Comparison of the B-Pentamers of Heat-Labile Enterotoxin and Verotoxin-1: Two Structures with Remarkable Similarity and Dissimilarity†

Titia K. Sixma,‡,§ Penelope E. Stein,‖ Wim G. J. Hol,*,‡,⊥ and Randy J. Read‖

BIOSON Research Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, and Department of Medical Microbiology and Infectious Diseases, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

Received August 17, 1992; Revised Manuscript Received October 12, 1992

ABSTRACT: We have compared the B-subunit pentamers of *Escherichia coli* heat-labile enterotoxin (LT) and verotoxin-1 (VT-1). The B-subunits of these bacterial toxins of the AB_5 class have virtually no sequence identity and differ considerably in size (69 amino acids in VT-1 versus 103 in LT). They share a number of functional properties: pentamer formation, association with an A-subunit, binding to carbohydrate-containing lipids, and interaction with membranes. The structures of these proteins are very similar in some respects and very different in others. They can be superimposed with an rms deviation of only 1.29 Å on the main chain atoms of 52 amino acids (0.98 Å on 47 C^{α}). Seven out of eight secondary structure elements are retained in the two toxins; only the N-terminal helix of LT is absent in VT-1. A disulfide bridge, which is essential for pentamer formation, is found in both structures, but in slightly different locations. However, the VT-1 B-subunit is much shorter on one side of the toxin, where the proposed membrane binding site of both VT-1 and LT is located. The monomer–monomer interface in the pentamer is much larger in LT than in VT-1, making the LT pentamer more stable. The central pores have a different character, and the sugar binding sites are not conserved between the toxins. The evolutionary relationship of the toxins is discussed.

The A–B class of bacterial and plant toxins is a heterogeneous group of proteins that includes the cholera-like enterotoxins [heat-labile enterotoxins (LT[1] and LT-II) and cholera toxin], the Shiga toxin family [Shiga toxin and verotoxins (VT) or Shiga-like toxins], pertussis toxin, diphtheria toxin, exotoxin A, and ricin. Members of this class of toxins have similar bipartite structures consisting of an A-subunit that is responsible for the enzymatic activity of the toxin and a B-component that binds to cell-surface receptors. However, these toxin families have nonhomologous amino acid sequences, they differ in quaternary structure, and they act by different pathogenic mechanisms. Toxic activities mediated by the different A-subunits include ADP-ribosylation of guanyl nucleotide binding regulatory proteins (cholera toxin, LT, LT-II, pertussis toxin), ADP-ribosylation of eukaryotic elongation factor 2 (diphtheria toxin, exotoxin A), and cleavage of a specific adenine from 28S ribosomal RNA (Shiga toxin, verotoxins, ricin). The B-components also differ among toxins in their quaternary structures and receptor binding specificities. Diphtheria toxin, exotoxin A, and ricin have a monomeric B-subunit. The pertussis toxin B-component is a heteropentamer made from four different subunits (single copies of S2,S3,S5 and two copies of S4). The B-components of the cholera toxin and Shiga toxin families are pentamers of identical subunits. The B-pentamers of cholera toxin and LT interact specifically with G_{M1} gangliosides [reviewed in Moss and Vaughan (1988)], while those of Shiga toxin and verotoxins bind to G_{b3} and G_{b4} [reviewed in Brunton (1990)].

The crystal structures of LT from a porcine variant of *Escherichia coli* (Sixma et al., 1991, 1992) and of the B-pentamer of verotoxin-1 (VT-1) (Stein et al., 1992) from *Escherichia coli* have both recently been solved by multiple isomorphous replacement. The LT structure has been refined to an *R*-factor of 18.2% at 1.95 Å (Sixma et al., submitted for publication) and the VT-1 B_5-subunit structure to an *R*-factor of 17.7% at 2.2 Å. The B-subunits of these toxins are different in size (103 residues for LT versus 69 residues for VT-1) and have nonhomologous amino acid sequences, but unexpectedly they share the same fold (Figure 1). Both B-subunits form pentamers in which the basic monomer consists of two three-stranded antiparallel β-sheets and an α-helix. β-Sheets from pairs of adjacent monomers form six-stranded antiparallel β-sheets around the outer surface of the pentamer, while the five helices line a "pore" in the center. The LT B-monomer has an additional α-helix at its N-terminus that is absent in VT-1. This close structural similarity suggests that the Shiga toxin and cholera toxin families may have evolved from a common ancestor, although the amino acid sequences of the two toxins show no clear evidence for such an evolutionary relationship. The basic monomer fold of the LT and VT-1 B-subunits has also been found (A. Murzin, personal communication; Sixma et al., 1992) in the monomeric protein *Staphylococcus aureus* nuclease (Arnone et al., 1971). This extracellular nuclease shows no sequence homology or functional similarity with either toxin.

Overall structural similarity between proteins which have nonidentical amino acid sequences may be explained either by divergent evolution from a common ancestor or by convergent evolution. Comparative analysis of the structures of families of proteins related by divergent evolution shows

† This project was supported by the Medical Research Council of Canada and the Alberta Heritage Foundation for Medical Research (to R.J.R. and P.E.S.), an International Reserach Scholars Award from the Howard Hughes Foundation (to R.J.R.), the Netherlands Foundation for Chemical Research (SON), and the Netherlands Organization for Scientific Research (NWO) (to T.K.S. and W.G.J.H).

* Author to whom correspondence should be addressed.

‡ University of Groningen.

§ Present address: Department of Molecular Biophysics and Biochemistry, Yale University, BCMM 154, 295 Congress Ave., New Haven, CT 06510.

