

Refined Structure of *Escherichia coli* Heat-labile Enterotoxin, a Close Relative of Cholera Toxin

Titia K. Sixma^{1†}, Kor H. Kalk¹, Ben A. M. van Zanten¹
Zbigniew Dauter³, Jaap Kingma², Bernard Witholt^{2†}
and Wim G. J. Hol^{1†‡}

¹*BIOSON Research Institute and* ²*Laboratory of Biochemistry*
University of Groningen, Nijenborgh 4, 9747 AG Groningen
The Netherlands

³*EMBL Outstation, 2000 Hamburg 52, Germany*

(Received 24 July 1992; accepted 22 October 1992)

Heat-labile enterotoxin (LT) from *Escherichia coli* is a bacterial protein toxin with an AB₅ multimer structure, in which the B pentamer has a membrane binding function and the A subunit is needed for enzymatic activity. The LT crystal structure has been solved using a combination of multiple isomorphous replacement, fivefold averaging and molecular dynamics refinement. Phase combination using all these sources of phase information was of crucial importance for the chain tracing. The structure has now been refined to 1.95 Å resolution, resulting in a model containing 6035 protein atoms and 293 solvent molecules with a crystallographic *R*-factor of 18.2% and good stereochemistry.

The B subunits are arranged as a highly stable pentamer with a donut shape. Each subunit takes part in approximately 30 inter-subunit hydrogen bonds and six salt bridges with its two neighbors, whilst burying a large surface area. The A subunit has higher temperature factors and less well-defined secondary structure than the B subunits. It interacts with the B pentamer mainly *via* the C-terminal A2 fragment, which runs through the highly charged central pore of the B subunits. The pore contains at least 66 water molecules, which fill the space left by the A2 fragment. A detailed analysis of the contacts between A and B subunits showed that most specific contacts occur at the entrance of the central pore of the B pentamer, while the contacts within the pore are mainly hydrophobic and water mediated, with the exception of two salt bridges. Only a few contacts exist between the A1 fragment and the B pentamer, showing that the A2 fragment functions as a ‘linker’ of the A and B parts of the protein. Interacting with the A subunit by the B subunits does not cause large deviations from a common B subunit structure, and the 5-fold symmetry is well maintained.

A potential NAD⁺-binding site is located in an elongated crevice at the interface of two small sheets in the A1 fragment. At the back of this crevice the functionally important Arg7 makes a hydrogen bond connecting two strands, which seems to be conserved across the ADP-ribosylating toxin family. The putative catalytic residue (A1:Glu112) is located nearby, close to a very hydrophobic region, which packs two loops together. This hydrophobic region may be important for catalysis and membrane translocation.

Keywords: crystal structure; cholera toxin; heat-labile enterotoxin; NAD⁺ binding; ADP-ribosylating

† Present addresses: T. K. Sixma, Department of Molecular Biophysics and Biochemistry, Yale University, BCM 154, 295 Congress Avenue, New Haven, CT 06510, U.S.A.; B. Witholt, Eidgenössische Technische Hochschule, Zürich, Switzerland; W. G. J. Hol, Department of Biological Structure SM-20, School of Medicine, University of Washington, Seattle, WA 98195, U.S.A.

‡ Author to whom correspondence should be addressed at: Department of Biological Structure SM-20, School of Medicine, University of Washington, Seattle, WA 98195, U.S.A.

1. Introduction

Heat-labile enterotoxin (LT \dagger) is a diarrheal agent produced by enterotoxigenic strains of *Escherichia coli*. It has high homology ($\sim 80\%$ identity) to cholera toxin (CT), produced by *Vibrio cholerae*. These bacterial protein toxins are the causative agents of important diseases, particularly in developing countries (Sack, 1975; Black, 1986; Gyr & Meier, 1987). They have an AB₅ subunit structure (Gill, 1976), where the five B subunits (103 amino acid residues each) are necessary for binding to the ganglioside receptor G_{M1} on the membrane of the target intestinal epithelial cell (reviewed in: Finkelstein, 1988; Moss & Vaughan, 1988). This binding is obligatory for the uptake of the enzymatic A subunit into this cell. The process may involve endocytosis (Hansson *et al.*, 1977; Joseph *et al.*, 1978; Janicot & Desbuquois, 1987; Janicot *et al.*, 1991), probably *via* non-coated invaginations (Montesano *et al.*, 1982; Tran *et al.*, 1987), but much attention has been drawn to the possibility of direct insertion of the A subunit into the membrane (Gill, 1976; Tomasi & Montecucco, 1981; Wisniewski & Bramhall, 1981; Dwyer & Bloomfield, 1982; Ribi *et al.*, 1988). Such an insertion is not necessarily contradictory to endocytosis, but may also take place as part of this process (Janicot *et al.*, 1991).

For activity the A subunit needs to be proteolytically cleaved (after approximately residue 192 or 194) and reduced at its single disulfide bond (186–199) to yield two fragments: an enzymatic A1 (1–192(194)) and an A2 linker fragment (193(195)–240) (Gill, 1976; Mekalanos *et al.*, 1979b; Moss *et al.*, 1981). In cholera toxin the proteolysis takes place during biosynthesis by an endoprotease, but in LT it occurs by extracellular processes. For both proteins the reduction is thought to take place at the membrane surface of the target cell.

The enzymatic action mediated by these toxins is an ADP-ribosylation reaction, splitting NAD⁺ into free nicotinamide and an ADP-ribose moiety (Gill, 1975; Kunkel & Robertson 1979; Moss & Vaughan, 1977; Mekalanos *et al.*, 1979a; Moss *et al.*, 1979). The latter is coupled covalently to the guanidinium group of an arginine in the α -subunit of the stimulatory G protein (G_{sz}) of the adenylate cyclase system (Cassel & Pfeuffer, 1978; Gill & Richardson, 1980). This process keeps the G_{sz} permanently in its activated GTP state, where it stimulates adenylate cyclase to produce large quantities of cyclic AMP. This leads eventually to the release of large quantities of ions and fluids from the affected intestinal cell. The *in vitro* reaction is stimulated by the presence of a member of a class of small GTP binding proteins, called ADP-ribosylating factors or ARFs

(Kahn & Gilman, 1984; Lee *et al.*, 1991; reviewed in: Moss & Vaughan, 1991).

The proteins are produced as monomers in the cytoplasm of the bacteria, and assembled in the periplasm, presumably at the periplasmic membrane (Hofstra & Witholt, 1984, 1985; Hirst *et al.*, 1984; Hardy *et al.*, 1988). The A subunit is not essential for, but has been shown to stimulate, the assembly of the B pentamer (Hardy *et al.*, 1988). In *Vibrio cholerae* there is a secretion machinery present for export across the outer membrane, but in *E. coli* the toxin remains in the periplasmic space (Hirst *et al.*, 1984), whereafter host intestinal factors may provide the signal for release (Hunt & Hardy, 1991).

We have recently solved the structure of the AB₅ complex of LT to 2.3 Å (1 Å = 0.1 nm) by X-ray crystallography (Sixma *et al.*, 1991). The complex contains a highly structured donut-like B pentamer, linked to the A subunit through the C-terminal A2 fragment. This A2 is inserted into a central pore in the B pentamer and emerges at the other side as a small helix (Sixma *et al.*, 1992a). Further studies have revealed the location of part of the sugar-binding site, indicating that the membrane-binding orientation of the complex is with the A subunit pointing away from the membrane (Sixma *et al.*, 1992a). The structure solved is in the unnicked, unreduced form and is therefore inactive, but a putative active site has been indicated by analysis of mutants with changed activity and by analogy to *Pseudomonas aeruginosa* exotoxin A (ETA), another ADP-ribosylating toxin (Allured *et al.*, 1986). This toxin, like diphtheria toxin, has the diphtamide residue in elongation factor-2 (EF-2) as its substrate. In its enzymatic A subunit ETA has sequence similarity to diphtheria toxin. Comparison of the structure of the A subunit of ETA with that of LT (Sixma *et al.*, 1991) has shown that they have a very similar common core in terms of the secondary structure, but very little sequence similarity, with as major exception the active site glutamic acid (Carroll & Collier, 1984, 1987; Douglas & Collier, 1987; Tsuji *et al.*, 1990, 1991).

There are many reasons to study the structure of LT. It can help to answer important biological questions about the membrane binding and transport of the protein into the target cell. The structures of exotoxin A from *Ps. aeruginosa*, LT and the recently solved diphtheria toxin (Choe *et al.*, 1992) are the first three-dimensional structures known of a new class of NAD⁺-binding proteins and further studies will be required to analyze this mode of NAD⁺ binding. In addition the LT/CT family makes use of the mammalian G-protein-linked signal transduction pathway, and it will be interesting to find out more about the complex formation of LT and G_{sz} as well as of LT and ARF. Finally the structure of LT and the detailed knowledge of its activity may be used to develop new drugs and vaccines against the still devastating diseases caused by this type of toxin (Black, 1986).

In this article the structure determination of the

\dagger Abbreviations used: LT, heat-labile enterotoxin; CT, cholera toxin; ARF, ADP-ribosylating factor; ETA, exotoxin A; PEG, polyethylene glycol; NAD, nicotinamide adenosine dinucleotide; MIRAS, multiple isomorphous replacement with anomalous scattering; r.m.s., root-mean-square.

native structure of LT is described in more detail, with emphasis on the influence of phase combinations during this process. The high resolution structure of the complex, now refined to 1.95 Å, is analyzed. This provides a detailed description of the interactions between subunits, which determine the existence of this unusual AB₅ complex.

2. Experimental procedures

(a) Crystallization and data collection

Heat-labile enterotoxin from porcine *E. coli* was over-expressed from *E. coli* C600 with the multicopy plasmid EWD 299 (Dallas *et al.*, 1979; Pronk *et al.* 1985). The protein was purified and crystals were grown according to the method of Pronk *et al.* (1985), in space group $P2_12_12_1$ with cell dimensions $a = 119.2$ Å, $b = 98.2$ Å, $c = 64.8$ Å. The crystals were stabilized in a solution of 6% (w/v) PEG 6000, 100 mM-Tris, 1 mM-EDTA and 0.02% (w/v) sodium azide (pH 7.5) (Pronk *et al.*, 1985). Data sets of native and derivative LT-crystals, collected on an Enraf-Nonius FAST area detector for use in the structure determination, have been described by Sixma *et al.* (1991).

High-resolution data of LT were collected at room temperature from a single crystal on the X11 synchrotron beam line at the EMBL-outstation at DESY in Hamburg. The data were collected using the oscillation method with the wavelength set to 0.96 Å. Detection was done on an image plate using the Hendrix-Lentfer image plate scanner. In total, 96 frames of 1.0 deg. oscillation were collected to a resolution of 1.90 Å from a single crystal. An additional low-resolution data set to 2.7 Å was collected 6 weeks later of the same crystal on the X31 beam line. The wavelength was set to $\lambda = 1.009$ Å and an updated version of the image plate was used. In total 60 frames of 1.5 deg. were collected at this lower resolution. Processing of the data took place with the Mosco suite of programs (Machin *et al.*, 1983) adapted to image plate, using the profile fitting option for image plates. The crystal diffracted initially to a resolution of 1.8 Å, but decayed rapidly to about 1.9 Å. During data processing the limit of resolution was set to 1.95 Å. The last 5 frames of the high-resolution set were skipped since they suffered from too much radiation damage in the high-resolution shells. The overall merging R -factor was 8.3% for the high-resolution setting, with R_{merge} in the last shell (2.0 to 1.95 Å) being 23.1%. Despite the clearly visible crystal decay of the high resolution reflections in the first setting and the long lag time between data collections, the second setting still had a reasonably good quality. The 2 settings scaled very well together, as is evident from the R_{merge} which was 8.3%, for the high resolution data alone and increased by only 0.5% after merging the 2 settings together, resulting in an $R_{\text{merge}} = 8.8\%$ for the complete data set (Table 1).

(b) Structure determination

The structure of LT was determined as described generally by Sixma *et al.* (1991), but some important aspects will be described here in more detail. An overview of the procedure is given in Table 2, and the various phase sets mentioned in this table are indicated in the text at each stage. Three heavy atom derivatives (Sixma *et al.*, 1991) were used to obtain initial multiple isomorphous replacement with anomalous scattering (MIRAS) phases

Table 1
Data collection and refinement statistics

Resolution (Å)	1.95
Completeness (∞ –1.95 Å) (%)	94.4
Completeness (1.96–1.95 Å) (%)	86.4
Number of measurements	230,816
Number of unique reflections	53,762
Number of unique reflections > 1 σ	53,040
R_{merge} (%)†	8.8
Refinement	
Resolution (Å)	8–1.95
R -factor (%)‡§	18.2
Number of reflections	52,397
Protein atoms	6035
Solvent molecules	293
r.m.s. deviation from ideality§¶	
Bond distances (Å)	0.15
Angles (degrees)	3.0
Dihedrals (degrees)	24.9
Improper dihedrals (degrees)	1.24
Mean temperature factors (Å ²)	
AB ₅	33.4
B #1	28.1
B #2	27.9
B #3	32.7
B #4	32.7
B #5	29.5
A1	38.3
A2	34.7
water	36.6

† $R_{\text{merge}} = \Sigma |I - \langle I \rangle| / \Sigma I$.

‡ R -factor = $\Sigma ||F_o| - |F_c|| / \Sigma |F_o|$ on all reflections 8–1.95 Å (53,040 reflections).

§ Values given for refinement model, including weak regions described in the text.

¶ Calculated by X-PLOR.

to 3.1 Å, with an overall figure of merit of 0.62 (between 15 and 3.1 Å). The anomalous differences allowed determination of the correct hand. The electron density map calculated with these phases was not interpretable in detail, but allowed the calculation of a solvent flattening envelope (radius = 12 Å, solvent content set to 40% out of ~45% present in the crystal) (Leslie, 1987). In this envelope features related by 5-fold symmetry were visible. These features were used to deduce an initial localization of the 5-fold axis (going through the point $x = 0.10$, $y = 0.45$, $z = 0.50$), while an initial orientation was obtained from self rotation function analysis, yielding polar angles $\varphi = 0^\circ$, $\psi = 94^\circ$ (as defined by Rossman & Blow, 1962), i.e. almost along the crystallographic a -axis. The orientation and position of the 5-fold axis, as well as the size and extent of the 5-fold symmetric features, were optimized by a procedure in which correlation coefficients were calculated between spheres of density in the MIRAS electron density map, related by the 5-fold axis. The calculations were varied as a function of the location and orientation of the 5-fold axis. Maximum correlation coefficients were found when the 5-fold went through the point ($x = 0.10$, $y = 0.442$, $z = 0.498$), with polar rotation angles ($\varphi = 0^\circ$, $\psi = 97^\circ$). Good agreement with 5-fold symmetry, as judged from the correlation coefficients, was found in a region extending perpendicular to the 5-fold axis by a radius of 32 Å and extending along the 5-fold axis from $x = 0.23$ to $x = 0.95$.

