a water-soluble protein complex by a single point mutation. *Nat Struct Biol.* 2002; 9:729–33. [PubMed: 12219082]
47. Briggs DC, et al. Crystallization and preliminary crystallographic analysis of the Clostridium perfringens enterotoxin. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2010; 66:794–7.
48. Leslie AGW. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography.* 1992; 26
49. Evans P. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr.* 2006; 62:72–82. [PubMed: 16369096]
50. Collaborative Computational Project, N. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994; 50:760–763. [PubMed: 15299374]
51. McCoy AJ, et al. Phaser crystallographic software. *J Appl Crystallogr.* 2007; 40:658–674. [PubMed: 19461840]
52. Cowtan K. 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography.* 1994; 31:34–38.
53. Afonine PV, Grosse-Kunstleve RW, Adams PD. The Phenix refinement framework. *CCP4 Newsletter on Protein Crystallography.* 2005; 42
54. Emsley P, et al. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:486–501. [PubMed: 20383002]
55. Bricogne, G., et al. *Buster.* Global Phasing Ltd; Cambridge, UK: 2010.
56. Schrodinger, LLC. *The PyMOL Molecular Graphics System.* Version 1.3r1. 2010.
57. Chen VB, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:12–21. [PubMed: 20057044]

Figure 1.
Cartoon illustration of the CPE monomer, coloured from blue at the N-terminus to red at the C-terminus. Figure drawn using Pymol ⁵⁶.

Figure 2.

CPE and HA3 and some members of the aerolysin-like β -pore-forming family as cartoons with topology diagrams highlighting with related domains in pink and pale green. (a) CPE, N-terminal domain coloured pink and green, C-terminal coloured cyan, residues 81-106 (predicted to insert into the membrane) coloured red. (b) HA3 A-chain (c) HA3 B-chain, both coloured as for CPE (d) *Laetiporus sulphurus* pore-forming lectin (PDB ID 1W3A) (e) *C. perfringens* ϵ -toxin (PDB ID 1UYJ) (f) *Bacillus thuringiensis* parasporin-2 (PDB ID 2ZTB), all with N-terminal receptor/substrate binding domains and C-terminal peptides coloured cyan, membrane-binding and oligomerization domains (domains II and III) coloured pink and light green and membrane-inserting β -hairpins coloured red.

Figure 3.

Surface representation of CPE with surface Ser/Thr tracks indicated by a circle. Serine and threonine residues are coloured green. Hydrophobic residues (alanine, cysteine, isoleucine, leucine, methionine, proline and valine) coloured yellow and aromatics (phenylalanine, tyrosine and tryptophan) coloured orange.

Figure 4.

CPE trimer, lighter shades showing C-terminal domain, darker ones N-terminal domain. Residues 81-106 coloured red. Aspartate 48 and tyrosines (residues 308, 310 and 312) in the claudin-binding pocket are shown in yellow. (a) Cartoon representation with Aspartate 48 side-chain in space-filling and claudin-binding pocket as sticks. (b) Surface representation viewed from the 'top' (claudin-binding pockets uppermost and (c) Surface representation viewed from the bottom.

Figure 5. Negative stain electron microscopy images of CPE in the presence of magnesium. Scale bar is 20 nm.

Figure 6. Centre of the CPE trimer showing the location of Glu 94 and 110, and positive difference electron density from the P2₁3 data that is likely attributable to Zn²⁺.

Figure 7. Model of CPE as a hexamer based on the dimer of trimers seen in HA3. Hinge angles of the CPE protein have been altered to match those of HA3, both at the mid-point of the N-terminal domain and between the N- and C-terminal domains. C-terminal domains in lighter shades, N-terminal domain darker, residues 81-106 in red. Aspartate 48 space-filling, claudin-binding pocket tyrosines in blue sticks.

Table 1

X-ray Data Collection and Refinement Statistics. Outershell values are given in parentheses.

	Crystal Form I	Crystal Form II
Synchrotron/beamline	ESRF ID14-EH2	ESRF ID14-EH4
Crystal parameters		
Space group	C2	P2 ₁ 3
Cell dimensions (Å)	a=211.1, b=119.5, c=74.7	a=b=c=159.7
Angles (°)	$\alpha=\gamma=90$, $\beta=110.6$	$\alpha=\beta=\gamma=90$
Data Collection		
Wavelength (Å)	0.9686	1.2823
Resolution limit (Å)	70.0-2.68 (2.82-2.68)	70.0-4.0 (4.5-4.0)
Mosaicity		
R _{merge}	0.089 (0.361)	0.171 (0.623)
Total number of observations	95985 (13648)	117419 (12325)
Total number unique	42488 (6259)	11712 (1657)
Mean I/ σ I	12.0 (2.2)	16.7 (3.2)
Completeness	95.8 (83.3)	99.5 (99.9)
Multiplicity	2.7 (2.4)	10.0 (7.4)
Refinement		
Protein atoms in model	6613	4416
Solvent atoms in model	237	0
R _{working}	0.211	0.249
R _{free}	0.249 ^a	0.299 ^b
rmsd from ideal geometry ^a		
Bond lengths (Å)	0.014	0.007
Bond angles (°)	1.532	1.03
Wilson B-factor	70.9	85.2
B-factor of protein atoms	55.7	Overall B: 115.4
Ramachandran plot ^c		
Most favoured (%)	94.6	97.7
Outlier (%)	0.35	0.0
PDB code	2XH6	2YHJ

$$R_{\text{merge}} = \sum |I_i - \langle I_i \rangle| / \sum I_i$$

$$R_{\text{working}} = \sum |F_o - F_c| / \sum F_o$$

R_{free} is the R-factor calculated for the cross-validated test set of reflections.

^aR_{free} is 5.1% of reflections

^bR_{free} is 4.74% of reflections

^cAs defined by MOLPROBITY