The Three-Dimensional Crystal Structure of Cholera Toxin

Rong-Guang Zhang*, David L. Scott‡, Mary L. Westbrook*, Sharon Nance*, Brenda D. Spangler*†, and Edwin M. Westbrook*§†.

*Center for Mechanistic Biology and Biotechnology, Argonne National Laboratory, Argonne, Illinois 60439.

‡Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511.

§Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston, Illinois 60602.

†Present Address: Department of Biological Sciences and Department of Chemistry, Northern Illinois University, Dekalb, Illinois 60115.

‖Author to whom correspondence should be addressed.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
Seven pandemics of cholera have been recorded since 1817, with a current major outbreak in South America. Efforts to prevent the spread of cholera have focused on improving sanitation, protecting water supplies, and developing effective vaccines. The clinical manifestations of cholera are largely attributable to the actions of a secreted hexameric AB5 enterotoxin (choleragen). We have solved the three-dimensional structure of choleragen at 2.5 Å resolution and compared the refined coordinates with those of choleragenoid (isolated B pentamer) and the heat-labile enterotoxin from Escherichia coli (LT). The crystalline coordinates provide a detailed view of the stereochemistry implicated in binding to GM1 gangliosides and in carrying out ADP-ribosylation. The A2 chain of choleragen, in contrast to that of LT, is a nearly continuous α-helix with an interpretable carboxyl tail (KDEL).

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
Cholera remains a major problem in developing nations which lack the sanitary facilities to prevent bacterial contamination of water and food supplies (1). Untreated, the profuse intestinal diarrhea characteristic of the classical form of cholera leads to dehydration and hypovolemic shock. Severe morbidity and death can usually be avoided by fluid therapy alone since the dysentery is self-limited after a course of 2 to 7 days. Current vaccines, using either purified toxin or peptide epitopes, have failed to provide effective long-term immunity (2).

The massive fluid release accompanying infection with large numbers of *Vibrio cholerae* results from the secretion of a potent enterotoxin (3). Cholera toxin is a heterohexameric AB5 complex (Mr = 85,620) that binds via its pentameric B subunits to GM1 gangliosides exposed on the luminal surface of intestinal epithelial cells (4). The A subunit is subsequently activated by proteolytic cleavage and reduction of an intrachain disulfide bond. The enzymatic A1 fragment (residues 1-192 or 1-194) released by reduction inserts into the membrane where it associates with one or more cytosolic factors (5). The A1 fragment binds NAD and catalyzes the ADP-ribosylation of Gsα, a GTP-binding regulatory protein associated with adenylate cyclase (6). The resulting increase in intracellular cyclic AMP activates luminal sodium pumps via a cyclic AMP-dependent protein kinase.

Cholera toxin (CT) is one of several microbial toxins that are composed of structurally independent A (enzymatic) and B (targeting) subunits (4, 7). The refined crystal structures of other A-B toxins including *Pseudomonas aeruginosa* exotoxin A, *Escherichia coli* heat-labile enterotoxin (LT), *E. coli* verotoxin-1 (B-pentamer only), and diptheria toxin have recently been reported (8-11). The A and B cistrons of the chromosomal CT gene are 75% and 77% homologous, respectively with those of the plasmid-borne LT enterotoxin. Differences in the inferred amino acid sequences are greatest near the A1-A2...
cleavability where the homology drops to 33% between amino acids 189 and 212 (4, 12).

As anticipated from the high sequence homology (Figure 1), the three-dimensional structures of CT and LT (9) are quite similar. In both cases the wedge-shaped A1 subunit is held above the pentameric ring of B subunits by the tethering A2 chain (Figure 2). However, notable differences occur in the arrangement of the putative ganglioside-binding site and in the conformation of the A2 chain. Surprisingly, comparison of the nicked CT and the uncleaved LT reveals little structural change at the A1-A2 junction.

OVERALL STRUCTURE

The crystal structure of nicked cholera holotoxin (choleragen) was solved by combining phases from molecular replacement using the previously determined B pentamer (choleragenoid (13)) and additional multiple isomorphous replacement (Table 1). The refined model includes all protein atoms with the exception of three residues at the A1-A2 cleavage site (residues A:193-195). The overall quality of the electron-density map is high (Figure 3), however, the density for residues A:192 and A:237-240 is fragmented. Nicking of the A chain in the crystallized toxin is supported by the large distance (9 Å) and the improbable geometry between the observed A1 carboxyl terminus (Arg192) and the amino-terminus of A2 (Ser196). The two chains of the nicked A subunit remain held together by extensive noncovalent forces and a single intrachain disulfide bond (Cys187=Cys199). The crystal structure of CT is well hydrated with over 100 obvious solvent molecules discernible at the current resolution. The three-dimensional coordinates agree well with previous low-resolution models derived from electron microscopy (14), atomic force microscopy (15), and infrared spectroscopy (16).
THE B SUBUNIT