Table 2
Stages in structure determination of LT

Stage	Starting phases or model†	End phases or model†	Model at end of stage
1. Multiple isomorphous replacement	—	MIRAS	—
2. Averaging I	MIRAS	AVE1	—
3. Model building B subunit	AVE1	MDL1	B pentamer (without loop 54–60)
4. Rigid body refinement B subunits	MDL1	MDL2	B pentamer (without loop 54–60)
5. Averaging II	MDL2+MIRAS	AVE2	—
6. Model building A subunit	AVE2	MDL3	B pentamer + 160 residues, 20 side-chains
7. TNT refinement B subunits	MDL2	MDL2b	B pentamer (without loop 54–60)
8. Averaging III	MDL4+1/2*MIRAS	AVE3	MDL4‡ = B ₅ (MDL2b) + A
9. Model building A subunit	AVE3	MDL5	B pentamer + ~ 231 residues, ~ 60 side-chains of A subunit
10. Averaging IV	MDL5+1/2*MIRAS	AVE4	—
11. Model building A subunit	AVE4	MDL6	B pentamer + ~ 220 residues, ~ 60 side-chains of A subunit
12. MD-refinement partial model	MDL6a	MDL7	B pentamer + ~ 206 residues, ~ 60 side-chains of A subunit
13. Tracing of A chain in phase combined electron density	MDL7+AVE2 +1/2*MIRAS		B pentamer including loop 54–60, A subunit, without second half A2

† The 2 central columns contain the following information: in case of averaging, start phases and end phases, respectively; in case of refinement, start co-ordinates, end co-ordinates; in case of model building, start phases (electron density), end co-ordinates.

‡ MDL4 = B₅(MDL2b) + current A model (MDL3).

(c) Fivefold averaging of B subunits

Real space electron density averaging (Bricogne, 1976) of the 5 B subunits and flattening of the solvent, were performed with DEMON, a suite of density-averaging programs written by F. Vellieux (unpublished results). An envelope for the averaging procedure was derived from the solvent flattening envelope by filling the protein region with random atoms (using a program written by M. Fujinaga). This co-ordinate file was edited to obtain 1 set of atoms corresponding to a single "B pentamer" (using the results from the 5-fold density correlation optimization) and another set of atoms describing a single "A subunit". From the combined co-ordinate set an averaging envelope was calculated having different regions for "B", "A" and "solvent". During the subsequent averaging procedure the "A" region was left unmodified, while the solvent was flattened and the B subunits were 5-fold averaged. A single cycle of averaging and map inversion at 3.1 Å produced a much improved electron density map, in which secondary structure was easily recognizable in the B subunit region (Fig. 1). In total 10 cycles of averaging were performed. After the averaging the electron density was inverted and the resulting phases were weighted using Sim weights (Sim, 1959, 1960) and combined (Hendrickson & Lattman, 1970) with MIRAS phases. This resulted in phase set AVE1 with an average phase difference of 45° with respect to the MIRAS phases and an *R*-factor of F_{invers} versus F_o of 26.8%.

(d) Model building and rigid body refinement of the B subunits

The B subunit sequence (Dallas & Falkow, 1980; with corrections as described by Leong *et al.* (1985) and Yamamoto *et al.* (1987)), could be traced easily in the resulting map. A model for 1 B subunit was built using the program FRODO (Jones, 1978) on an Evans & Sutherland PS390 graphics system, with skeletonized electron densities produced by the program BONES

(Jones & Thirup, 1986) as a background help, yielding co-ordinate set "MDL1". At this stage the loop containing residues B:54–61 could not be interpreted. The symmetry-related B subunits were generated by application of the 5-fold symmetry. Their position was optimized by rigid body refinement using the program TNT (Tronrud *et al.*, 1987) giving co-ordinate set "MDL2". Each subunit was allowed to move as a separate rigid group. The crystallographic *R*-factor ($\frac{\sum |F_o - F_c|}{\sum |F_o|}$) decreased from an initial value of 50.2% to 49.7% for MDL2, indicating that the symmetry parameters had been previously determined reasonably accurately.

(e) Fivefold averaging of B subunits to improve the A subunit

At this stage the electron density for the A subunit was still quite poor and could not be interpreted, except for a single long helix. Electron density maps of type $2F_o - F_c$, were calculated using 3 types of phases: (1) model phases derived from the B subunit model alone (MDL2); (2) model phases combined with MIRAS phases (MIRAS+MDL2); (3) model phases combined with phases, obtained after 5-fold averaging of the B subunits (MDL2+AVE1). Phase combination and weighting schemes used were according to the methods of Sim (1959, 1960) or Read (1986). To improve the density an iterative procedure of averaging and model building was started.

The averaging was reinitiated with an updated envelope using the B pentamer atom co-ordinates. As starting phases the combined MIRAS and model phases (MDL2), after TNT rigid body refinement, without loop 54–61, were used. The MIRAS phases were weighted down by taking half the Hendrickson–Lattman coefficients (Vellieux *et al.*, 1990). Averaging was done during 18 cycles, with gradual extension of the phases from 3.1 to 2.9 Å. The phases from map inversion were combined with the starting MIRAS+MDL2 phases at every step, initially using Sim, later Sim*Rayment (Rayment, 1983) weights for the inverted phases, in an attempt to lower

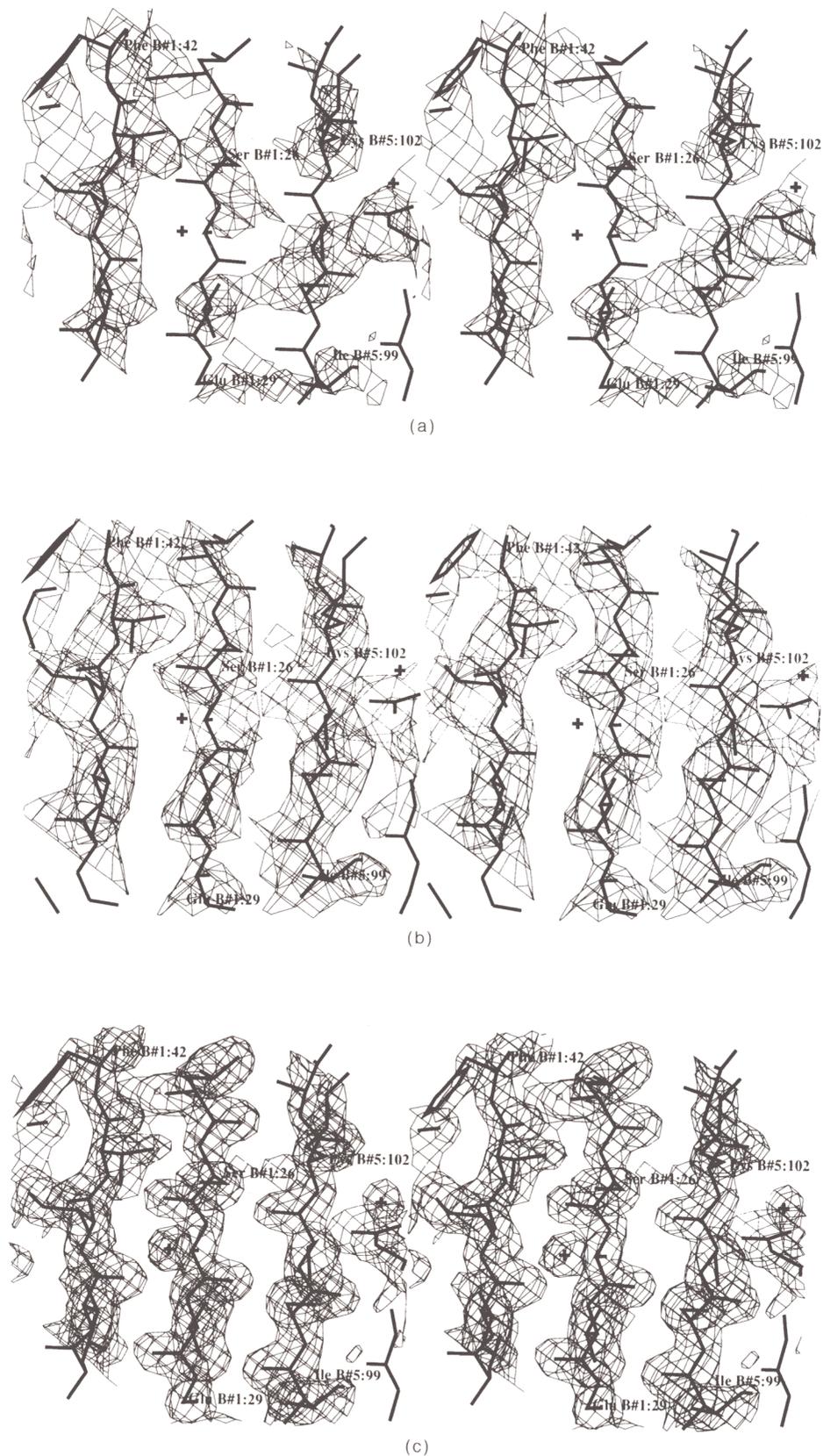


Figure 1. Sample of electron density at 3.1 Å resolution, showing the improvement due to 5-fold averaging of the B subunits around the interface of B#1 and B#5, in the β -sheet, contoured at 1 σ level of the map. The co-ordinates depicted are from the final refined model. (a) F_0 Fourier with figure-of-merit weighted MIRAS phases. (b) Sim weighted F_0 electron density after initial averaging of 3 B subunits (AVE1), 7 cycles with phase combination of the averaged phases with the MIRAS phases after each cycle. (c) $2F_0 - F_c$, SIGMAA (Read, 1986) weighted Fourier, using refined model phases and F_c derived from the final 1.95 Å refined model.

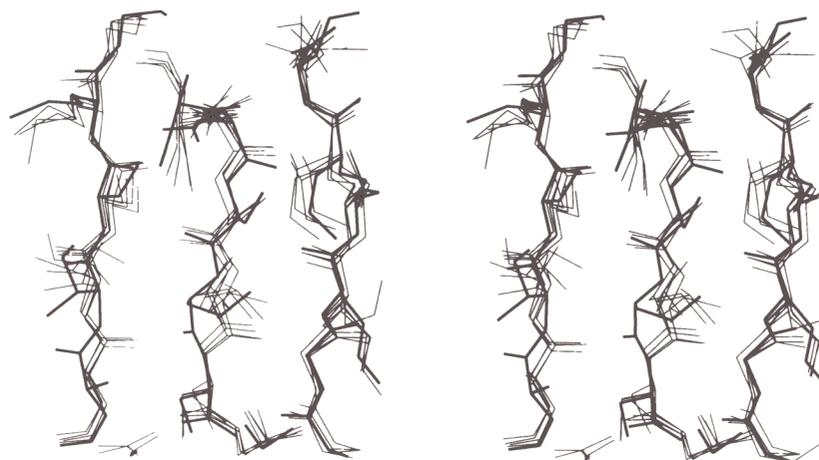


Figure 2. Effect of Gromos molecular dynamics refinement (Fujinaga *et al.*, 1989) on individual B subunits: starting model for all 5 B subunits (thick lines) superimposed with the 5 B subunits (thin lines) after initial Gromos MD-refinement of a model consisting of the B pentamer and a partial A model (MDL7). Note that the 2 peptide flips in this sheet have occurred automatically in all 5 B subunits independently, showing a good reproducibility of the molecular dynamics procedure.

the weight of the inverted map relative to the starting phases (Rossmann, 1988). This procedure resulted in phase set "AVE2". In the electron density map, calculated after this averaging step, a β -sheet and some fragments of helix could be traced in the A subunit, and a stretch of 20 residues could be correlated with the amino acid sequence (Dykes *et al.*, 1985; Yamamoto *et al.*, 1987). This resulted in a partial model of the A subunit, containing approximately 140 "alanine residues" and 20 residues with side-chains, which was added to the B pentamer co-ordinates to form the next model (MDL3).

A new cycle of averaging was started using an updated envelope by addition of the current model to the "atoms" used for the envelope calculation. In addition, the radius of the atoms used to create the A subunit envelope was increased, in the hope of avoiding the flattening of density for loops at the surface of this subunit. Phases from this partial AB₅ model were combined with the weighted down MIRAS phases. During this averaging step (21 cycles, with phase extension to 2.7 Å) Sim*Rayment and Rayment weights were used (AVE3). In the resulting map, the model of the A subunit was extended by model building to 231 residues of which approximately 60 had side-chains, while the others were modeled as polyalanine (MDL3).

The next averaging step (16 cycles) was started from phases obtained by phase combination of the weighted down MIRAS phases and a model (MDL4) containing a partially refined B pentamer (MDL2b: refined with TNT to an *R*-factor of 35% and geometry far from ideal (r.m.s. deviation on angles: 6°)) and the current partial A model (MDL3). The electron density calculated from the resulting phase set (AVE3) allowed rebuilding of the A subunit model. Some regions of the model had to be deleted, and not much sequence could be added at this stage (MDL5).

One more averaging step (16 cycles) was carried out, starting from the current model (approximately 230 residues, 60 side-chains for A; MDL5) and MIRAS phases. In this averaging cycle no phase combination was performed (AVE4). The resulting map was clearly different from previous maps, but did not help much to improve the structure of the A subunit (MDL6).