‖ University of Alberta.

⊥ Present address: Department of Biological Structure SM-20, School of Medicine, University of Washington, Seattle, WA 98195.

[1] Abbreviations: LT, heat-labile enterotoxin; VT-1, verotoxin-1; rms, root mean square.

0006-2960/93/0432-191$04.00/0

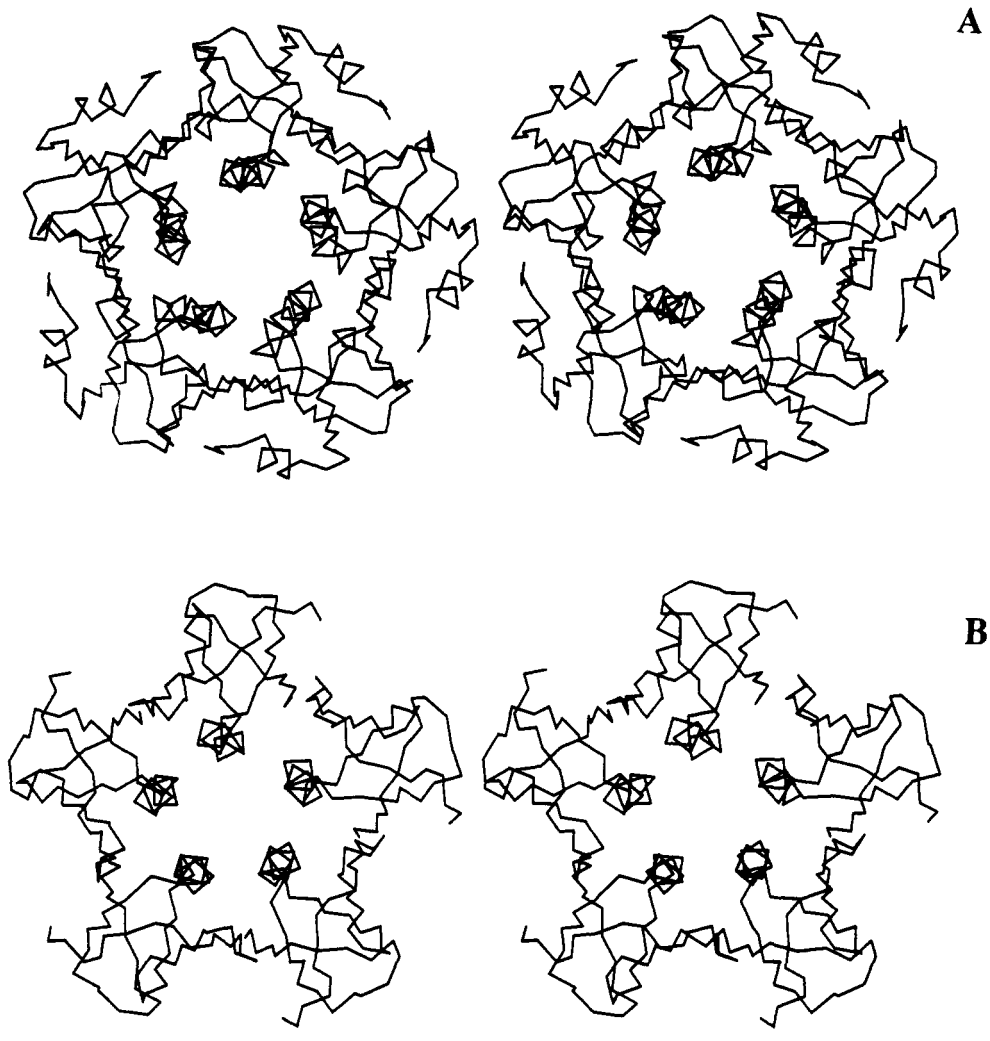


FIGURE 1: Stereoviews of C^{α} tracings of the B-subunit pentamers of LT viewed along the 5-fold rotation axis (A) and VT-1 viewed along the 5-fold screw axis (B). The side closest to the viewer is the A-subunit binding side. Monomers are numbered clockwise around the pentamer with monomer 1 at the top.

that the extent of structural change is typically related to the extent of sequence change (Chothia & Lesk, 1986). However, several families of proteins that have evolved under selective constraints on function show remarkable structural adaptation to sequence divergence, as in the globin, cytochrome, and immunoglobulin families (Lesk & Chothia, 1980; Hubbard & Blundell, 1987).

While overall similarity of protein backbone fold usually results from divergent evolution, convergent evolution typically involves a particular motif that forms only part of the structure. The serine proteases subtilisin BPN′ and chymotrypsin provide a well-known example of convergent evolution. Their active sites and catalytic mechanisms are similar, but their amino acid sequences and overall secondary and tertiary structures are unrelated (Kraut et al., 1971).

It is sometimes difficult to decide if structural similarity between proteins results from divergent or convergent evolution. The TIM barrel structural motif was found first in triosephosphate isomerase (Banner et al., 1975) and subsequently in many other enzymes that have nonhomologous sequences and different functions. In different TIM barrel enzymes, the active-site residues are found in different positions in the primary sequence. This may be an example of convergent evolution to produce a stable fold (Goldman et al., 1987). However, Farber and Petsko (1990) classified 17 TIM barrel structures into 4 groups on the basis of their crystal structures and suggest that cyclic permutation of the gene could provide a mechanism for divergent evolution from a single ancestor. See, however, Brändén (1991) for a different point of view. Unexpected structural similarity with virtually no sequence homology has also been found between methylamine dehydrogenase (MADH) and neuraminidase (Vellieux et al., 1990), between ubiquitin and ferredoxin (Vriend & Sander, 1991), and between muscle actin (in a complex with deoxyribonuclease I) and the N-terminal domain of the 70-kDa bovine heat-shock protein (Kabsch et al., 1990).

