# The three-dimensional structure of "escherichia coli" heat-labile enterotoxin in relation to its function

Sixma, Titia Karen

## Summary

Heat labile enterotoxin (LT) from *Escherichia coli* is a major virulence factor of enterotoxigenic *E. coli* strains (ETEC), which cause a generally mild diarrhea in human beings and domestic animals. It is one of the major causes of traveller's diarrhea and as such usually harmless. In small children in third world countries, however, it can lead to dehydration and death. Through the high incidence of the disease this occasional effect still causes larges numbers of victims. Heat labile enterotoxin has a very high similarity, both in sequence and activity, to cholera toxin (CT), the major virulence factor of the cholera-causing bacteria *Vibrio cholerae* O1.

LT and CT are $AB_5$ multimer proteins. Five B subunits of 103 amino acids each, combine to pentamers. The function of these B pentamers is binding to $G_{M1}$ gangliosides on host cell membranes. This is the first step in toxin action, and a prerequisite for translocation of the enzymatic A subunit into the cytoplasm. This process probably involves endocytosis, followed by an unknown membrane translocation step. The A subunit has 240 amino acids, and for activity it has to be proteolytically nicked at approximately residue 192 and reduced at its single disulfide bridge between Cys 187 and Cys 199, to form two fragments: an A1 'enzyme' (from residue 1 to ~192) and an A2 'linker' fragment (from residue ~193 to 240). The enzymatic A1 fragment uses NAD for the ADP-ribosylation of $G_{s\alpha}$, whilst releasing nicotinamide. The ADP-ribosylation brings the GTP-binding protein into its activated (GTP-bound) state, where it continuously activates adenylate cyclase, to produce large quantities cyclic AMP. This leads to the loss of fluids and electrolytes from the cell.

Interest in the three dimensional structure of these toxins is partially based on purely scientific reasons, the wish to know more about this unusual $AB_5$ complex, about the events upon assembly, secretion, binding to the target cell, membrane translocation, and activity. But structural information has also a very practical application, in the rational design of drugs and vaccines against the disease caused by this type of toxins, as well as in studies into the possibilities to use these toxins as immune adjuvants and immuno-toxins.

This thesis describes the crystallographic three dimensional structure determination of heat labile enterotoxin, followed by a detailed analysis of the structure itself, a comparison to exotoxin A and verotoxin, and a number of structure-based functional studies on the binding of the membrane receptor and the relative flexibility of the subunits. In the general introduction (chapter 1) an overview is given of the literature about LT and CT.

The X-ray structure determination is described in chapters 2-4. Good X-ray grade crystals of LT were grown by Sylvia Pronk, but due to the fact that they showed individual differences it was difficult to solve the structure. In chapter 2 a number of possible reasons for this variability are given: changes at the subunit interface of A1 and A2 or changes at the A/B interface or deamidation of asparagines upon aging of the protein, all may have caused the variability. In addition it could be shown that methionine residues were partially oxidized, which made the heavy atom substitution lower and the high noise level relatively more important. The problem could be solved in two different ways: by using one crystal multiple times and by the use of a fresh batch of protein. In chapter 3 the multiple isomorphous replacement is described, which provided the starting phases for the structure. Density modification by an extensive averaging

and phase combination procedure was necessary for the chain tracing of the structure (described in chapter 4). The structure has been refined, initially to 2.3 Å and later to 1.95 Å, to an R-factor of 18.2% and good stereochemistry.

The structure of LT is described in chapters 3 and 4. Each B monomer consists of two small three-stranded β–sheets and two α–helices, one short and one long, with a topology similar to that of *Staphylococcus aureus* nuclease. The B subunits combine to a pentamer by forming six-stranded inter-subunit sheets with each neighbour, surrounding a central pore formed by a long helix from each subunit, while the short helices pack against the sheets on the outside. The secondary structure of the A subunit is much less orderly than that of the B subunits, with as main feature two small β–sheets, at an angle to each other, and a number of helices and loops around it. A long loop winds around the subunit and ends in the A2 fragment. This 'linker'-fragment forms a long helix in contact with the A1 fragment, and enters the pore of the B pentamer. There it converts to an extended strand which transverses this pore, to emerge at the other end again as a small helix. This helix was recognised first in the lactose bound version of the structure (chapter 5), but is also present in the native toxin (chapter 4).The A subunit interacts with the B subunits almost exclusively by the C-terminal half of the A2 fragment. The interaction hardly disturbs the five fold symmetry of the B subunits. Many of the specific contacts are localised at the point where the A2 subunit enters the pore. The pore itself is lined by a high number of charged residues of the B subunits, and it has a net positive charge. Most contacts of A2 with B within this pore, however, are water mediated, with only two A/B salt bridges in this region. Hydrophobic contacts occur mainly at both ends of the pore.