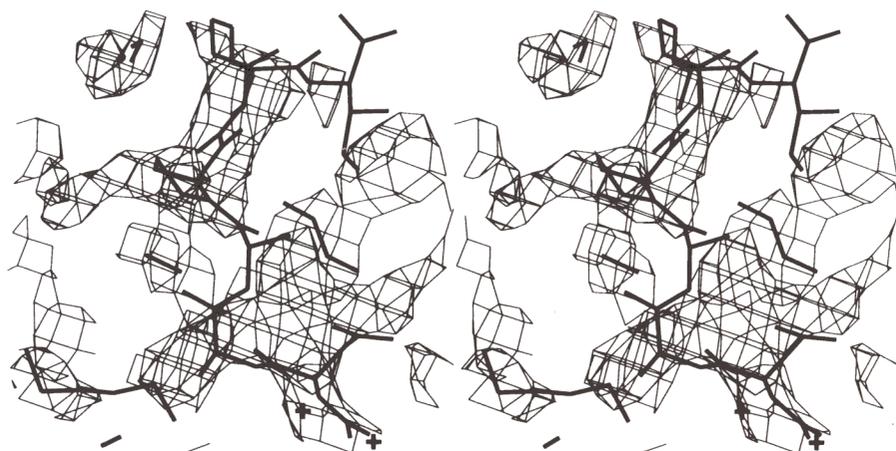
(f) Initial molecular dynamics refinement

A model containing the complete B subunits and the partial A subunit, the latter consisting of 14 fragments with a total of 206 amino acid residues out of 240 and about 60 side-chains, was subjected to GROMOS molecular dynamics refinement (Fujinaga *et al.*, 1989) (1 fragment was left out compared to the previous model) (MDL6a). The procedure of Gros *et al.* (1989*a,b*) was followed, starting the refinement at low resolution (8 to 5 Å) with gradual extension to 3 Å. The starting *R*-factor (8 to 3 Å) was 45.5% and the refinement was finished after energy minimization, with an *R*-factor of 32.8% (MDL7). No restraints were put on the 5-fold symmetry, to allow proper refinement of potential A-B interactions. An example of the quality of this refinement, and the reproducibility, is given in Fig. 2, in which 2 peptide planes have "flipped" in all 5 B subunits. The resulting electron density map showed some clearly improved regions in the A chain, but side-chain assignment and loop tracing were still not possible.

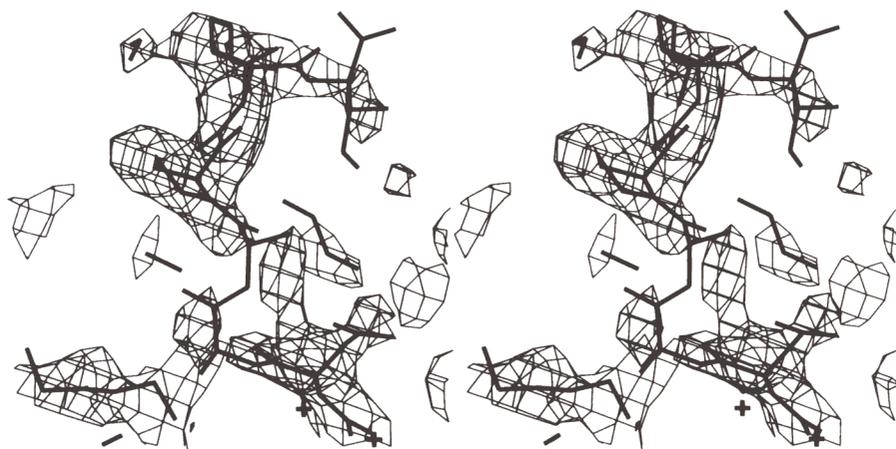
(g) Phase combination

Phase combination of the refined model phases with phases obtained from either MIRAS or averaging, respectively, did not result in an interpretable electron density map. However, simultaneous phase combination using all 3 sources of phase information, i.e. combining MIRAS (15 to 3.1 Å), averaging (AVE2; 15 to 3.0 Å) and model (MDL7; 7 to 3.0 Å) phases with the SIGMA phase combination routine resulted in an electron density map which was quite easily traceable for almost the entire A subunit. A good example of the improvement obtained by this procedure is shown in Fig. 3. The loop around ProA1:12 was not present in the refined partial model, as is obvious from the poor electron density in Fig. 3(c), and it had not yet been recognized as a loop. In the combined map the connecting density for the loop region is very clear (Fig. 3(e)) and allowed correct tracing of the sequence in this area.

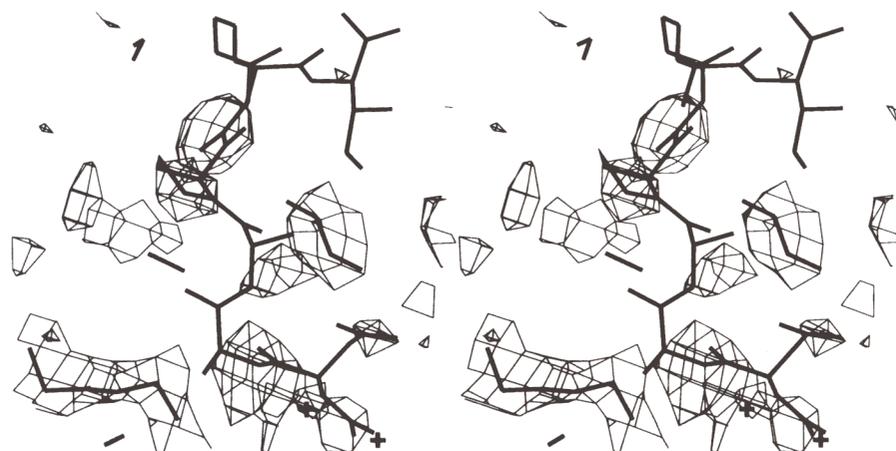
Obviously these 3 phase sets are not truly independent, but it should be noted that they all contain different types



(a)



(b)



(c)

Fig. 3.

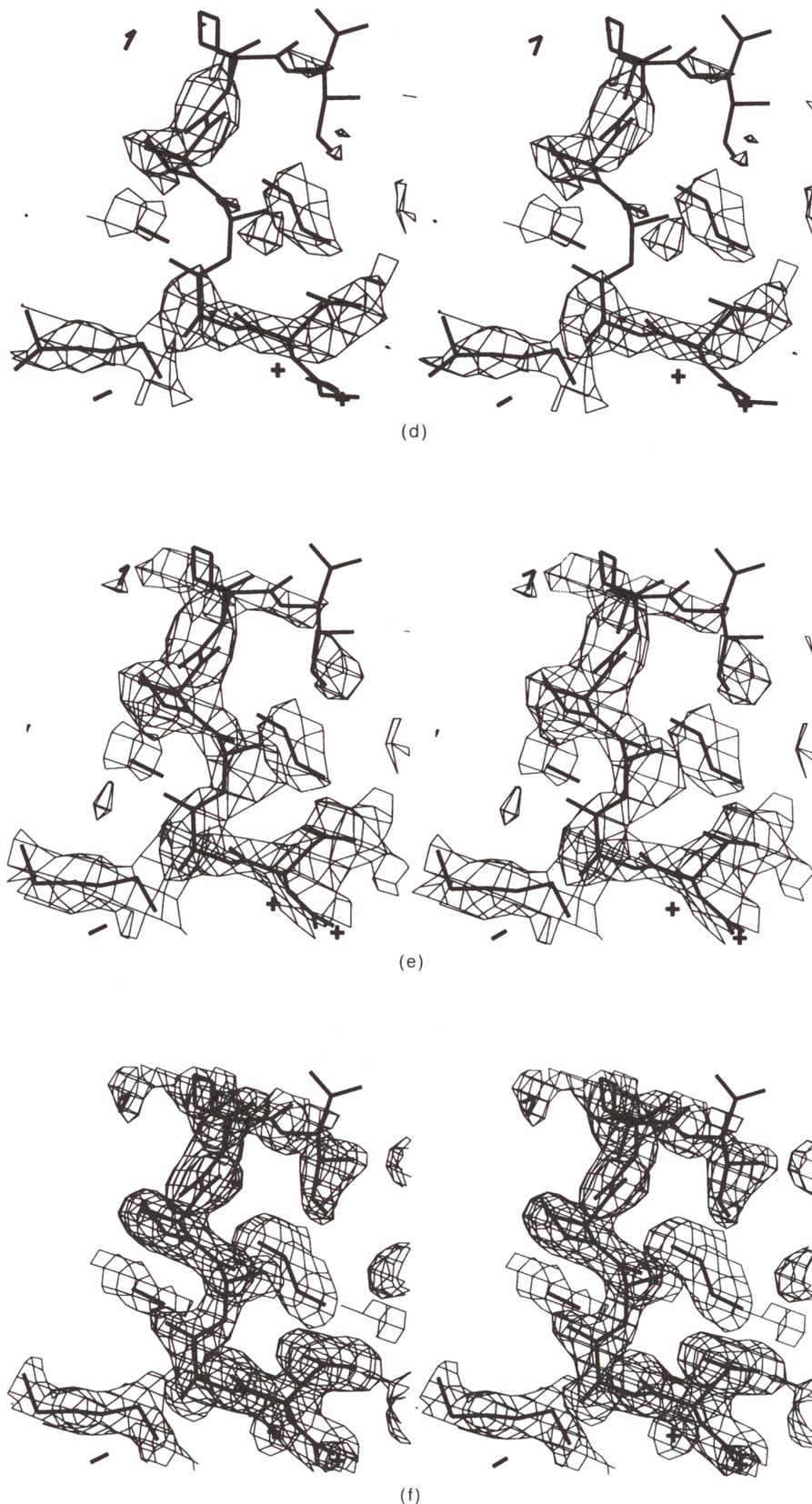


Figure 3. Illustration of electron density improvement at 3.0 Å resolution in the A subunit due to phase combination. Maps are contoured at 1σ level. Co-ordinates are from the final refined high-resolution model. The region shown is around ProA1:12 and this region was not present in the partial model MDL7 (see Table 2 for definitions of phase sets). (a) MIRAS phases; (b) AVE3 averaging phases; (c) MDL7 partial model phases; (d) combined phases from MDL7 partial model and AVE2 averaging phases; (e) "3-way" combined phases of MIRAS + partial refined model (MDL7) + averaging phases (AVE2); (f) final 1.95 Å refined model phases.

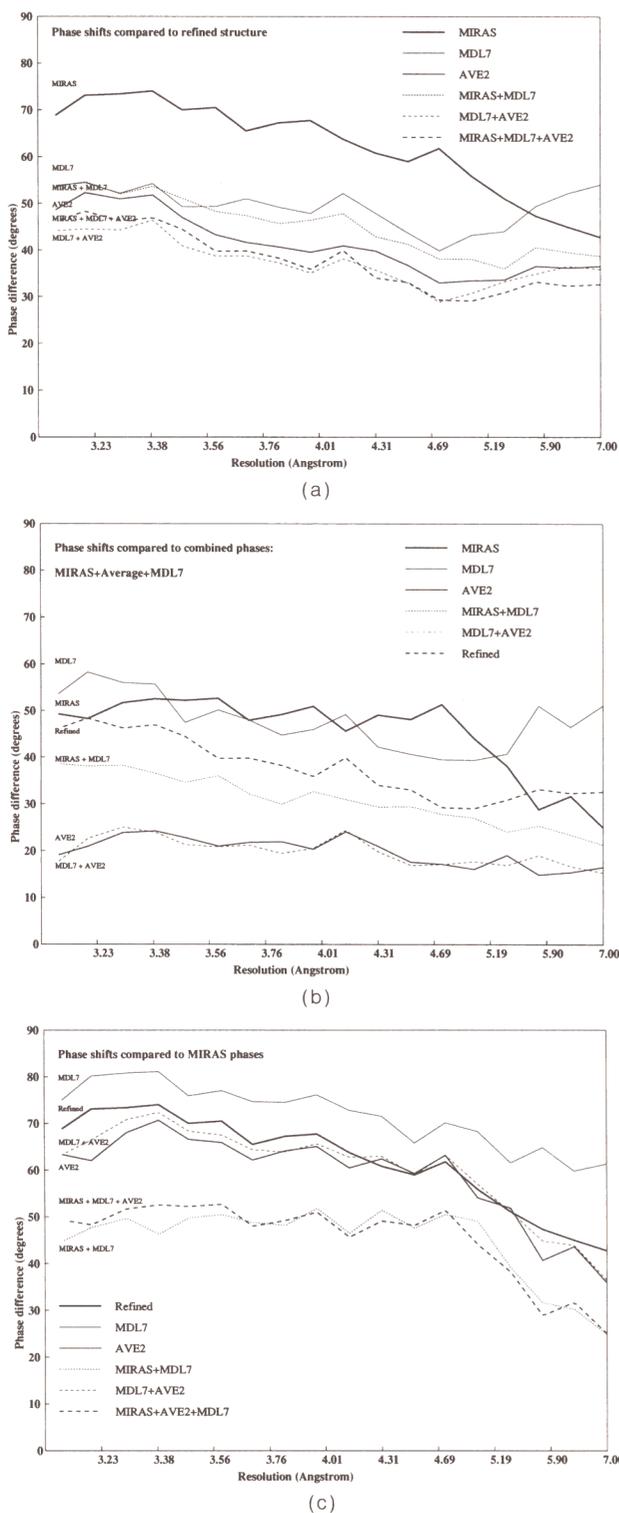


Figure 4. Phase shifts (degrees) of different phase sets as function of resolution (\AA) used to trace the A subunit. Phase differences are given between 7 and 3.1 \AA and calculated in all cases for the same set of reflections. Continuous lines: individual phase sets; broken lines: phase combinations (but note that the averaging phases are themselves the result of phase combination between MIRAS and averaging, as described in the text). Further description in the text and Table 2. (a) Phase differences with respect to the final refined model. (b) Phase difference compared to the “3-way” combined phase set. (c) Phase difference compared to the MIRAS phases.

of information. The MIRAS phases contain information about the whole structure, including the unknown part, but they are not very accurate. This can be seen in Fig. 4(a), which shows a comparison of different phase sets at this stage against the final “true” structure. The phase difference of the MIRAS phases to the true phases is 63° , much larger than for any other phase set at this stage. The model phases (MDL7), in contrast, are much closer to the final phases, but contain very little information about the unknown part of the structure. The averaging phases are already a mixture of the first 2 phase sets, by virtue of the phase combination during the averaging procedure, but they have some independent extra information obtained through the 5-fold averaging of the B subunits. In Fig. 4(a) it can be seen that the averaging phases (AVE2) alone are closer to the final refined phase set than the combined set of MIRAS + model phases. In general, it can be seen in Fig. 4(a) that all combined phase sets show an improvement with respect to their parent phase sets.