The conventional way of detecting families of related proteins is to search for amino acid sequence homology. This works well for closely related sequences, but may fail to detect sequence identity below about 25%. More complex methods have been described to detect distantly related proteins. These include profile methods in which the amino acid sequences, and sometimes also the three-dimensional structures, of members of a family of proteins are used to make a "fingerprint" that will detect other related proteins. An early example is a fingerprint concerning NAD binding proteins (Wierenga et al., 1986). These methods are most successful when the fingerprint incorporates information about the environment of residues in the structure, including their solvent accessibilities and secondary structure locations (Bowie et al., 1991; Lüthy et al., 1991; Overington et al., 1992).

FIGURE 2: Stereoview of superposed C^{α} tracings of representative B-subunit monomers of LT (thick lines) and VT-1 (thin lines), viewed perpendicular to the 5-fold axis. The N-terminus (N) and C-terminus (C) of LT are marked.

We have carried out a detailed structural comparison of the B-subunits of LT and VT-1 to try to understand the evolutionary relationship between these proteins and the possible functional significance of their close structural similarity.

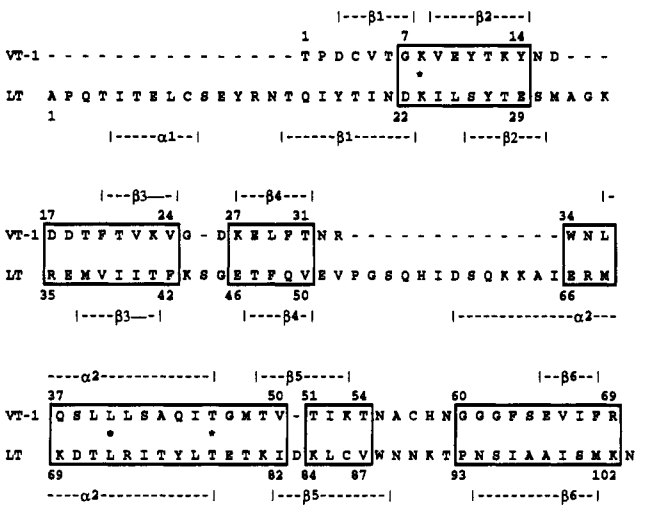
RESULTS

Monomer Superposition. Equivalent secondary structure elements of representative B-subunits of VT-1 (monomer 3) and LT (monomer 2) were superimposed by two different methods: first, by "hand", choosing equivalent residues by visual inspection, followed by optimization of the superposition with the superimpose option of INSIGHTII (software program from Biosym Technologies of San Diego); second, by automated superposition using the three-dimensional alignment option of WHATIF (Vriend & Sander, 1991). These methods gave essentially the same result (Figure 2). A total of 208 equivalent main chain atom pairs in 6 different segments (Table I) were chosen to define the alignment. The rms difference for superposition of these 208 atom pairs between the two structures was 1.29 Å (1.17 Å for C^{α} atoms in 52 residues). Optimization of the superposition by removing the worst agreeing atoms (statistical outliers) gave an rms difference of 0.92 Å for 178 main chain pairs (0.98 Å for C^{α} atoms in 47 residues).

The B-subunit sequences of VT-1 (Seidah et al., 1986) and LT (Dallas & Falkow, 1980; Yamamoto et al., 1987) were aligned on the basis of this structural superposition (Table I). The alignment showed only 3 pairs of identical residues among the 52 aligned pairs (VT-LT pairs: Lys8–23, Leu40–72, and Thr46–78). No relation between the sequences could be identified by a method generating suboptimal alignments (Zuker, 1991). Adding information about the three-dimensional environment of side chains in the structure also did not result in a significant score for the comparison of LT and VT-1 sequences (Bowie et al., 1991). Apparently, environments of the structurally equivalent regions are not sufficiently comparable for this method to work.

The alignment of LT and VT-1 (Table I) coincides only partially with the secondary structure elements in the two

Table I: Alignment of VT-1 and LT B-Subunits According to Structural Superposition[a]



[a] Boxes indicate residues used in superposition. Asterisks are given for the three identical residues in the two toxins.

structures. The secondary structure elements of VT-1 are shorter than their equivalents in LT, but a complete VT-1 B-subunit can be closely fitted to the "lower" part of an LT monomer (as viewed in Figure 2). In this orientation, the upper surface of the structure is thought to be the membrane binding side (Sixma et al., 1992; Stein et al., 1992), while the lower surface interacts with the A-subunit, in LT (Sixma et al., 1991) and in Shiga toxin (M. E. Fraser and M. N. G. James, personal communication; note that the B-subunit sequence of Shiga toxin is identical to that of VT-1). The N-termini of the two structures differ significantly: LT has an N-terminal α-helical extension that is absent in VT-1. Strands $\beta 2$, $\beta 3$, $\beta 4$, and $\beta 6$ and helix $\alpha 2$ of VT-1 can be closely aligned with the parts of the equivalent (longer) structures in LT that lie near the monomer surface. The rms difference for main chain atoms in each of these closely matched segments is around 1 Å (Table II). The conformations of the turns

Table II: Rms Difference in Positions of Equivalent Main Chain Atoms (N, C^{α}, C, O) in Secondary Structure Elements of LT and VT-1[a]

secondary structure	residues in LT	residues in VT-1	no. of atoms	rms difference (Å) local alignment	rms difference (Å) global alignment
strand β1	19–23	4–8	20	1.13	2.86
strand β2	24–28	9–13	20	0.34	0.65
strand β3	38–41	20–23	16	0.35	0.87
strand β4	46–50	27–31	20	0.54	0.78
helix α2	68–78	36–46	44	0.25	0.92
strand β5	81, 82, 84–86	48, 49, 51–53	20	1.25	2.58
strand β6	98–101	65–68	16	0.43	1.20

[a] See Table I for residues used to obtain the global alignment.