The membrane binding mode of LT and CT has been extensively studied by many groups. A major question was the orientation of the complex upon binding to the pentasaccharide of $G_{M1}$. A putative sugar binding site, given in chapter 3, could be verified in chapter 5 by cocrystallisation studies with lactose showing part of the $G_{M1}$-pentasaccharide binding site. The orientation and the location of the galactose moiety in the five binding sites on the B subunits of LT gives a clear indication of the orientation of the toxin upon membrane binding, indicating that the A subunit is most likely far from the membrane upon binding to the target cell. This gives new information to the debate on the mode of translocation of the A subunit into the target cell. Interestingly the extreme C-terminus of the A subunit is now the only part interacting with the membrane. This region may have an important role in the translocation process. The KDEL/RDEL sequence at this C-terminus in CT and LT respectively could even indicate a possible interaction with a KDEL receptor, the protein responsible for protein retention in the endoplasmic reticulum.

The active site of the protein is discussed in chapter 3 and 4. Surprisingly, the core of the A subunit structure is very similar to the enzymatic domain of another ADP-ribosylating toxin, Exotoxin A from *Pseudomonas aeruginosa*. In this area $C^{\alpha}$ atoms of 44 residues, in seven fragments, could be superimposed with an rms value of 1.5 Å, without sequence homology. One of the few identical residues is the catalytically active Glu 112 (in LT). Except for this one residue the active sites in the two toxins were very different. In LT the Glu 112 is close to another residue with importance for activity, Ser 61. These residues are on the edge of a long crevice, which may form the NAD binding site. Arg 7, a residue with importance for NAD binding, is also located in this area. In ETA this residue is a histidine, but makes the same side

chain hydrogen bonds. The residue at this position seems to have the same structural function in all ADP-ribosylating toxins.

In chapter 6 soaking studies of lanthanide ions are described. Samarium and erbium ions bind, interestingly, at the A/B interface of LT in one or two sites, causing a small rearrangement of the subunits. This binding is discussed in the light of literature studies which indicate a possible role for calcium binding of the toxin. It is clear that lanthanide binding is much stronger than calcium binding, but a weak calcium binding site at this position is possible.

The relative orientation of the A1 fragment with respect to the B pentamer is not fixed. In chapters 5-7 a number of LT crystal forms are described with different relative orientations, changing up to seven degrees. In chapter 7 a second native crystal form of LT is solved by molecular replacement. In this crystal form two LT molecules are found in the asymmetric unit, with different A/B orientations. The five structures: the three LT native structures, LT with samarium bound and LT with lactose bound, are compared. The soaking study with samarium seems to be an exception: all other structures have the same hinge area, around Gln 221 in the A2 fragment, while the C-terminus of this A2 fragment rotates with the B pentamer. The change in relative orientation does not cause any major changes in the interaction, and seems to be possible because there is a limited interaction area.

The structure of the LT B subunit is extremely similar to that in *E. coli* verotoxin (VT). This toxin also forms an $AB_5$ complex, in which the B subunits bind $G_{b3}$ sphingolipids, and an enzymatic A subunit with similarity to that in ricin. The similarity of the structures of LT and VT is remarkable since LT B (103 amino acids) is much larger than VT-B (69 amino acids). In chapter 8, a comparison of the individual subunits shows that 52 amino acids (75 % of the VT-1 B subunit sequence and 50% of LT B) can be superimposed with an rms variability of 1.29 Å on the main chain atoms, while there is no sequence identity (only 3 identical residues). Many of the functionally relevant area's of the toxin: the sugar binding site, the residues in contact with the A subunit, and the surface interacting with the membrane are different in the two toxins. The three-dimensional structures of these two enzymes have generated a large number of questions on the similarity and difference within the bacterial protein toxin families.