However, it is not clear in this Fig. that the “3-way” combination of MIRAS + model + average phases is better (i.e. closer to the refined value) than the combination of model + averaging phases alone. These 2 combined phase sets differ by only 20 deg. as can be seen in Fig. 4(b). In Fig. 4(c) it can be seen that this difference is indeed caused by the addition of the MIRAS phases, since the “3-way” combined phase set is much closer (by approximately 20 deg.) to the MIRAS phases than the model + averaging phase set. Apparently the addition of the MIRAS phases in the phase combination results in a phase set which is not actually more accurate overall (Fig. 4(a)), but which contains more phase information for the unknown part of the structure, than a phase set which combines only the model and the averaging phases.

The “3-way” phase combination was tried using a number of different averaging phase sets (AVE2, AVE3 or AVE4). The best results were obtained when an averaging phase set was used, for which only MIRAS and B_5 model phases were used as starting phases (AVE2). When the averaging phases included information about the partial A subunit model (AVE3), they were less effective, probably because they reinforced the errors in the A part of the structure. The averaging without phase combination (AVE4) had not resulted in a very good density overall. At the edges of the envelope, however, use of the AVE3 and AVE4 phase sets showed improved density compared to use of the AVE2 phase set in the triple phase combination because these phases were obtained after updating the envelope for the A subunit had expanded the total A region and improved the loop regions.

The region of the A2 subunit that inserts into the B subunit pore could not be traced in any of these maps, since it was wiped out by the averaging procedure. This region was identified later, during the refinement.

(h) Refinement

Crystallographic refinement of the structure was initially performed against a 2.3 \AA FAST data set (Sixma et al., 1991) using the standard procedure of iterative cycles of refinement and model building. No restraints were ever put on the 5-fold non-crystallographic symmetry during the refinement, except that the initial model was composed of 5 identical B subunits. During the model building steps, however, the co-ordinates of the symmetry-equivalent subunits were sometimes used to check whether differences were true variations or due to errors in the refinement. Initially, refinement was carried

out by molecular dynamics refinement with GROMOS, yielding a final R -factor of 18.6%, and in later stages by least-squares refinement with the TNT package. In the early stages of the refinement individual temperature factors were refined with Agarwal's FFT routine (1978), mainly as a check on errors indicated by very high B -factors. The individual B -factors were reset to an overall value before each round of refinement. In later stages of the refinement the individual values were retained and only in the last stages of TNT refinement a correlation on the B -factors was imposed along the chains resulting in an R -factor of 17.8%. During refinement the C-terminal 20 residues of the A2 subunit became gradually more clear. Much effort was put into the refinement of the C-terminal 7 residues, but the density at this site remained very bad, even after molecular dynamics refinement with this region was omitted. High-resolution refinement using the 1.95 Å synchrotron data was started using the least squares minimization with the TNT package, followed by X-PLOR simulated annealing refinement (Brünger *et al.*, 1987, 1990; version 2.1). The latter program package was chosen for the ease of incorporation of bound ligands during refinement, mainly for a parallel refinement of the lactose-LT complex (Sixma *et al.*, 1992a). The refinement of the lactose-LT complex also revealed the correct position of the A2 chain, which could subsequently be incorporated in the high-resolution native LT structure. In the final high-resolution model several loops are poorly defined. Attempts were made to improve these by an omit refinement run, but this did not result in a significant map improvement. In the course of the refinement 293 water molecules were incorporated into the structure.

(i) Analysis of the structure

Since the interactions of the 5 B subunits with the A subunit vary, the individual subunits have been given a number, indicated with B#1 through B#5, anti-clockwise when viewed along the 5-fold axis toward the A subunit (Fig. 7). Surface accessibility calculations were performed with the method of Kabsch & Sander (1983), by the program DSSP. Hydrophobic contributions to the accessible surface were calculated with the Connolly procedure (1983), with an atomic radius consisting of the van der Waals' radius + 1.4 Å, and a probe radius of zero. Hydrogen bonds were calculated with X-PLOR and included in Tables only if they were shorter than 3.2 Å and had a deviation from linearity of less than 65°. Only those salt bridges with a maximal distance of 4.0 Å were included in tables (Barlow & Thornton, 1983). Hydrophobic interactions were given if a carbon-carbon distance is less than 4.0 Å. Crystal contacts were cut off at a distance of 3.5 Å. These criteria for interactions used for compiling the tables are fairly stringent; a more general impression of subunit interactions can be obtained from the changes in accessible surface as depicted in Fig. 11.

3. Results

(a) Quality of the structure

A model of the AB₅ complex of heat-labile enterotoxin has been obtained, with a crystallographic R -value of 18.2% ($\sum |F_o - F_c| / \sum |F_o|$), on all data from 8 to 1.95 Å, and with good geometry (see Table 1). The estimated co-ordinate error, determined by the SIGMAA procedure, is 0.21 Å. A number of loops

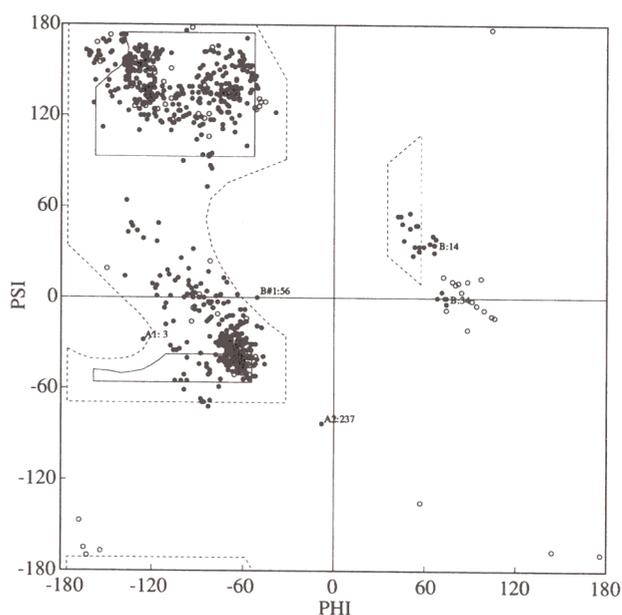


Figure 5. Ramachandran plot of the final refined model of the complete AB₅ complex showing ϕ and ψ angles in degrees (Ramachandran & Sasisekharan, 1968). Open circles: glycine residues; filled circles: all other residues. Residues outside the allowed regions are discussed in the text.

are not very well visible in the refined structure: there is no density for the N-terminal (A1:1-3) three and the C-terminal (A2:237-240) four residues of the A subunit as well as for residues A:189-195, which form the loop connecting the A1 and the A2 fragment. Furthermore, there is weak density in the main-chains of the regions B:54-59, A1:33-38, A1:79-80 and A1:110 and A1:137-138. The dihedral angles along the main-chain are given in Figure 5. There are three residues completely outside the allowed regions (Ramachandran & Sasisekharan, 1968): A1:3, B#1:56 and A2:237. All three are in the badly defined regions mentioned above. In addition all five Lys B:34 residues have a somewhat unusual phi/psi combination. There is one *cis*-proline in each B subunit, at residue B:Pro93, and one in the A1 subunit, A1:Pro178. All six *cis*-prolines have well-defined densities, with average B -factors ranging from 21.8 to 38.4 Å².

The crystallographic temperature factors were refined in a correlated fashion along the chain, but the mean B -factor value of the high resolution data set was relatively high: 33 Å² for the entire AB₅ complex (*versus* 20 Å² in the 2.3 Å structure: Table 1). A plot of the B -factors given as mean values per residue for main-chain and side-chain atoms is given in Figure 6. The values for the individual B subunits correlate quite nicely (Fig. 6(b) and (c)). Apparently the A subunit has a higher flexibility than the B subunits as can be seen in the mean values per subunit (Table 1), with an average of 30 Å² for the B subunits and a value of 38 Å² for A1.

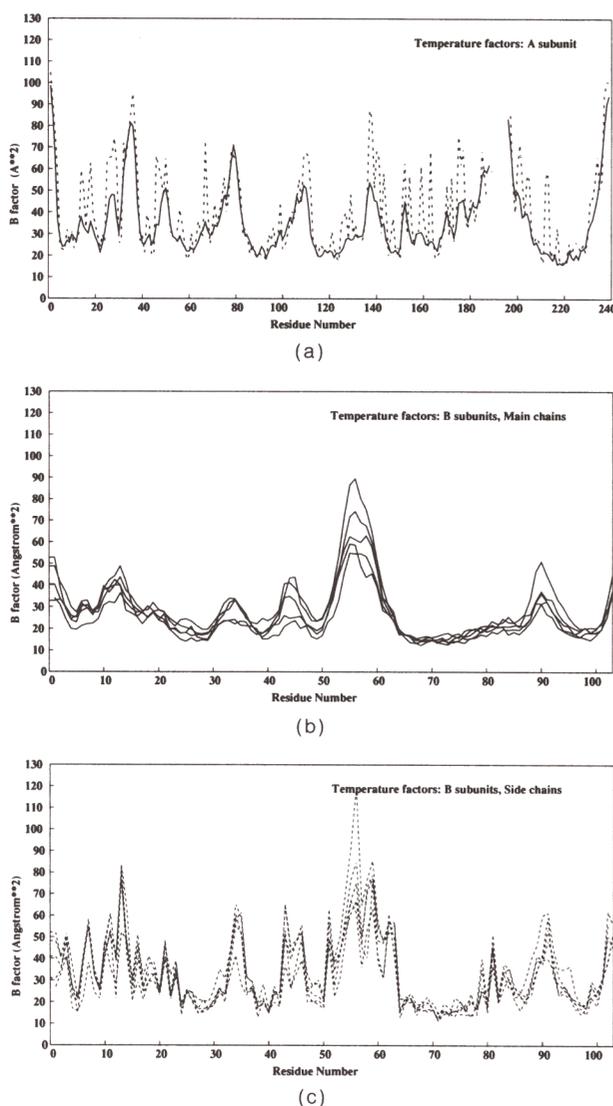


Figure 6. Mean crystallographic temperature factors (\AA^2) for main-chain (continuous lines) and side-chain (broken lines) atoms, averaged per residue. (a) A subunit; (b) main-chain for 5 B subunits superimposed; (c) side-chains for 5 B subunits superimposed. Mean overall values per subunit are given in Table 1. Note the good correlation between the B subunits.

(b) *The B monomer*

The B monomer of LT (Figs 7(a), 8(a)) is a small, compact, highly structured subunit with dimensions of 38 Å along the 5-fold axis and about 28 Å by 42 Å in directions perpendicular to the 5-fold. The secondary structure has been described by Sixma *et al.* (1991), and consists of a small N-terminal helix ($\alpha 1$), two three-stranded anti-parallel sheets (sheet I, composed of strands $\beta 2$, $\beta 3$, $\beta 4$ and sheet II, strand $\beta 1, \beta 5$ and $\beta 6$) and a long α -helix ($\alpha 2$) (Table 3). A schematic main-chain hydrogen bond description is given in Figure 9(a). The two sheets interact with each other *via* three main-chain hydrogen bonds, between strand $\beta 4$ and $\beta 6$, running parallel to each other in this short stretch. This is

possible because the two strands $\beta 5$ and $\beta 6$ bend “backwards” to make this interaction, as shown in Figure 7(a), and are connected *via* a loop which contains the *cis*-ProB:93. In this way the two sheets form a β -barrel, although one side of the long helix enters into the hydrophobic side of this barrel.

The LT B topology is identical with that of the recently elucidated structure of B subunits of *E. coli* verotoxin (or shiga-like toxin) (Stein *et al.*, 1992) although there is no detectable sequence homology. This protein misses only the $\alpha 1$ helix. Verotoxin is another AB_5 toxin, but with a ricin-like activity (reviewed by Brunton, 1990). A detailed comparison will be described elsewhere (Sixma *et al.*, 1993). In addition, the same topology is also found in the first 99 residues of *Staphylococcus aureus* nuclease (Arnone *et al.*, 1971), a protein which exists in a monomer form. In the nuclease, however, strand $\beta 1$ and $\beta 2$ are commonly assigned as one single strand and the $\alpha 1$ helix does not enter into the barrel.

The N-terminal helix in LT is linked to strand $\beta 5$ *via* the disulfide bridge between Cys9 and Cys86. The loops in the LT B subunit can be divided into two classes, referring to the two ends of the sheets. On one end of the subunit, the “narrow” (or “A”) end, the loops are generally short, involving the connections $\beta 1$ – $\beta 2$, $\beta 3$ – $\beta 4$ and $\alpha 2$ – $\beta 5$ as well as the C terminus. The subunit widens at the other end, with much longer loops connecting secondary structure elements $\alpha 1$ – $\beta 1$, $\beta 2$ – $\beta 3$, $\beta 4$ – $\alpha 2$, $\beta 5$ – $\beta 6$. The N terminus is also more or less located on this end of the subunit. The loops at the “narrow” end involve only one β -hairpin in the loop $\beta 3$ – $\beta 4$, which has a standard type of turn, the I(1–4)+G1 β -bulge conformation (Sibanda *et al.*, 1989). The loops at the wider end are more complicated; they are all involved in shaping the cavity that forms the binding site for the galactose part of the membrane receptor (Sixma *et al.*, 1992a). Most of the turns have either a glycine or a proline nearby (Gly33 between $\beta 2$ and $\beta 3$; Pro53 and Gly54 between $\beta 4$ and $\alpha 2$ and the *cis*-Pro91 between $\beta 5$ and $\beta 6$). The longest loop, which connects $\beta 4$ and $\alpha 2$ from Val52 to Asp59, extends below the plane of the β -sheets. This loop is quite flexible in the native structure, but becomes distinctly less mobile after lactose binding (Sixma *et al.*, 1992a).