Table III: Comparison of Turn Conformations (Crawford et al., 1973) between Adjacent Antiparallel β-Strands in Representative B-Monomers of VT-1 (Monomer 3) and LT (Monomer 2)

	residues		turn	main chain torsion angles				turn between
	i	*i*+3	type	ϕ_2	ψ_2	ϕ_3	ψ_3	strands
VT-1	15	18	I	−61	−28	−84	0	β2–β3
	24	27	II′	57	−124	−107	4	β3–β4
	58	61	II	−55	140	77	3	β5–β6
LT	32	35	II	−49	127	71	9	β2–β3
	42	45	III	−67	−12	−67	−20	β3–β4
	88	91	?	−86	5	−101	25	β5–β6

between adjacent antiparallel β-strands differ in the two structures (Table III).

Construction of a Perfect VT-1 B-Pentamer. Lack of 5-fold symmetry in the VT-1 B-subunit crystal structure was a problem when trying to superimpose the structures of the VT-1 and LT B-pentamers. In LT, the B-pentamer is nearly perfect, and neighboring monomers are connected by β-sheet interaction (Sixma et al., 1991). This β-sheet interaction exists for only four of the five neighboring pairs in the VT-1 B-pentamer. This is due to a screw component in the 5-fold axis (Stein et al., 1992), which creates a lock–washer structure having a gap between monomers 1 and 2. The screw component was defined by superimposing α-carbon atoms from two copies of the VT-1 B-pentamer (monomer 1 on 2, 2 on 3, etc.) using the method of Kabsch (1976) followed by that of Rao and Rossmann (1973) as implemented in the computer program SUPPOS, from the Groningen BIOMOL package. In iteratively discarding outlier coordinate pairs, all 69 monomer 1 on 2 pairs, and 20 other pairs, mostly from surface loops were rejected. The resulting coordinate transformation consists of a rotation of 72.3°, plus a translation of −1.32 Å along the rotation axis.

The lack of symmetry in the VT-1 B-subunit crystal structure is probably an artifact of crystallization. In the crystal structure of Shiga toxin, the B-subunit pentamer is much closer to having perfect 5-fold symmetry (M. E. Fraser and M. N. G. James, personal communication). A "perfect" VT-1 B-pentamer, which was felt to be more likely to represent the in vivo structure, was made to use in some comparisons with LT. The construction of the perfect pentamer proceeded in two stages. First a rough model was constructed manually; then it was optimized using the program X-PLOR (Brünger et al., 1987). Monomer 4 was chosen as the basis of the perfect pentamer, since it is furthest from the 1–2 interface. First it was reoriented so that it was centered on the *x*-axis, with the 5-fold screw along the *z*-axis. If a second monomer were generated at this point by a 72° rotation about the *z*-axis, it would be too high in *z* to make the correct β-sheet interactions with the first monomer. In order to form these interactions after a pure 5-fold rotation, the monomer was rotated to change the relative *z*-value of the two interacting strands by 1.3 Å. Since the interacting β-strands are about 18 Å apart in the monomer, this required a rotation of approximately 4.2° about the *x*-axis. The rough model of a perfect pentamer was then generated by 5-fold rotations about the *z*-axis. Finally, X-PLOR was used to optimize the perfect pentamer by rigid-body refinement, minimizing the deviations of the β-sheet interactions from the values observed in the crystal structure for monomers 4 and 5. The rms deviation from the target values was 0.03 Å. Although no other restraints were applied, the resulting model is 5-fold symmetric and has no serious van der Waals clashes.

Pentamer Superposition. The LT B-subunit structure was superimposed on the perfect VT-1 B-pentamer structure using the 208 main chain atoms in each monomer that were used in the monomer superposition (1040 atoms in total). The rms difference for superposition of these 1040 atom pairs between the 2 structures was 1.84 Å. Removing the worst agreeing atoms gave an rms difference of 1.74 Å for 1009 atoms.

Superposition of the VT-1 and LT B-pentamers agrees with the results obtained from superimposing the individual monomers (Figure 3). The structures superimpose closely at their "lower" surface, but the LT structure extends beyond the VT-1 structure at the "upper" surface (in the orientation of Figure 2). The β-sheet interaction between adjacent monomers is similar in both structures, but in LT, the N-terminal helical extension (α1) buries this interface, whereas the shorter N-terminus in VT-1 causes the surface to be exposed.

Pentamer Stability. The amount of surface buried in formation of the pentamer (calculated using the program DSSP; Kabsch & Sander, 1983) is quite different in the two toxins: in LT, each subunit interaction buries about 2710 Å^2 (1355 Å on each side), whereas in VT-1, the four regular subunit interactions each bury about 1320 Å^2. About half the difference in buried surface can be attributed to interactions provided by the N-terminal extension on the LT B-subunit; when the first 15 residues are omitted from the calculation, the surface area buried by each subunit interaction in LT drops to about 2000 Å^2. The remainder is due to the difference in length of secondary structural elements, particularly the helices lining the pore. This difference in interaction area is consistent with the difference in stability of LT and VT-1 (Gill & King, 1975; Ramotar et al., 1990).