As discussed in an accompanying report (13), the B subunits of CT form a highly stable pentamer (choleragenoid) with nearly perfect fivefold symmetry. The central pore (11 Å diameter x 40 Å long) is lined by five amphipathic α-helices that are intimately involved in pentamer stabilization. The pentamer contains five equivalent GM1 binding sites with published dissociation constants ranging from $10^{-9}$ to $10^{-10}$ M (17). Small angle X-ray diffraction studies suggest that the negatively charged pentasaccharide of GM1 (Gal-β1-3-GalNAc-β1-(NeuAc-α2-3)-4Gal-β1-4-Glc-1-ceramide) extends outwards from the cell membrane, reaching a distance of 21 Å away from the lipid/water interface (18). Co-crystals, respectively of LT with lactose (Gal-β1-4-Glc), and choleragenoid with galactose, have confirmed that a small cleft adjacent to B:Trp88 is the likely site for binding of the distal saccharides (19, 20). However, there remains considerable controversy over whether choleragen binds with the A subunit towards or away from the membrane surface (4).

THE A SUBUNIT

The A subunit of CT is translated as a single chain of 240 amino acids that is nicked by a bacterial endoprotease to form two chains (A1 and A2) associated by an interchain disulfide bond and multiple weaker interactions (27). Full activation of the toxin requires reduction of this disulfide bond and dissociation of A1 (5).

The A1 chain catalyzes the reaction $[\text{NAD}^+ + \text{Acceptor} \rightarrow \text{ADP-ribose-Acceptor} + \text{Nicotinamide} + \text{H}^+]$ and is both an ADP-ribosyltransferase and an NAD-glycohydrolase (22). It is organized into three distinct substructures (Figure 4). The first 132 amino acids form a compact globular unit composed of a mixture of
α-helices and β-strands (A1₁). The N-terminal nitrogen, absent from the crystal structure of LT, hydrogen bonds with the backbone carbonyl oxygen of residue 153 stabilizing a helix-strand transition. Catalysis is thought to occur in a well-defined cleft on the free surface of A1₁ (Figure 5). This cleft is remote from both the A1₁-A2 and A-B interfaces and is presumably the binding site for both NAD and substrate (23).

A second globular substructure (A1₂) is formed from the 31 carboxyl terminal residues of the A1 chain. This substructure surrounds the disulfide bridge linking A1 to A2 and is notable for a high density of hydrophobic residues including a cluster of four prolines (Pro168, Pro169, Pro184, and Pro185) and two tryptophans (Trp174 and Trp179). The distal A1₂ chain is quite flexible and partially disordered in the crystal structure as it approaches the nick site located along a remote free edge. Comparison of the crystalline coordinates of the nicked CT and the unnicked LT (9) suggests that this modification does not, in itself, significantly perturb the local protein structure.

The A1₃ substructure comprises the intervening 29 residues (residues 133 to 161) that form a suspension bridge between the globular A1₁ and A1₂ constructs. The A1₃ linker extends 23 Å from the distal free face of A1₁ near the catalytic site to the A1₁-A1₂ interface. Its structure suggests a functional role similar to that ascribed to the A2 chain, i.e., as a molecular tether. The A1₁-A1₂ interface is almost entirely hydrophobic with only a few polar interactions on the molecular surface.

The A2 chain anchors the enzymatic A1 chain to the B pentamer. The A2 chain of LT consists of three discrete segments; a long amino-terminal helix (residues 197 to 224), a length of extended chain that passes through the pore of the B pentamer (residues 225 to 231), and a small carboxyl-terminal helix (residues 232 to 236⁺) (19). In contrast, the A2 chain of CT consists of a nearly
continuous α-helix broken only by a kink, or bend, of 52° at residue 228. This kink is stabilized by a hydrogen bond between the γ-oxygen of Ser228 and the peptide nitrogen of Asp229. This kink is significant because it redirects the helix prior to its descent into the pentamer pore. The sequence of the last four residues of the A2 chain (KDEL) mimics that of an endoplasmic retention signal (HDEL)(24). These residues lie outside the ventral opening of the central pore with little or no stabilization by the B subunits. The carboxyl terminal of LT is completely disordered in its crystal structure, but it is clearly visible in CT (Figure 6). This presumably is due to the more compact nature of the A2 chain of CT (α-helix) relative to that of LT (extended chain). This permits the KDEL residues of CT to lie closer to its pentamer. Deletion of these residues has little effect on pentamer formation of the B subunits, but loss of the carboxyl-terminal KDEL sequence reduces the stability of the A1-pentamer complex (25).