The long $\alpha 2$ helix has a very amphipathic character: on one side it has charged residues (8 charges in total), while on the other side it is mainly hydrophobic. The last turn of the helix is an exception, having hydrophobic residues also on the “charged side” (Ile74, Leu77). The uncharged side of the helix interacts with the side-chains of the β -sheets in the monomer whereas the charged side is exposed to the central pore of the pentamer (see section (c)). As is quite common in helices (Barlow & Thornton, 1988), this helix is curved, with a curve towards its hydrophobic side, due to additional hydrogen bonds of the main-chain atoms in the helix on the pore side to water molecules and hydrophilic side-chains. The O γ atoms of Thr78 and several water molecules are involved in such hydrogen bonds, in addition to the

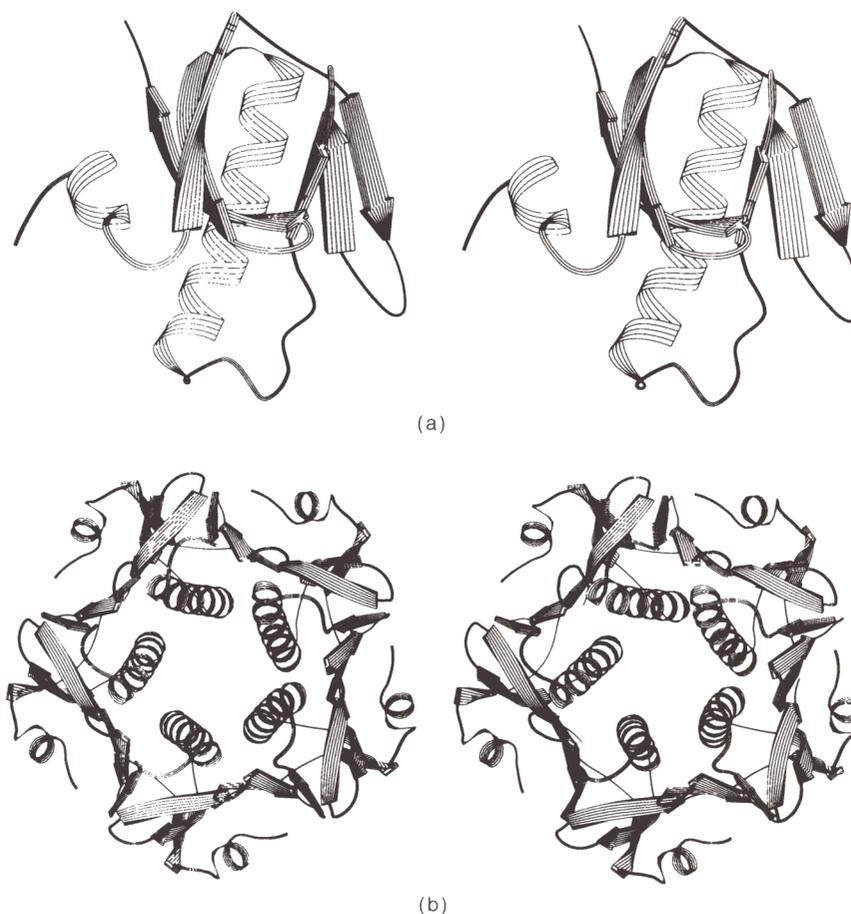


Figure 7. Stereo ribbon drawings showing the secondary structure of the B subunits of LT. (a) B monomer, viewed approximately perpendicular to 5-fold axis of pentamer; (b) B pentamer, viewed down the 5-fold axis, looking towards the A subunit. The Figure was made using the program Ribbon (Priestle, 1988). Numbering of B subunits is anti-clockwise when viewed toward the A subunit.

guanidinium group of Arg67 of the helix in the neighboring subunit $N+1$ (see section (c)). This curve is the origin of the "narrowing" of the pore within the pentamer.

(c) The B pentamer

The B pentamer has an unusual donut shape, with the five monomers arranged symmetrically around the central, 5-fold axis, leaving a pore in the middle. The short loops on the narrow end of the individual subunits combine to give a very flat surface on one side of the pentamer, which is the side where the A subunit is located. The longer loops on the opposite side produce a much more convoluted surface, which was shown to form part of the membrane receptor binding site (Sixma *et al.*, 1992). A ribbon diagram of the B pentamer is given in Figure 7(b), viewed along the 5-fold axis, where it can be seen that the three-stranded anti-parallel β -sheets of the B monomers combine with the opposite sheet of the neighboring monomers to form six-stranded anti-parallel sheets (seen end on in

Fig. 7(b)), with $\beta 2$ of subunit N and $\beta 6$ of N1 forming adjacent strands (numbering of B subunits defined in legend of Fig. 7). Mutants with deletions at the C terminus of the B subunits were shown to interfere with pentamer formation, presumably due to loss of main-chain hydrogen bonds with strand $\beta 2$ of the neighboring subunit (Sandkvist *et al.*, 1990).

The small $\alpha 1$ helices pack each against a six-stranded sheet on the outside of the molecule, with the N terminus exposed to the solvent. A very polar central pore is formed by the long $\alpha 2$ helices, which pack closely together. The radius of the pore can be characterized by the distance of the helix axis to the 5-fold axis, which varies between 12.5 Å at the N termini of the helices, on the "convoluted" surface, to 10.4 Å at the C termini, near the flat or "A" side of the pentamer. The distance from the closest side-chain atoms to the 5-fold axis varies between 4.9 Å (Ile74 C ^{δ 1}) at the C-terminal end and 6.0 Å (Asp70) in the middle of the pore. Many charged residues are located within the central pore, giving the pentamer a very asymmetric distribution of charges (Fig. 10). The central helices have more positive (Lys62, Lys63, Arg67, Lys69 and Arg73) than negative

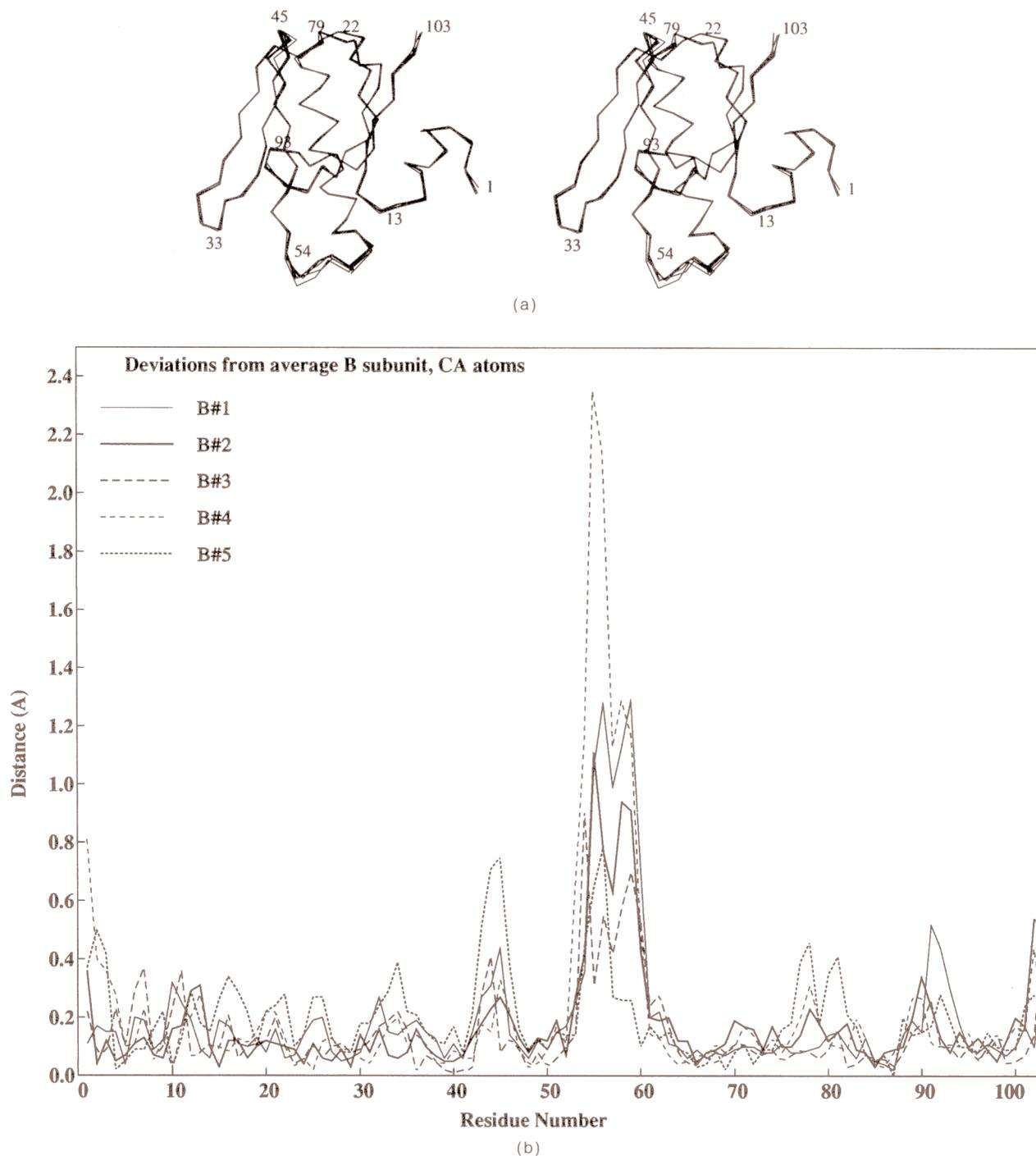


Figure 8. Fivefold symmetry in B subunits of LT. All 5 B subunits are superimposed on an “average B subunit” calculated as the true average after superposition of all B subunits on B#1. (a) C^{α} atom trace of all 5 B subunits. Residue numbers are given in each loop. (b) Deviations from average C^{α} atom position for the 5 B subunits, giving an indication of the differences between individual B subunits. The r.m.s. deviations are 0.20 to 0.28 Å. Note the flexible loop from residue 52 to 59. Interestingly, the long helices (residues B:59–78) that form the central pore are very similar.

(Asp59, Glu66 and Asp70) residues, resulting in a net positive environment. In the region from Asp59 to Lys63 the pore is more open and residues Asp59 and Lys62 point somewhat outward. These lysine residues may also interact with the negative residues at the C terminus of A2. Within the actual pore there are thus at least 5×3 positive and 5×2

negative residues, i.e. 25 charged residues, with a net positive charge of +5. If the lysine 63 residues are also taken into account the net positive charge goes up to +10 in the pore.

The B pentamer of LT is stabilized by a variety of interactions, of which the hydrogen bonds between B subunits are listed in Table 4A. Note that each

Table 3
Secondary structure elements in LT

B subunits		B #1		B #2		B #3		B #4		B #5	
		Start	End	Start	End	Start	End	Start	End	Start	End
α -Helices	$\alpha 1$	5	10	5	10	5	9	5	9	5	9
	$\alpha 2$	61	77	59	78	59	78	59	78	59	77
β -Strands	$\beta 1$	15	22	15	22	15	22	15	23	15	22
	$\beta 2$	26	30	26	30	26	30	26	30	26	30
	$\beta 3$	37	41	37	41	38	41	38	41	38	41
	$\beta 4$	47	50	47	50	47	50	47	50	47	50
	$\beta 5$	82	88	82	88	82	88	81	88	82	88
	$\beta 6$	94	102	94	102	94	102	94	102	94	102
A subunit											
	β -Strands		α -Helices								
	Start	End	Start	End							
$\beta 1$	4	9	$\alpha 1$	13	19						
$\beta 2$	21	22	$\alpha 2$	41	46						
$\beta 3$	59	63	$\alpha 3$	66	76						
$\beta 4$	82	89	$\alpha 4$	97	101						
$\beta 5$	94	96	$\alpha 5$	102	104	3_{10}					
$\beta 6$	112	116	$\alpha 6$	108	110	3_{10}					
$\beta 7$	119	120	$\alpha 7$	121	123	3_{10}					
$\beta 8$	124	131	$\alpha 8$	147	152						
$\beta 9$	134	135	$\alpha 9$	158	161						
$\beta 10$	140	141	$\alpha 10$	162	164	3_{10}					
			$\alpha 11$	172	175	3_{10}					
			$\alpha 12$	179	182	3_{10}					
			A2: $\alpha 1'$	197	222						
			A2: $\alpha 2'$	223	226	3_{10}					
			A2: $\alpha 3'$	232	236						

Values determined using DSSP (Kabsch & Sander, 1983). Beginning and end of each strand are given. Note that the numbering of secondary structure elements is different for A1 from that published in Sixma *et al.* (1991). The main difference is the addition of 2 small separate 3_{10} helices involving residues A1:108–110 and A1:102–104 (in fact the latter is a continuation of helix $\alpha 4$, but at an angle of about 90°). Helix $\alpha 10$ and A2: $\alpha 2'$ are continuations of helix $\alpha 9$ and A2: $\alpha 1'$, respectively.

subunit forms both the hydrogen bonds in the first column (with subunit $N-1$) and in the second column (with subunit $N+1$). The total number of hydrogen bonds of a single B subunit with its neighboring molecules is therefore at least 26. The best-defined hydrogen bonds are between strand $\beta 2$ and $\beta 6$ (involving 7 hydrogen bonds), between the central helices (involving 4 hydrogen bonds) and between the turn 31–35 of one subunit and Glu11, Tyr12 as well as Gln61 of another subunit. Each subunit forms at least four salt bridges with each neighbor, giving a total of more than 20 inter-subunit salt bridges in the pentamer, with ten of those localized in the central pore.

In addition to all these polar contacts there is extensive packing of subunits against each other. Figure 11(a) gives a representation of this subunit packing by showing the difference in accessible surface of the monomer and the dimer subunits. Approximately 40% of the total accessible surface of the B monomer is buried upon pentamer formation, involving about $2 \times 1350 \text{ \AA}^2$. Of the buried surface, approximately 59% is hydrophobic, somewhat more than the overall value of 54% for the complete monomer.

Mutation of a residue in the long helical pore, Ala64 to Val in porcine LT, resulted in the defective formation of B oligomers (Iida *et al.*, 1989), and prevented the formation of the AB_5 complex. This mutation probably interferes with the packing of the C^β atom of Ala64 to the C^β atom of Met31 subunit ($N+1$) and may also prevent the formation of a 5-fold repeated water channel (see section (h), below), which could stop the correct formation of the pentamer.