The presence of the disulfide bridge is necessary for pentamer formation in LT (Jobling & Holmes, 1991) and in Shiga toxin (Jackson et al., 1990), but it is, however, in a slightly different location in the two proteins (see Table I). In the monomeric protein *Staphylococcus aureus* nuclease (Arnone et al., 1971), no such disulfide is present. The topology of the first 100 residues of this molecule is identical to that of LT and VT-1, but the optimal superposition is not nearly as clear as in the LT/VT-1 comparison. This is mainly due to different orientations of secondary structural elements. The closer similarity of LT and VT-1 monomers may be because they both have to fit in a pentamer, and are stabilized by interactions from their neighbors.

The Central Pore. A distinct common feature of the B-subunit structures of LT and VT-1 is the central pore formed by the five long helices. This region has functional importance for the assembly of the toxin, both for pentamer formation, by side chain packing, and for the interaction with the A-subunit, as seen in the structure of LT. At the same time,

FIGURE 3: Stereoview of the C^{α} tracing of the B-subunit pentamer of LT (thick lines) superposed on the main chain tracing of the "perfect" B-subunit pentamer of VT-1 (thin lines), viewed along the 5-fold rotation axis.

it presents some of the more striking examples of dissimilarity between the two structures. The helices are much larger in the LT B-subunits (~19 amino acids) than in VT-1 (~12 residues). The overlapping region is at the C-terminal half of the LT helix (VT, 34–47; on LT, 66–79), the part of the helix which is closest to the A-subunit in LT. The axes of the VT-1 helices in the perfect pentamer differ in orientation by only 8–11°, as calculated with the BIOMOL program SECSTR (Ploegman et al., 1978). In LT, the helices are curved, with an average difference in orientation of about 19°; in the N-terminal half (residues 60–69), this is about 24–27°, and in the C-terminal half (residues 68–78), it is about 12–15°. The difference between LT and VT-1 is therefore small in the overlapping region, and this small difference may be caused by errors in creating the "perfect pentamer". The radius of the pore, when measured as the distance of the helix axis to the 5-fold axis, varies between 10.4 and 12.5 Å in LT (going from the N- to the C-terminus of the helix) and between 10.4 and 8.5 Å in VT-1. If the pore size is measured as the distance from side chain atoms to the 5-fold axis, both pores have a radius of between 5 and 6 Å, except in the vicinity of residue Trp34 in VT-1. Since Trp34 adopts three different conformations in the five VT-1 B-monomers, it should not be considered to constrict the pore permanently.

The major difference between the pores of the LT and VT-1 pentamers resides, however, not in the size but in the character of the side chains lining the walls. In LT, all side chains pointing into the pore are charged except for the C-terminus turn of the helix (Figure 4a). This pore is therefore very highly charged with in total 40 charges in alternating positive and negative layers, but with an excess of positively charged residues. Many salt bridges are formed between B-subunit amino acids, and there is a large number of bound water molecules visible in the electron density in the pore. In contrast, the central pore in verotoxin has no charges at all within the helix: the exposed residues are Asn35, Ser38, Ser42, Ile45, and Thr46 (Figure 4B).

It can be concluded that the character of the central helix pore is not of major importance for a property shared by these two toxins. The notably charged character of the surface does not seem to be essential for pentamer stabilization.

Architecture of the AB_5 Complex. In both LT (Sixma et al., 1991) and Shiga toxin (M. E. Fraser and M. N. G. James, personal communication), the A-subunit is located at the flat surface of the B-pentamer, at the C-terminal end of the long helices. Both toxins have an A2-fragment, that is inserted

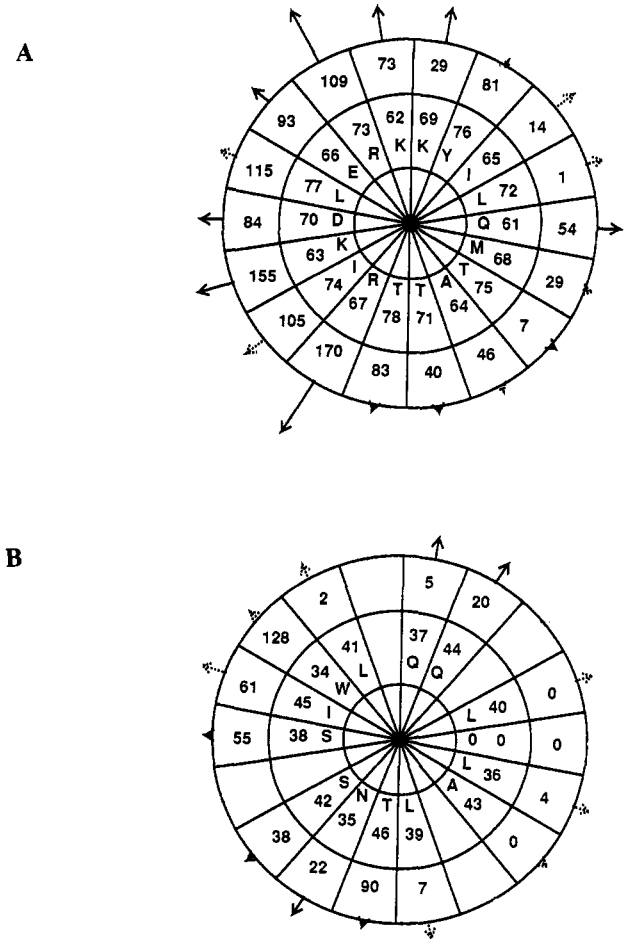


FIGURE 4: Helical wheels showing the remarkable difference in character of the pore-forming central helices of (A) LT and (B) VT-1. Wheels were created with the program Hera (Hutchinson & Thornton, 1991), showing in the inner circle the residue numbers, next the residue name in one-letter code, and in the outer circle the accessibility according to DSSP (Kabsch & Sander, 1983). The arrows indicate the hydrophobicity (dotted lines) or hydrophilicity of the amino acid type.