Despite the relatively small size of the A2 chain, it shares an extensive interface with A1 (Figure 4). The long N-terminal A2 helix (residues 196-228) lies in a shallow groove that extends from one corner of the A1 subunit to the A1-B interface. The carboxyl terminus of the A1 chain and the N terminus of the A2 chain are partially disordered in the electron-density map, suggesting that this region contributes minimally to the stability of the nicked A subunit. Since the interactions throughout the length of the A1-A2 interface appear relatively tenuous, the disulfide bridge between A1(187) and A2(199) probably serves as the primary anchor between the two chains.

THE A SUBUNIT / B PENTAMER INTERFACE

The arrangement of the B subunits in choleragen differs only modestly from that described for choleragenoid (13). Introduction of the A chain affects only the B side chains that extend into the central pore of the pentamer or that lie
directly at the A-B interface. The presence of the A subunit does not appear to have any effect on the putative ganglioside binding site which lies on the remote outer edge of the B subunit.

Few direct stabilizing interactions occur between the A1 chain and the B pentamer. Three arginine side chains (Arg33, Arg143, and Arg148) located along the ventral surface of the A1 subunit, however, form multiple hydrogen bonds with groups (O78, O79, Oe1 79 and Oe2 79) located along the top of the pentameric α-helices. Since these contacts are all located adjacent to the central pore, substantial rotation of the A subunit with regard to the B pentamer is possible without disruption of the interface (26).

In contrast to the A1 chain, the A2 chain interacts with all five B subunits. The A2 subunit passes through the central pore of the B pentamer either as a continuous helix (CT) or as an extended chain anchored at its carboxyl terminus by a short turn of helix (LT). The pore diameter is just sufficiently wide to accommodate the A2 chain as a helix, therefore, the binding of CT's A subunit displaces most of the solvent molecules noted to fill the choleragenoid pore (13). Stabilizing contacts within the pore between the A2 chain and the B subunits are largely hydrophobic with very few specific hydrogen bonds noted at the current resolution. However, several B side chains that form intersubunit contacts in choleragenoid are recruited to hold the A2 chain. The decreased stability of the B pentamer in the presence of the A subunit noted experimentally is likely to be a consequence of these latter rearrangements (27).

CONCLUSIONS

The crystal structure of CT provides a detailed structural context for a rich collection of biochemical, genetic, and physical studies. The arrangement and contacts between subunits suggests that the B pentamer serves as a targeting
device for delivering the A1 chain to susceptible cell surfaces. Unfortunately, the crystal structure alone provides no direct indication for the mechanism by which the enzymatic fragment enters the cytoplasm. Co-crystals of CT with intact GM1 and substrate analogues are an obvious next step to defining the molecular basis of toxicity.
Table 1. Multiple isomorphous replacement (MIR) of intact cholera toxin. Crystals of choleragen (0.3 x 0.3 x 0.3 mm) were grown from batches of freshly isolated, isoelectrically pure cholera toxin (28). The crystals are of space group \( P2_1 \) \((a = 73.0 \, \text{Å}, \ b = 92.2 \, \text{Å}, \ c = 60.6 \, \text{Å}, \ \alpha = 90.0^\circ, \ \beta = 106.4^\circ, \ \gamma = 90.0^\circ)\), with one AB\(_5\) hexamer in the asymmetric unit. Native diffraction data were collected to a nominal resolution of 2.5 Å with a Siemens-Xentronics multiwire proportional counting detector using graphite monochromated radiation from a Rigaku RU-200 rotating anode generator set at 5 kW. Molecular replacement with the computer program Merlot (29) permitted the refined coordinates of the choleragenoid B pentamer to be correctly positioned in the \( P2_1 \) cell. Conservative rigid body and least-squares refinements (30, 31) of the molecular replacement model reduced the crystallographic R-factor below 30%, but electron-density maps calculated from the inverse Fourier transformation of the partial structure failed to provide an unambiguous interpretation of the A subunit. Three heavy atom derivatives were subsequently prepared by conventional soaking techniques. Heavy-atom binding sites and occupancies were refined with a modified version of PHARE (32) that (i) used anisotropic temperature factors to better resolve closely spaced or clustered sites, and (ii) reduced the bias contributed by a particular heavy atom to its own refinement. Although MIR phasing alone only gave a figure of merit of 45% at 2.7 Å, combination of MIR phases with those from molecular replacement of the B pentamer provided an electron-density map in which the A subunit was readily interpretable. An iterative process of simulated annealing refinement of the initial AB\(_5\) model with X-PLOR (31), followed by density modification through solvent flattening (33), improved the map quality sufficiently to visualize most side chains. Several small areas of ambiguous density benefitted from comparison with the published structure of heat-labile enterotoxin (9). After
several rounds of manual rebuilding and least-squares refinement with PROFFT (30), the crystallographic R-factor for all data from 10 to 2.4 Å now stands at 18.5%. On average, bond lengths, bond angle distances, and planarity deviated from ideal values by less than 0.015, 0.032, and 0.045 Å, respectively.
Figure 1. Comparison of the amino acid sequences of the A subunits of cholera (CT) and heat-labile (LT) enterotoxins (12). The LT isozyme shown here corresponds to that present in the crystal structure (9). Asterisks occupy positions of sequence identity; secondary structure for CT is noted below the corresponding sequence segment (34). Acidic residues (Asp and Glu) appear in red, basic residues (Lys, Arg, and His) are shown in blue, and regions not observed in the respective crystal structures are shown yellow. Solid squares denote residues implicated in catalysis (6, 8, 21, 23, 35); the single disulfide bridge is shown by a dotted line.
Figure 2. Overview of the crystal structure of cholera toxin.