(d) The A1 fragment

The enzymatic A1 fragment consists of a single domain with a wedge shape (Fig. 12), formed by numerous but not very regular secondary structure elements (Table 3). A main-chain hydrogen bond scheme is given in Figure 9(b). Only 65.4% of the possible main-chain hydrogen bonds are formed, in contrast to 74.4% in the B pentamer, which fits with the high flexibility and low levels of secondary structure of the A subunit as identified by FTIR methods (Surewicz *et al.*, 1990). The secondary structure is essentially as given in Sixma *et al.* (1991), but most secondary structure elements have

Table 4
Hydrogen bonding

A. Hydrogen bonds between B subunits†									
B(N)		B(N-1)		Distance (Å)‡					
25	O	103	N	2.6	2.8	2.8	2.8	2.8	
27	N	101	O	2.9	3.0	3.0	2.8	2.9	
27	O	101	N	2.9	2.9	2.9	2.9	2.9	
29	N	99	O	3.0	3.0	2.9	3.0	3.0	
29	O	99	N	2.9	3.0	2.9	3.1	3.1	
31	N	97	O	2.9	2.9	3.0	3.0	2.8	
92	O	1	N	2.7	2.6	2.7	2.5	—	
23	Lys	N ⁵	103	Asn	O	—	3.2	—	3.1
31	Met	O	61	Gln	N ²	3.0	2.9	2.9	3.1
33	Gly	O	12	Tyr	O ⁿ	2.7	2.6	2.6	2.9
35	Arg	N ²	11	Glu	O ²¹	2.9	—	2.9	—
35		N ⁿ¹	12	Tyr	O ⁿ	—	—	—	2.9
66	Glu	O ²²	63	Lys	N ⁵	—	3.0	—	—
66		O ²¹	67	Arg	N ⁿ²	2.8	3.0	3.1	3.1
66		O	67		N ⁿ²	3.3	3.0	3.2	3.1
70	Asp	O ²²	67		N ⁿ¹	2.9	2.8	2.8	2.9
70		O ²²	67		N ⁿ²	—	—	—	3.2
73	Arg	N ⁿ¹	71	Thr	O ⁿ¹	3.0	2.8	2.8	2.7

B. Hydrogen bonds between A1 and B, A2 and B and A1 and A2§									
A1			B			Distance (Å)			
148	Arg	N	79	Glu	O ²¹	B#4			3.2
		N ⁶	103	Cterm	OXT	B#3			2.8
		N ⁿ²	76	Tyr	O	B#4			3.0
151	Arg	N ⁿ²	79	Glu	O ²²	B#4			2.9

A2			B			Distance (Å)			
217	Lys	N ⁵	103	Cterm	OXT	B#2			2.8
220	Arg	N ⁶	77	Leu	O	B#3			2.8
		N ⁿ²	76	Tyr	O	B#3			3.0
221	Gln	N ²	78	Thr	O	B#3			2.8
229	Glu	O ²¹	73	Arg	N ⁿ²	B#3			2.9
235	Arg	N ⁿ²	70	Asp	O ²¹	B#1			3.0

A2			A1			Distance (Å)			
204	Gln	N ²²	37	Met	O				2.8
214	Tyr	O ⁿ	123	Gln	O ²¹				2.9
225	Asp	O ²¹	146	Arg	N ⁿ¹				3.1
		O ²²			N ⁿ²				3.0

† Cut off at 3.2 Å and maximum deviation from linearity of 65°. The first 7 hydrogen bonds are main-chain-main-chain contacts within the β -sheet.

‡ Distances are given for the five subunit contacts, starting with N = 1, i.e. B#1-B#5; B#2-B#1; B#3-B#2, etc.

§ Cut off at 3.2 Å and maximum deviation from linearity of 65°.

|| Numbering of B subunits is defined in the legend to Fig. 7.

been slightly extended by the high resolution refinement. An exception is strand $\beta 9$, which is now split into two parts. This region is, however, somewhat less well defined in this high-resolution structure than in the 2.3 Å electron density. There is a small additional 3_{10} helix comprising residues 108 and 110.

The overall structure of A1 is not very simple to describe: there are two anti-parallel β -sheets, formed by strands $\beta 5$, $\beta 6$, $\beta 3$ and $\beta 1$, $\beta 4$, $\beta 8$, $\beta 9$ which make up the core of the subunit. These are linked together *via* a single pair of hydrogen bonds between $\beta 1$ and $\beta 3$ (Tyr6-Ser 63). These sheets can also be visualized as one single anti-parallel sheet with a very strong twist between $\beta 1$ and $\beta 3$. The

third, tiny sheet links “strand” $\beta 2$ and $\beta 7$ by a single pair of hydrogen bonds. This sheet closes off an open area originating from the large angle between strand $\beta 1$ and $\beta 3$. There is an equivalent empty area at the other side of the main sheet. This is near the residues that have been implicated in the activity of the subunit (Sixma *et al.*, 1991). Most loops connecting strands in the main sheet are quite long and contain α -helices, leaving only one simple β -hairpin, between strand $\beta 8$ and $\beta 9$, with a classical type I' conformation (Sibanda *et al.*, 1989). The region from residue 131 to 164 makes a long loop around the subunit ending in a cluster of three small helices, comprising residues 157 to 182, which forms a small “protrusion” at one end of the subunit (Fig. 12).

(e) The A2 fragment

The A2 fragment (A2:195-240) consists of four parts: (1) a long helix (residues 197 to 226), with residues 223 to 226 adopting a 3_{10} helical conformation; (2) a piece of extended strand (residues 227 to 231); (3) a small helix (residues 232 to 236); and (4) the last four residues (237 to 240), which are not clear in the electron density and have not been included in the analysis. A sample of the electron density of A2 within the B pore is given in Figure 13. Surprisingly perhaps, the long helix of the A2 subunit is not particularly amphipathic. This helix is kinked near the C terminus, due to residue Ser A2:224, whose O^y atom makes a hydrogen bond to the carbonyl group of Arg A2:220. At approximately this site it enters the central pore of the B pentamer (Fig. 12).

(f) The complete A subunit

In the complete A subunit almost the entire length of the long helix of A2 packs with one side against the A1 fragment. The disulfide bridge and nicking site between A1 and A2 are located in the small protrusion on one side of the A1 subunit (Fig. 12). The disulfide bridge is just visible in the electron density, whereas most residues around the nicking site (A1:189-A2:195) could not be identified, possibly because they are too flexible. A schematic diagram (Fig. 14) indicates the location of the hydrogen bonds and salt bridges between the A2 fragment and the rest of the structure. Another indication of the contacts between A1 and A2 is given in Figure 11(c), which depicts the difference in accessible surface for the A2 subunit in the presence and absence of the A1 subunit. It is very obvious from this Figure that the interactions take place on one side of the helix. Specific interactions of A1 and A2 are given in Tables 4, 5 and 6. A number of the contacts between A1 and A2 are mediated *via* water molecules, as can be seen from the data in Table 7, but there are actually very few direct hydrophilic contacts, only four hydrogen bonds and one salt bridge. The packing of A1 and A2 is, however, also not excessively hydrophobic. The

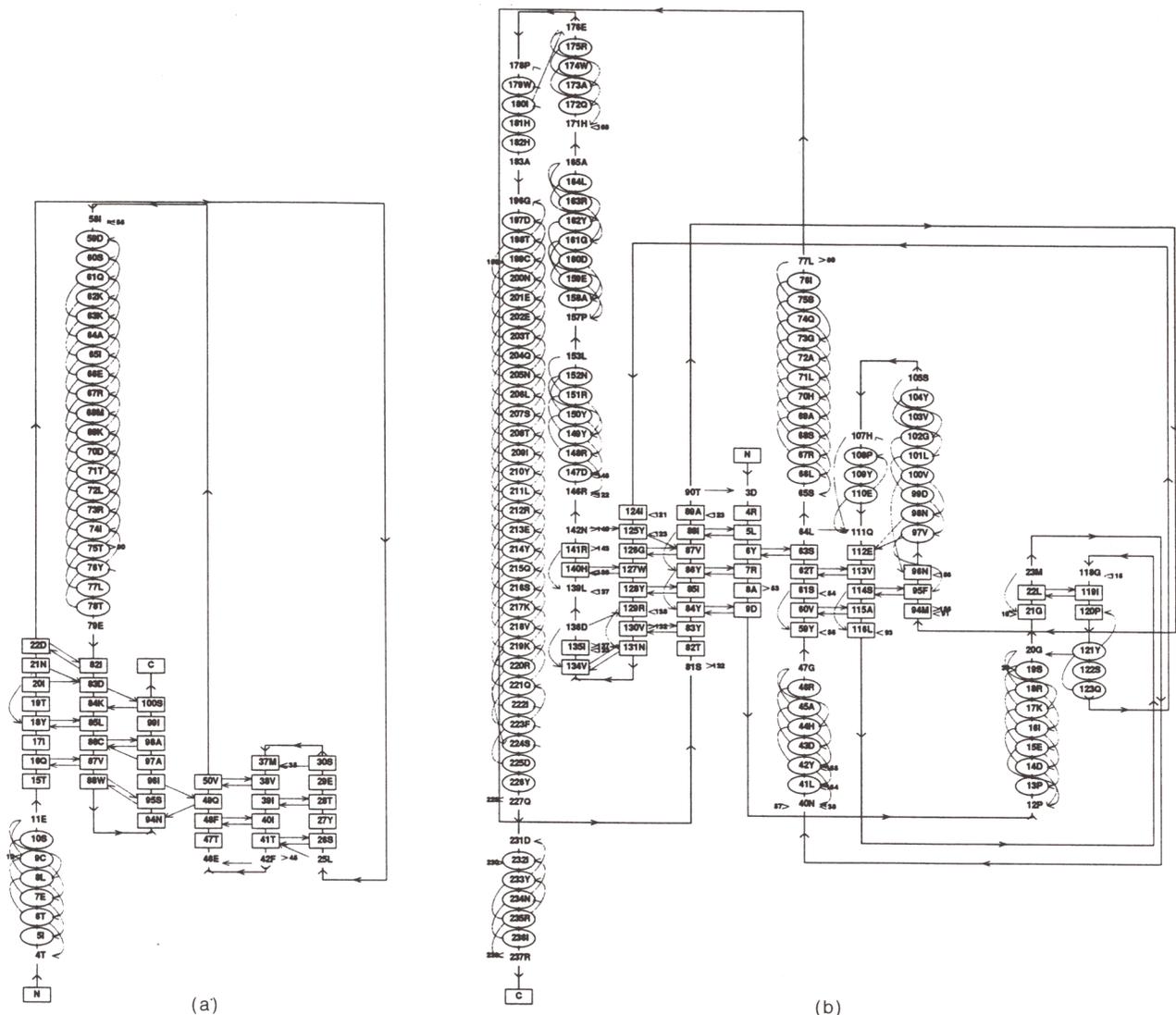


Figure 9. Main-chain hydrogen bonds in LT, determined by DSSP method (Kabsch & Sander, 1983). Hydrogen bonds with energy below a cut off at $-0.7 \text{ kcal mol}^{-1}$ ($1 \text{ cal} = 4.184 \text{ J}$), are shown in this Fig., which is produced with the program HERA (Hutchinson & Thornton, 1991). Arrows point from hydrogen bond donor to acceptor. Residues taking part in β -strands are given in boxes, residues in α -helices are given in circles. (a) B monomer, the hydrogen bond between 32 and 35 is not drawn. (b) A subunit. Note that the helix A1 : 97–104 is in reality made out of 1 helix A1 : 97–101 and 1 3_{10} helix A1 : 102–104 making an angle of about 90° with each other.

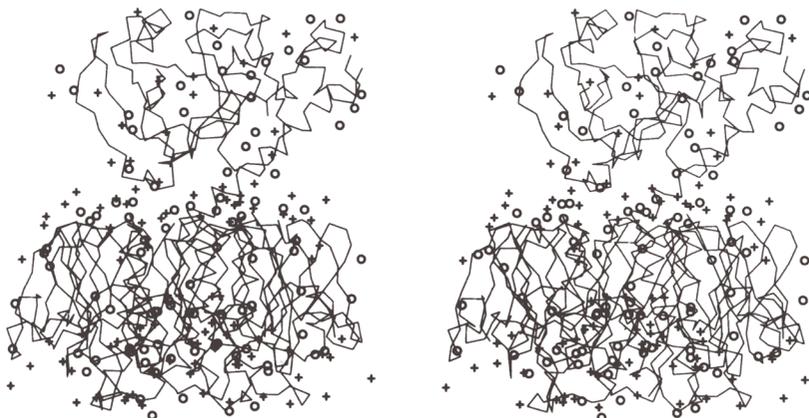


Figure 10. Stereo Fig. showing charge distribution in LT. Positively charged residues (Lys and Arg) and N termini are shown as crosses, negatively charged residues (Glu and Asp) and C termini as circles. The A subunit is shown on top, B₅ below. Two regions are not included in the Fig., A1 : 189–A2 : 196 and A2 : 237–240. Both of these regions contain additional charges, but were not visible in electron density, and are therefore omitted here. Note the asymmetry of the localization of charges in the B pentamer: most charged residues are localized on the flat surface close to A or in the central pore. Note that since this Fig. is only concerned with charges and not with polar residues in general, it does not give an indication of the hydrophobicities of the surfaces.