into the central pore. In at least LT, the A2-fragment passes completely through the pore, continuing on the other, convoluted, side as a small helix. The last four residues, RDEL (KDEL in cholera toxin), are invisible in the electron density map, probably due to flexibility. This region is thought to interact with the membrane after binding (Sixma et al., 1992). Surprisingly, this is a sequence that has been identified as an

endoplasmic reticulum retention signal (Pelham et al., 1989), but its importance in LT and cholera toxin has yet to be proven. The C-terminus of the VT-1 A-subunit ends in the sequence RRTISS, very different from a "KDEL"-like sequence except that it is also quite polar. Even if this C-terminal region interacts with the membrane, it may be in a quite different way than in LT and CT.

Sugar Binding and Membrane Interaction. LT and verotoxin both bind to the sugar chain of a sphingolipid membrane receptor, LT binding to G_{M1} ganglioside [Galβ1–3GalNAcβ1–(NeuAcα2–3)4Galβ1–4Glc cerebroside] and verotoxin to G_{b3} [Galα1–4Galβ1–4Glc cerebroside]. The binding site of the terminal galactose in LT has been determined (Sixma et al., 1992), with residues 51, 56, 61, 90, and 91 as hydrogen bond donors and acceptors, while residues 88 and 57 have van der Waals interactions with the hydrophobic side of the sugar. Surprisingly, this region is completely absent in the verotoxin molecule after superposition. All atoms of the sugar as well as the hydrogen-bonding ligands (except Glu51) are outside the protein region superimposable on verotoxin.

However, an analogous region is present in the verotoxin secondary structure too. The galactose binding site in LT is located in a pocket which has as its base the turn between strands β5 and β6 in sheet II (residues 89–93). This turn bends away from sheet II and has main chain interaction with the other three-stranded sheet (sheet I) within the monomer. The turn connecting β5 and β6, in VT-1 (residues 56–60), is similar in structure, also connecting the sheets via main chain interactions but different in orientation (see Figure 2). Apparently, the position of this loop has changed, which may be due to the difference in exact position of the disulfide bridge between LT (Cys86, at the end of β5) and VT-1 (Cys57, within the turn). There is no clear cavity at this structural location in VT-1, but a small depression can be seen near the peptide link of Gly62–Phe63 flanked by the N-terminal Thr, region 30–33, Ala56, Glu65, and Asp16 of a neighboring subunit. It is therefore not impossible that this site in VT-1, which has the same structural function as the galactose binding site in LT, also forms part of a carbohydrate binding site. In addition, an alternative binding site has been proposed for VT-1, from mutant and structural analysis, on the side of the pentamer (Stein et al., 1992). This region in LT is occupied by the small N-terminal helix which is not present in VT-1, and it is not likely to be part of the ganglioside binding site of LT. The two proposed sites for VT-1 are, however, not necessarily mutually exclusive. A single sphingolipid molecule could potentially bind both at the side of the pentamer and in the slight cavity that is topologically equivalent to the LT galactose binding site.

Both potential glycolipid binding sites in VT-1 imply that the same surface of the pentamer interacts with the membrane. This is the surface analogous to that which is furthest from the A-subunit in LT. This surface is very different in LT and VT-1. In LT, it is quite convoluted, with the region 52–64 stretching out from the surface, whereas this loop is missing completely in VT-1. The difference in surface implies that membrane interaction may be quite different in the two toxins.

DISCUSSION

The B-subunits of heat-labile enterotoxin and verotoxin have a strikingly similar fold, with an rms value of 1.29 Å for the main chain atoms of 52 amino acids (which is 75% of the VT-1 B-chain and 50% of LT-B). These B-subunits share a number of functional properties: (1) they form pentamers; (2) they interact with an enzymatic A-subunit; (3) they recognize the oligosaccharide chain of a glycolipid membrane receptor; (4) their A-subunit needs to be internalized through a membrane. In contrast to these clear similarities, there are a number of major differences between these proteins: (i) VT-B is only two-thirds of the size of LT-B; (ii) there is no significant sequence identity (3 identical residues in 52 superimposed residues); (iii) the LT galactose binding region is missing in VT; (iv) the proposed membrane binding side of the VT-1 and LT B-pentamer is completely different in shape; and (v) the central pore, which contains the regions reponsible for interaction with the A-subunit, has a very different character.

The similarity between LT and VT-1 is noteworthy for the perfection of the superposition combined with the lack of sequence identity. Protein families with structures that can be superimposed within 2.5 Å generally have a clear sequence identity, in which closeness of the superposition of the common core can be related to the percentage sequence identity by a simple formula (Chothia & Lesk, 1986). On the basis of this empirical rule, the expected rms main chain distance for 6% sequence identity would be 2.3 Å, whereas the actual rms deviation of 1.3 Å would be expected for sequences with 37% identity. There are other protein families, however, which have clear structural similarity, but no apparent sequence identity. Within those families, though, the rms superposition is usually worse, and to align these structures, other properties must be considered (Sali & Blundell, 1990; Pastore & Lesk, 1990; Sander & Schneider, 1991; Bowie et al., 1991).