(A) Alpha-carbon trace of the nicked AB5 toxin emphasizing functional features. The green side chains of B:Trp88 mark the position of the ganglioside binding site on each of the yellow B subunits. The A1 (blue) and A2 (magenta) chains are shown joined by their single disulfide bridge (yellow). Three residues of the A1 chain (Arg7, Ser61, and Glu112) are shown in gold to indicate the crevice where NAD binding and catalysis is thought to occur (6, 8, 21, 23, 35). Amino and carboxyl termini are colored cyan and red, respectively. Note the limited contact surface between the A1 chain and the B pentamer and the tethering action of the A2 chain.

(B) Representation of CT similar to that shown in (A) but rotated 90° to reveal the ventral surface. The five long α-helices (residues B:58-79) surrounding the central pore of the B pentamer form a tight cage around the interposed A2 α-helix. Residues 236-240 of the A2 chain emerge from the ventral surface of the B pentamer but are partially disordered in the crystal structure.
Figure 3. Representative electron-density for cholera toxin. Stereo view of the $2F_o - F_c$ electron-density map at the junction between the A subunit and the B pentamer. The long A2 $\alpha$-helix (orange) can be seen as it begins its descent into the central pore. Residues belonging to the A1 chain or to the B pentamer are indicated (Pro120 and Ala475, respectively).
**Figure 4.** View of the A subunit demonstrating the two A chains and their contact surface. The orientation chosen is similar to that shown in Figure 2(A). The A1 chain can be divided into three substructures; a globular amino terminus that includes the catalytic residues (A1₁-blue), a linking bridge of residues (A1₂-gold), and a second globular region (A1₃-green). The A2 chain (magenta) is composed of two long helices that form an angle of 52°. The N-terminal helix of A2 lies in a shallow groove along the ventral surface of the A1 chain and is anchored proximally by a disulfide bridge (yellow). The second A2 helix lies almost entirely within the central pore of the B pentamer.
Figure 5. Schematic stereo view of the putative catalytic site. The view presented is from the "top" looking down as defined in Figure 2. The residues implicated in catalyzing the NAD-dependent ADP-ribosylation of substrate line the left wall of a narrow (7 Å) cleft at the top of the A1 subunit. The cleft lacks the traditional βαβ NAD-binding fold. Both Glu112 and Ser61, which share a side-chain hydrogen bond, are important for catalysis (6, 8, 21, 23, 35). The side chain of Arg7, which forms the floor of the cleft, hydrogen bonds to the backbone carbonyl oxygen of Phe52 and the side-chain carboxylate of Asp9. Replacement of Arg7 with a lysine prevents catalysis; the shorter lysine side chain could fail to hydrogen-bond with the normal partners and instead interact with Glu112. The left wall of the crevice is composed of a short stretch of charged α-helix (residues 65 to 77).
Figure 6. Carboxyl terminus of A2 peptide: KDEL.

Electron density map is clearly interpretable but fragmented in places. These four residues are $\alpha$-helical and constitute the end of the long (229-240) carboxyl terminal helix of A2 which passes through the central channel of the B-subunit pentamer.
References:


32. J. E. Ladner and G. Bricogne, version 9A2 of PHARE kindly provided by the authors (1974); Z. Otwinowski, manuscript in preparation.


Acknowledgements: The coordinates will be deposited in the Brookhaven Protein Data Bank. We wish to thank Drs. Graham Shipley, Prakas Maulik, and Robert Reed of Boston University for collaborating on the structure of choleragenoid (B pentamer). Dr. Paul B. Sigler, of Yale University, generously provided computational and graphics facilities during a portion of this work. These studies were supported by the NIH under grant R01-Al28535 and by the U.S. Department of Energy, Office of Health and Environmental Research, under contract W-31-109-ENG-38.