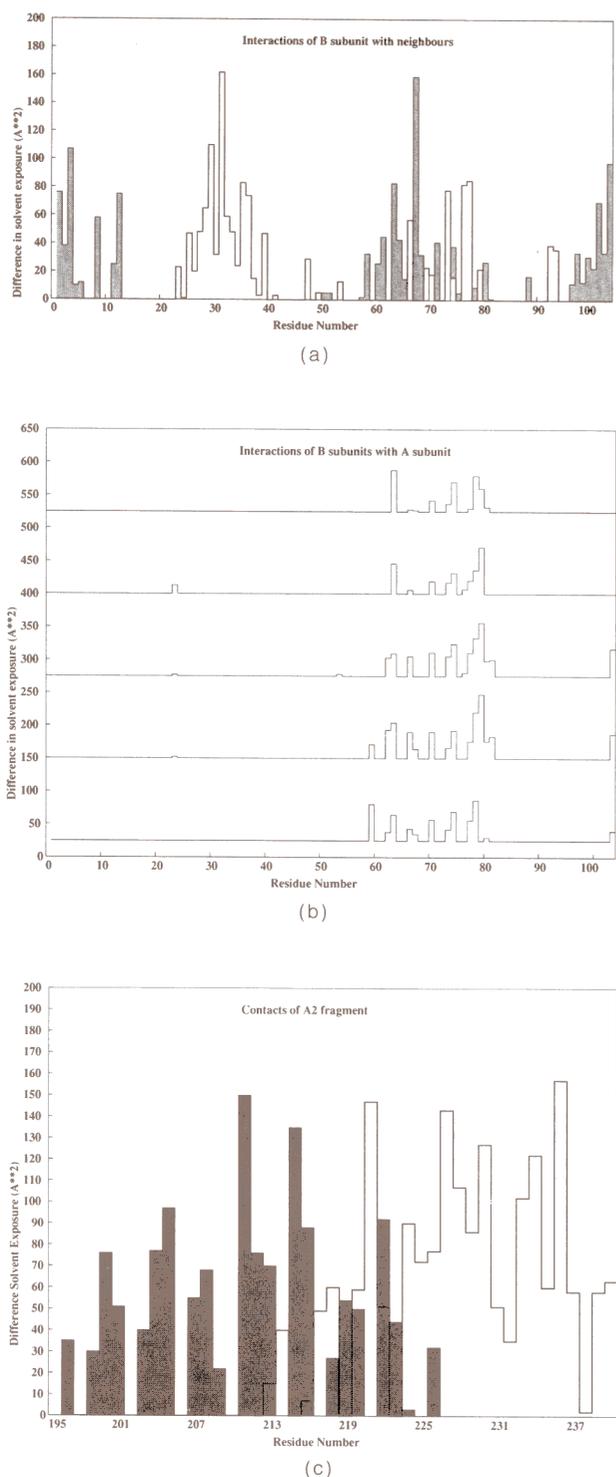


Figure 11. Subunit interactions, shown as the difference in accessible surface area in the presence and the absence of the neighboring subunit. Accessibilities were calculated with DSSP (Kabsch & Sander, 1983). (a) Interactions of B monomer with neighboring B subunits. Shaded bars represent the interactions with subunit $N+1$, open bars represent interactions with subunit $N-1$. The numbering is as in Fig. 7. (b) Interactions of B subunits with A shown for each individual B subunit. B#1 is at the bottom, B#5 at the top. (c) Interactions of A2 with the rest of the protein. Light shading: interaction of A2 with B subunits; heavy shading: interaction of A2 with A1 fragment.

Table 5
Salt bridges in LT

Internal salt bridges in B subunits				Distance (Å)†				
22	Asp	Lys	81	—	—	3.4	—	—
22	Asp	Lys	43	3.6	3.6	—	—	—
23	Lys	Glu	79	4.0	—	3.1	3.1	3.1
51	Glu	Lys	91	3.2	3.5	3.1	4.0	3.6
66	Glu	Lys	69	3.0	2.9	3.2	3.1	2.9
70	Asp	Arg	73	4.0	3.1	3.7	3.5	3.5

Internal salt bridges in A subunit				Distance (Å)				
3	Asp	Arg	4	3.8				
7	Arg	Asp	9	2.9				
9	Asp	Arg	11	2.8				
14	Asp	Arg	18	4.0				
17	Lys	Glu	144	3.2				
25	Arg	Asp	56	4.0				
43	Asp	Arg	46	2.9				
54	Arg	Glu	112	3.1				
54	Arg	Glu	110	3.6				
137	Glu	Arg	138	2.9				
143	Arg	Glu	144	3.6				
147	Asp	Arg	151	3.0				

A1 residue	B residue		Distance (Å)					
33	Arg	Glu	79	B#2	3.0			
148	Arg	Glu	79	B#4	4.0			
148	Arg	Cterm	103	B#3	3.4			
151	Arg	Glu	79	B#4	2.9			
151	Arg	Cterm	103	B#3	4.0			

A2 residue	B residue		Distance (Å)					
213	Glu	Lys	81	B#2	3.1			
217	Lys	Glu	79	B#3	3.9			
217	Lys	Cterm	103	B#2	2.8			
219	Lys	Glu	79	B#2	2.9			
220	Arg	Cterm	103	B#2	3.2			
229	Glu	Arg	73	B#3	2.9			
235	Arg	Asp	70	B#1	3.0			
235	Arg	Glu	66	B#2	3.9			

A1 residue	A2 residue		Distance (Å)					
146	Arg	Asp	225	3.0				

B(N) residue†	B(N+1) residue†		Distance (Å)					
11	Glu	Arg	35	3.3	3.0	2.7	—	2.9
103	Cterm	Lys	23	3.2	—	—	3.1	—
63	Lys	Glu	66	3.0	—	—	—	—
67	Arg	Glu	66	3.0	3.1	3.1	2.9	2.8
67	Arg	Asp	70	2.9	2.8	2.8	2.9	2.9

Cut off at 4 Å (Barlow & Thornton, 1983).

† Where 5 distances are given they refer either (top) to subunits B#1 to B#5, respectively or (below) to the interaction of subunit B(N) with B(N+1), for N=1 to 5, respectively.

solvent-accessible surface area buried by the A1–A2 interaction is 1150 Å², with 58% of this being hydrophobic surface (Connolly, 1983). The complete A subunit is also not very hydrophobic. Solvent accessibility calculations gave a value of 52.9% hydrophobic surface for the entire A subunit (of 12,253 Å²) and 52.7% for A1 alone (9832 Å²), which is much lower than found for a hydrophobic protein,

Table 6
Hydrophobic interactions

A. Hydrophobic interactions of A subunit with B subunits							
A1 residue		B residue				Distance (Å)	
148	Arg	C ^β	C ^δ	Glu	79	B#4	3.7
148	Arg	C ^ε	C ^γ	Glu	79	B#4	3.9
A2 residue		B residue				Distance (Å)	
219	Lys	C ^β	C ^α	Thr	78	B#2	4.0
220	Arg	C ^α	C ^{γ2}				4.0
		C ^γ	C ^β				3.7
223	Phe	C ^γ	C ^{γ2}	Thr	78	B#1	3.7
		C ^{δ1}	C ^{γ2}				3.6
		C ^{ε1}	C ^β				3.8
		C ^{ε1}	C ^{γ2}				3.9
		C ^ε	C ^β				3.8
224	Ser	C ^β	C ^{γ2}	Thr	78	B#3	3.6
225	Asp	C ^β	C ^{γ2}	Thr	78	B#5	3.9
226	Tyr	C ^{δ1}	C ^β	Leu	77	B#1	3.7
		C ^{δ1}	C ^{δ1}				3.7
		C ^{δ1}	C ^{γ2}	Ile	74	B#5	3.9
		C ^{δ1}	C ^{γ2}	Thr	78	B#5	3.8
		C ^{ε1}	C ^β	Leu	77	B#1	3.9
		C ^{ε1}	C ^β	Thr	78	B#5	3.6
		C ^{δ2}	C ^{γ2}	Thr	78	B#1	3.6
227	Gln	C ^γ	C ^{δ1}	Ile	74	B#3	3.8
228	Ser	C ^β	C ^{δ1}	Ile	74	B#5	3.7
232	Ile	C ^{γ2}	C ^α	Lys	63	B#2	3.6
		C ^{γ2}	C ^β				3.5
		C ^{δ1}	C ^β	Arg	67	B#2	3.9
233	Tyr	C ^{ε1}	C ^δ	Lys	63	B#2	3.8
		C ^{ε2}	C ^β	Lys	62	B#3	3.9
		C ^{ε2}	C ^γ	Glu	66	B#3	3.5
		C ^ε	C ^δ	Lys	63	B#2	3.7
235	Arg	C ^δ	C ^β	Lys	63	B#1	3.7
235	Arg	C ^ε	C ^β	Glu	66	B#1	3.8
236	Ile	C ^{γ2}	C ^γ	Glu	66	B#2	3.7
		C ^{δ1}	C ^α	Lys	63	B#2	3.9
		C ^{δ1}	C ^β		63		3.6
B. Hydrophobic contacts between A1 and A2							
A2 residue		A1 residue				Distance (Å)	
196	Gly	C ^α	C ^γ	Pro	169		4.0
198	Thr	C ^{γ2}	C ^β	Gln	185		3.9
199	Cys	C ^β	C ^β	Pro	169		3.6
		C ^ε	C	Gly	186		3.5
	Cys	S ^γ	C ^α	Pro	169		3.6
		S ^γ	C ^β				3.8
202	Glu	C ^γ	C ^β	Pro	184		3.6
		C ^δ	C ^β	Gln	185		4.0
203	Thr	C ^{γ2}	C ^α	Gly	166		3.7
204	Gln	C ^δ	C ^α	Asn	38		3.6
204	Gln	C ^δ	C	Asn	38		3.9
206	Leu	C ^β	C ^{δ2}	Leu	164		4.0
		C ^α	C ^{δ2}				3.8
207	Ser	C ^α	C ^{δ2}	Leu	164		3.8
		C ^β	C ^β	Ile	39		3.9
		C ^β	C ^{γ2}				3.8
210	Tyr	C ^{δ1}	C ^{δ2}	Leu	116		3.8
		C ^{ε1}	C ^ε	Phe	95		4.0
		C ^{ε2}	C ^ε				4.0
		C ^ε	C ^ε				3.7
211	Leu	C ^{δ1}	C ^α	Gly	117		3.8
214	Tyr	C ^γ	C ^β	Asn	93		4.0
		C ^{δ2}	C ^β				3.8
		C ^{ε2}	C ^β				3.8
		C ^{ε2}	C ^γ				3.6
		C ^ε	C ^β				3.9
		C ^ε	C ^γ				3.6
215	Gln	C ^γ	C ^{ε2}	Tyr	30		3.8
		C ^δ	C ^ε				3.9
218	Val	C ^{γ1}	C ^ε	Phe	31		3.8

Table 6 (continued)

B. Hydrophobic contacts between A1 and A2							
A2 residue		A1 residue				Distance (Å)	
219	Lys	C ^γ	C ^{ε1}	Phe	31		3.8
		C ^γ	C ^ε				3.8
		C ^ε	C ^δ				3.8
		C ^ε	C ^{ε1}				3.9
221	Gln	C ^β	C ^δ	Arg	146		3.7
		C ^γ	C ^γ	Tyr	149		3.8
		C ^γ	C ^{δ2}				3.7
221	Gln	C ^δ	C ^β	Tyr	149		4.0
		C ^δ	C ^γ				3.6
		C ^δ	C ^{δ1}				3.7
		C ^δ	C ^{δ2}				4.0
222	Ile	C ^{δ1}	C ^{ε2}	Phe	31		3.9

All carbon-carbon, or carbon-sulfur distances less than 4 Å are shown. Cys-Cys contacts have been excluded from the table.

such as crambin, where 67% of the solvent-accessible surface is hydrophobic (Miller *et al.*, 1987).

(g) *The AB₅ structure*

The interactions of the A subunit and the B pentamer are detailed in Tables 4 through 7 and illustrated in Figure 14. The major contacts are localized at the C-terminal end of the central pore in the B pentamer, where the A2 fragment enters it. Here the few direct contacts of A1 with the B pentamer can be found. These are made by three arginine residues of A1 (A1:33, 148 and 151) involved in salt bridges (Table 5), while the guanidinium group of arginine A1:148 is also involved in two hydrogen bonds to the C terminus of the central helix of B#4, implying an electrostatic interaction with the helix dipole. Except for the single hydrophobic interaction between the side-chains of Arg A1:148 and B#4:Glu79, all A1-B interactions are thus electrostatic interactions.

In the A2 fragment, the helix dipole is also used: Arg A2:220 interacts with the long helix of B#3, where the carbonyl groups of the C terminus of the helix make hydrogen bonds to the arginine side-chain. In this region there are five more salt bridges between A2 and B pentamer, involving residues Glu A2:213, Lys A2:217 and A2:219 and Arg A2:220. Evidently, many interactions at the entrance of the pore are charged, which may explain the flexibility of the A subunit with respect to the B pentamer seen in different crystal forms of the AB₅ complex (Sixma *et al.*, 1992c). There are two regions of hydrophobic interactions between A2 and the B pentamer. One of these is at the narrow end of the central pore where the A2 fragment forms a helix (up to residue A2:222) and a ₃10 helix (A2:222-226). There the pore is quite hydrophobic, lined by Leu B:77 and Ile B:74, as well as by the hydrophobic methyl group of Thr B:78 (while its O^γ atom forms a hydrogen bond to the carbonyl oxygen atom of B:74, partly causing the curve of the long helices). No water molecules are found in

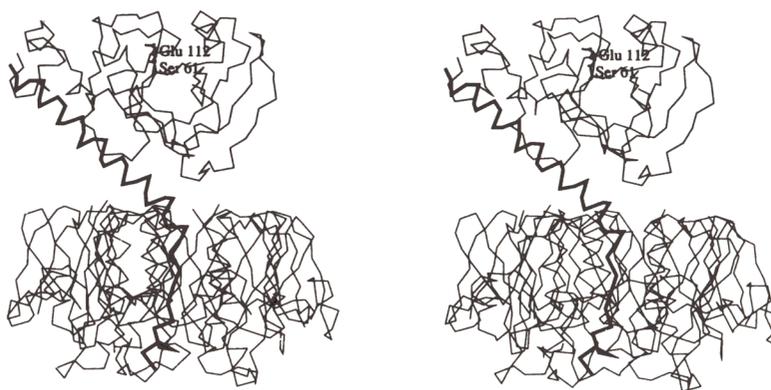


Figure 12. C α atom tracing of heat labile enterotoxin, with the A2 fragment drawn as thick lines, showing the complete AB₅ structure. The A subunit is shown on top, the B subunit is below. Regions which are not visible in electron density, i.e. residues A1 : 1–3, A1 : 188–195 and A2 : 237–240, have been omitted. Residues A1 : Ser61 and A1 : Glu112 are shown in full, to indicate the putative active site.

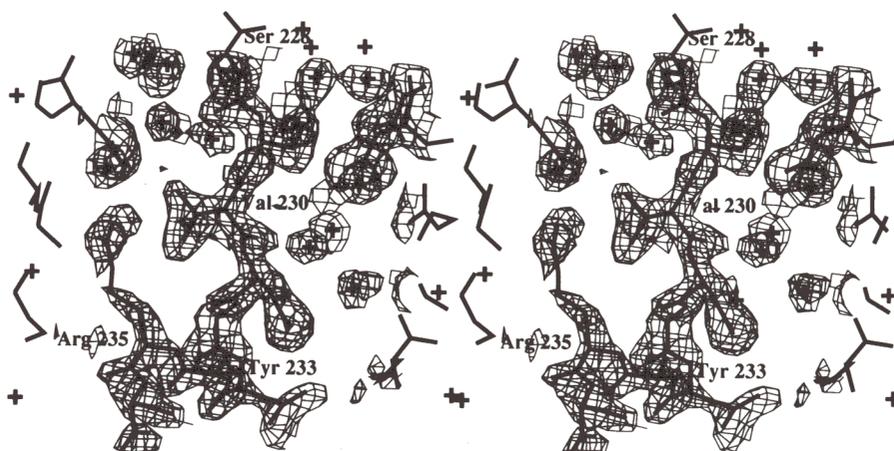


Figure 13. Sample of final electron density of the A2 fragment, passing through the central pore, and some of the many water molecules located in the pore. Coefficients were $2F_o - F_c$, with SIGMAA weights and phases calculated from the model. The start of the small C-terminal helix is visible, showing that the electron density for Arg A2 : 235 is not perfect.