The extent of the functional resemblance between the LT and VT-1 B-monomers is not obvious. The clearest similarity is found in the β-sheet interactions, that stabilize both pentamers. Side chain interactions, however, are completely different, most notably so in the central pore, and the B-subunit interface is much more extensive in LT. Nevertheless, the A-subunits in both toxins are located on the same surface of the pentamer and interact similarly, by an A2-fragment inserted into the central pore (Sixma et al., 1991; M. E. Fraser and M. N. G. James, personal communication). Both toxins should therefore interact with the membrane through the same surface, which we believe is the side of the B-pentamer opposite to the A-subunit. However, these interactions must be very different, since both the B-subunit surface and the C-terminal region of the A2-fragments are very different in LT and VT-1. The extent of conservation of the sugar binding site is even less well established; at best, it will be located at a generally similar structural site.

The toxin B-subunit fold has now been found in at least three different proteins, two pentameric and one monomeric. All three proteins, or just the two toxin B-subunits, may have diverged from a common ancestor, or all three proteins may have converged on the same fold following different paths in evolution. There are thus two issues to address: whether *Staphylococcus aureus* nuclease has evolved by divergence from the two toxins, and whether the two toxins have evolved by divergence. There may not yet be enough information to resolve these questions, but it is intriguing to speculate.

If the evolution of *Staphylococcus aureus* nuclease converged to the same fold as the B-subunits, there must be some respect in which this fold is particularly favorable, either in its folding pathway or in its structural stability. Stability seems unlikely to have been the driving force. The B-subunits are probably stabilized by pentamer formation, and the extra two helices in the nuclease structure might well be required for stability.

A potential argument against divergent evolution to the nuclease fold is the difference in quaternary structure. However, there is sufficient precedent. For instance, in the globin family, myoglobin is a monomer while other members exhibit a bewildering variety of quaternary structures (Vinogradov, 1985). An argument in favor of divergence is the shared functional site; the enzymatic active site of *Staphylococcus aureus* nuclease is in the same location as the carbohydrate binding site of LT. If the structures are derived from divergent evolution, divergence may have taken place either from an original pentameric protein with a sugar binding site to a monomeric enzyme or vice versa. The conversion would have been accompanied by the addition (or loss) of a disulfide bond. The exact location of this bond is not conserved between LT and VT-1, but that is similar to the situation in other protein families, e.g., the family of closely related serine proteases (Delbaere et al., 1979).

Considering the striking differences between the two toxin B-subunits, one could even imagine that they have arisen by convergent evolution. As with the nuclease, this could be attributed to some intrinsic advantage to the fold. In addition, there might be some advantage to a pentameric binding component. Multivalency would multiply the binding affinity of five carbohydrate binding pockets. Also, the packing of five helices around a 5-fold axis creates a pore of appropriate size to interact with a polypeptide chain. The A–B toxins must solve the problem of forming a stable complex that allows the enzymatic component to dissociate under appropriate conditions. A strand of polypeptide fitting into a pore achieves a strong interaction with a minimal amount of extra protein. Because the rest of the A-subunit barely interacts with the B-pentamer, dissociation is readily achieved by the cleavage of the covalent bonds between the A1- and A2-subunits. It is not difficult to imagine that evolution has found this solution at least twice.

If the two B-subunits have diverged from a common ancestor, we must explain the differences in the ways they carry out their common functions. It seems likely that the site of carbohydrate binding is different in the two toxins. One could imagine a primordial protein with a sugar binding site roughly in the position of those in LT and VT-1. By a series of small mutations, the shallow sugar binding site may have been extended at one end to bind an additional sugar residue, while by a similar process it may have been shortened at the other end. By a series of such small steps, two related proteins with nonoverlapping binding sites could be generated. One might have expected the membrane binding surface to be highly conserved, but clearly this need not be so, since both toxins do bind to membranes despite their differences on this surface.

Finally, regardless of the evolutionary questions, we suggest that the formation of a central pore in the B-pentamer, with which the "linking" A2 chain interacts, is a crucial feature that is likely to be found in all of the AB_5 toxins. The LT-II's (LT-IIa and LT-IIb) are AB_5 toxins with an A-chain homologous to that of LT, and a B-chain that is similar in size but not in sequence to LT (Pickett et al., 1987, 1989). A central pore may even exist in pertussis toxin which has five B-subunits forming a heteropentamer. One should note, however, that some of these subunits are much larger than those in LT (molecular mass 12–22 kDa; Nicosia et al., 1986; Locht & Keith, 1986), so there will be substantial differences in structure even if they are related. Both LT-II and pertussis toxin have an "A2"-like region in the A and S1 subunits, respectively, that is likely to interact with the B-components.

ACKNOWLEDGMENT

We thank Glen D. Armstrong and James L. Brunton for stimulating discussions.

REFERENCES

Arnone, A., Bier, C. J., Cotton, F. A., Day, V. W., Hazen, E. E., Jr., Richardson, D. C., Richardson, J. S., & Yonath, A. (1971) *J. Biol. Chem. 245*, 2302–2316.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D., & Waley, S. G. (1975) *Nature 255*, 609–614.

Bowie, J. U., Lüthy, R., & Eisenberg, D. (1991) *Science 253*, 164–170.

Bränden, C.-I. (1991) *Curr. Opin. Struct. Biol. 1*, 978–983.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458–460.

Brunton, J. L. (1990) *Bacteria 11*, 377–397.

Chothia, C., & Lesk, A. M. (1986) *EMBO J. 5*, 823–826.

Crawford, J. L., Lipscomb, W. N., & Schellman, C. G. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 538–542.

Dallas, W. S., & Falkow, S. (1980) *Nature 288*, 499–501.

Delbaere, L. T. J., Brayer, G. D., & James, M. N. G. (1979) *Nature 279*, 165–168.

Farber, G. K., & Petsko, G. A. (1990) *Trends Biochem. Sci. 15*, 228–234.

Gill, D. M., & King, C. A. (1975) *J. Biol. Chem. 250*, 6424–6432.

Goldman, A., Ollis, D. L., & Steitz, T. A. (1987) *J. Mol. Biol. 194*, 143–153.

Hubbard, T. J. P., & Blundell, T. L. (1987) *Protein Eng. 1*, 159–171.

Hutchinson, E. G., & Thornton, J. M. (1990) *Proteins: Struct., Funct., Genet. 8*, 203–212.

Jackson, M. P., Wadolkowski, E. A., Weinstein, D. L., Holmes, R. K., & O'Brien, A. D. (1990) *J. Bacteriol. 172*, 653–678.

Jobling, M. G., & Holmes, R. K. (1991) *Mol. Microbiol. 5*, 1755–1767.

Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110–119.

Kabsch, W. (1976) *Acta Crystallogr. A32*, 922–923.

Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577–2637.

Kabsch, W., Mannherz, H. G., Suck D., Pai, E. F., & Holmes, K. C. (1990) *Nature 347*, 37–44.

Kraut, J., Robertus, J. D., Birktoft, J. J., Alden, A., Wilcox, P. E., & Powers, J. C. (1971) *Cold Spring Harbor Symp. Quant. Biol. 36*, 117–123.

Lebioda, L., Hatada, M. H., Tulinsky, A., & Mavridis, I. M. (1982) *J. Mol. Biol. 162*, 445–458.

Lesk, A. M., & Chothia, C. (1980) *J. Mol. Biol. 136*, 225–270.

Locht, C., & Keith, J. M. (1986) *Science 232*, 1258–1264.

Lüthy, R., McLachlan, A. D., & Eisenberg, D. (1991) *Proteins: Struct., Funct., Genet. 10*, 229–239.

Moss, J. & Vaughan, M. (1988) *Adv. Enzymol. Relat. Areas Mol. Biol. 61*, 303–379.

Nicosia, A., Perugini, M., Franzini, C., Casagli, M. C., Borri, M. G., Antoni, G., Almoni, M., Neri, P., Ratti, P., & Rappuoli, R. (1986) *Proc. Natl. Acad. Sci. U.S.A 83*, 4631–4635.

Overington, J. P. (1992) *Curr. Opin. Struct. Biol. 2*, 394–401.

Pastore, A., & Lesk, A. M. (1990) *Proteins: Struct., Funct., Genet. 8*, 133–155.

Pelham, H. R. B. (1989) *Annu. Rev. Cell Biol. 5*, 1–23.

Pickett, C. L., Weinstein, D. L., & Holmes, R. K. (1987) *J. Bacteriol. 169*, 5180–5187.

Pickett, C. L., Twiddy, E. M., Coker, C., & Holmes, R. K. (1989) *J. Bacteriol. 171*, 4945–4952.

Ploegman, J. H., Drenth, J., Kalk, K. H., & Hol, W. G. J. (1978) *J. Mol. Biol. 123*, 557–594.

Ramotar, K., Boyd, B., Tyrrell, G., Gariepy, J., Lingwood, C., & Brunton, J. (1990) *Biochem. J. 273*, 805–811.
Rao, S.T., & Rossmann, M. G. (1973) *J. Mol. Biol. 76*, 241–256.
Sali, A., & Blundell, T. L. (1990) *J. Mol. Biol. 212*, 403–428.
Sander, C., & Schneider, R. (1991) *Proteins: Struct., Funct., Genet. 9*, 56–68.
Seidah, N. G., Donohue-Rolfe, A., Lazure, C., Auclair, F., Keusch, G. T., & Chretien, M. (1986) *J. Biol. Chem. 261*, 13928–13931.
Sixma, T. K., Pronk, S. E., Kalk, K. H., Wartna, E. S., van Zanten, B. A. M., Witholt, B., & Hol, W. G. J. (1991) *Nature 351*, 371–378.
Sixma, T. K., Pronk, S. E., Kalk, K. H., van Zanten, B. A. M., Berghuis, A. M. & Hol, W. G. J. (1992) *Nature 355*, 561–564.
Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L., & Read, R. J. (1992) *Nature 355*, 748–750.
Stuart, D. I., Levine, M., Muirhead, H., & Stammers, D. K. (1979) *J. Mol. Biol. 134*, 109–142.
Vellieux, F. M. D., Kalk, K. H., Drenth, J., & Hol, W. G. J. (1990) *Acta Crystallogr., B46*, 806–823.
Vinogradov, S. (1985) *Comp. Biochem. Physiol. B82*, 1–15.
Vriend, G., & Sander, C. (1991) *Proteins: Struct., Funct., Genet. 11*, 52–58.
Wierenga, R. K., Terpstra, P., & Hol, W. G. J. (1986) *J. Mol. Biol. 187*, 101–107.
Yamamoto, T., Gojobori, T., & Yokota, T. (1987) *J. Bacteriol. 169*, 1352–1357.
Zuker, M. (1991) *J. Mol. Biol. 221*, 403–420.