Table 7
Water-mediated contacts involving A2

A. Waters shared between A2 and A1 subunits									
Residue		Distance (Å)		Residue		Distance (Å)			
200	Asn	N ^{δ2}	2.8	Water	O	167	Phe	3.1	
203	Thr	O ^γ	2.8	Water	N	166	Gly	2.8	
B. Waters shared between A2 and B subunits				Residue		Distance		B subunits	
Residue		Distance (Å)							
217	Lys	O	3.0	Water	O ^{ε1}	79	Glu	2.9	B # 3
221	Gln	N ^{ε2}	2.9						
227	Gln	N ^{ε2}	2.9	Water	O	70	Asp	2.8	B # 2
228	Ser	O ^γ	2.6	Water	N ^{η2}	73	Arg	2.8	B # 5
233	Tyr	O ^η	2.5	Water	O ^{ε2}	66	Glu	2.9	B # 3
219	Lys	N ^ε	3.2	Water	O	77	Leu	2.7	B # 2

Hydrogen bonds are cut off at 3.2 Å and maximal deviation from linearity of 65°.

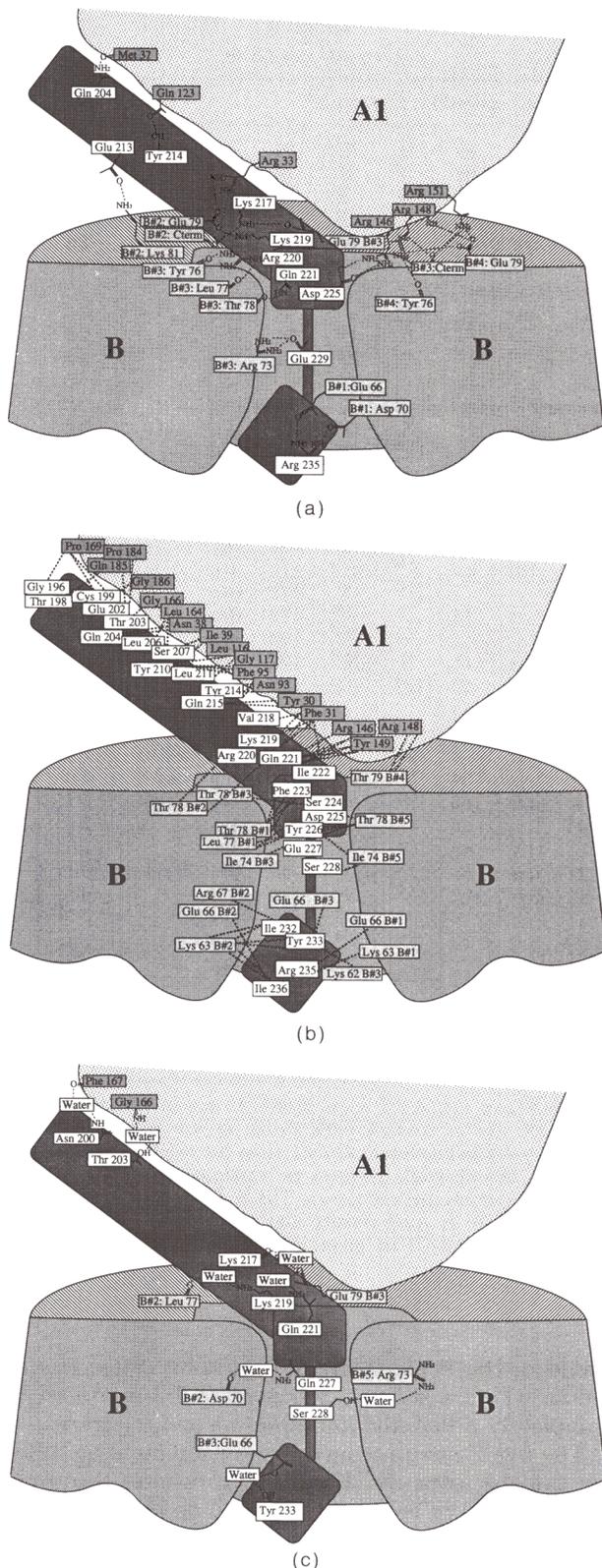


Figure 14. Schematic view of A2 fragment (A2:196–240, heavy shading), showing its interactions with A and B subunits. The A2 fragment mediates the interaction between A1 and B. Note that in this view subunit B#3 should be behind the A2 fragment, while B#1 should be in front of it. (a) Hydrogen bonds and salt bridges; (b) hydrophobic contacts; (c) water-mediated contacts. See Tables 4 to 6 for criteria and distances of the interactions.

this hydrophobic region of the pore (extending 7 Å along the 5-fold axis). The second region of hydrophobic A2:B interactions occurs at the wide end of the pore where residues A2:232–236 form a small helix. Several residues of this helix are involved in contacts with Lys B:62, Lys B:63 and Glu B:66. In the central part of the pore residues A2:227–231 are in an extended conformation. Further contacts are mainly indirect *via* a number of water molecules, listed in Table 7, and discussed below in section (h), except for two salt bridges: Glu A2:229 makes a salt bridge to Arg B#3:73 and Arg A2:235 to both Asp B#1:70 and Glu B#2:66.

The excess positive charge (+5) in the middle part of the central pore is partially compensated by the A2 fragment. Residues Glu A2:229, Asp A2:231 and Arg A2:235 are buried in the pore. The glutamic acid and the arginine residues are involved in the salt bridges mentioned above, while Asp A2:231 is located at the N terminus of the small A2 helix, formed by residues A2:232–236, in electrostatic interaction with the helix dipole of the small helical region, while it does not have direct contact with the B subunits. This leaves a net charge of +4 in the central section of the pore.

The fact that the A subunit cannot assemble with complete B pentamers to form AB₅ complexes *in vitro* (Hardy *et al.*, 1988), can be explained by the alternation of the hydrophobic and hydrophilic contacts in the central pore, which might hinder the insertion of A2 into a preformed pore. Another possible reason is that the C-terminal helical fragment of the A2 may be too “wide” to fit into the narrow end of the central pore. This possibility assumes a significant stability of the C-terminal A2 helix, which is by no means certain.

Jobling & Holmes (1991) showed that mutants of Arg B:35 to Glu or Asp affected holotoxin formation. This is surprising because this arginine is located very far from the A subunit, close to the galactose-binding site. The effect of the mutation is the loss of the inter-B-subunit salt bridge between Arg B:35 and Glu B(N+1):11 (Table 5) and the introduction of a repulsion between two negatively charged residues. This may destabilize the B–B interactions and cause the formation of an imperfect pentamer, which consequently may loosen the interaction with the A subunit.

Introduction of extensions of the C terminus in the B subunits by Sandkvist *et al.* (1987) also interferes with the AB₅ formation. This is probably mainly due to steric hindrance, because the A subunit and the C terminus of the B subunits are located on the same face of the pentamer. The disappearance of the salt bridges, which are normally formed between the C termini of subunits B#2 and B#3 and the A subunit (Table 5), may also affect the complex formation.

(h) Water molecules

During refinement, 293 water molecules were positioned in the electron density, whereas no ions

could be identified. There are large numbers of water molecules located in the various subunit interfaces and in the crystal contacts. Fewer water molecules are found on the surface of the A1 subunit, except near the putative active site. This fact may be related to the higher mobility of the A subunit as has been noted in section (d), or possibly to a higher hydrophobicity of the A surface, although the hydrophobicity was not particularly high in solvent accessibility calculations (see section (f)). A number of the water positions near a B subunit are repeated in all five subunits. Sixteen different water molecules ($\times 5$) obey the 5-fold symmetry, as defined by having a distance of less than 0.5 Å from at least two of the four symmetry-related molecules, after 5-fold superposition. An additional set of four water molecules is also 5-fold repeated, but absent in one of the five subunits. Out of the 16 5-fold waters eight ($\times 5$) are located within the central pore of the pentamer. Four of these form a small water channel around each central B helix near residue Ala B:64, Glu66, Arg67 and Lys69 in the helix and near residues Glu B:29 and Ser B:30 of the subunit $N+1$ from the sheet. In addition to these 40 water molecules, there are 26 "non-5-fold" water molecules in the pore, showing a total of at least 66 resolved water molecules within the central pore.

(i) Crystal contacts

A list of the major direct crystal contacts is given in Table 8. In addition to the contacts given, there are 23 water molecules mediating crystal contacts. The crystal packing is such that the 5-fold axes are almost parallel to the a -axis, and that there are extensive crystal contacts involving the A subunit, as can be seen from the results presented in Table 8. These contacts explain why the A subunit in the crystal is, fortunately, found in a single orientation with respect to the pentamer and is not 5-fold degenerate.

(j) Fivefold symmetry

The B subunits do not form an exact pentamer, there is a small deviation from a perfect 5-fold position for the five subunits. Rotations needed to superimpose successively (Rao & Rossman, 1973; Kabsch, 1976) each subunit (N) on subunit ($N+1$) are: 72.8°, 71.6°, 71.8°, 72.5° and 71.6° (starting from $N=1$, as defined in the legend of Fig. 7), with translations along the axis of -0.18, 0.34, -0.26, 0.25 and -0.18 Å, respectively. The final orientation of the 5-fold axis is $\phi = 0.4^\circ$, $\psi = 96.0^\circ$ (using polar co-ordinates as defined by Rossmann & Blow, 1962). Superposition of the complete pentamer after rotation of 72° around this axis results in a r.m.s. co-ordinate difference of 0.6 Å (C^α atoms) and 1.0 Å (for all atoms). This deviation is very small and does not prevent the subunits from having very similar interactions with the neighboring subunits (see Tables 4A and 5). A very different situation was

Table 8
Crystal contacts

Symmetry operation: $x, y, z \pm 1$						Distance (Å)		
B #1	Asn	89	O	N ^{$\delta 2$}	89	Asn	B #3	3.0
	Asn	90	O	C ^{γ}	13	Arg	B #3	3.3
				O	14	Asn		2.8
	Thr	92	C ^{$\gamma 2$}	O ^{γ}	10	Ser		3.3
B #3	Ser	10	O ^{γ}	N ^{$\epsilon 2$}	3	Gln	B #5	2.9
Symmetry operation: $-x+1/2, y, z+1/2$								
A1	Asn	1	C ^{β}	C ^{$\epsilon 2$}	52	Phe	A1	3.3
			N	O ^{η}	55	Tyr		2.6
			O	N ^{$\eta 1$}	25	Arg		3.2
Symmetry operation: $x+1/2, -y+1/2, -z$								
B #1	Ala	1	N	O ^{$\epsilon 1$}	172	Gln	A1	3.4
B #2	Asn	14	O ^{$\delta 1$}	C ^{γ}	168	Pro		3.3
	Ser	55	C ^{β}	O	188	Gly		3.4
	Gln	56	O ^{$\epsilon 1$}	N ^{$\eta 2$}	175	Arg		2.9
			O ^{$\epsilon 1$}	O	169	Pro		2.9
			O ^{$\epsilon 1$}	O	188	Gly		3.0
	Asn	90	N ^{$\delta 2$}	O ^{$\delta 2$}	170	Asp		3.0
			N ^{$\delta 2$}	N	172	Gln		3.1
	Lys	91	C ^{α}	N ^{$\epsilon 2$}	172	Gln		3.2
			N ^{ζ}	O	170	Asp		3.4
	Thr	92	N	N ^{$\epsilon 2$}	172	Gln		2.8
B #5	Arg	35	N ^{$\eta 1$}	O	132	Phe		2.8
Symmetry operation: $-x, y+1/2, -z+1/2$								
B #1	Thr	4	O ^{$\gamma 1$}	O	44	Ser	B #5	2.9
	Lys	34	N ^{ζ}	O ^{$\delta 2$}	197	Asp	A2	3.3
	Lys	102	N ^{ζ}	O ^{$\epsilon 1$}	16	Gln	B #5	3.1
B #2	Lys	43	O	N	17	Ile		3.0
	Gly	45	C ^{α}	O	17	Ile		3.4
	Lys	84	N ^{ζ}	O ^{$\epsilon 1$}	46	Glu	B #4	3.1
	Lys	102	N ^{ζ}	N ^{$\epsilon 2$}	16	Gln		3.2
			O ^{η}	O ^{η}	18	Tyr		3.2
B #4	Arg	13	N ^{$\eta 1$}	C ^{γ}	157	Pro	A1	3.4
				O ^{$\delta 2$}	160	Asp		2.9
			N ^{$\eta 2$}	O ^{$\delta 1$}	160	Asp		3.0
				C ^{β}	159	Glu		3.4
	Ser	55	O ^{γ}	C ^{ϵ}	181	His		3.0
B #5	Asn	14	O ^{$\delta 1$}	N ^{$\epsilon 2$}	36	Gln		3.2

All contacts within 3.5 Å are listed. Atoms on the left side of the Table are in contact with those on the right side of a symmetry-related molecule. Only 1 atom per residue is listed; the one with the shortest distance to a related molecule. Water mediated contacts are not shown. The N-terminal residue of the A1 is not really in good density and therefore the contacts for symmetry operation 2 are given only as a general indication.

found in the B pentamer of verotoxin (Stein *et al.*, 1992), where a screw component of 1.3 Å caused a completely different interaction at one interface.

The five B subunits have a very similar structure, as can be seen in Figure 8(a), where the five monomers have been superimposed on an average B subunit. The deviations from this average subunit are shown in the plot of Figure 8(b), with r.m.s. differences in co-ordinate positions varying between 0.21 to 0.28 Å for all 103 C^α atoms and between 0.60 Å and 0.68 Å for all 824 non-hydrogen atoms. Deviations from the average subunit are mainly in the loop regions, around residues B:35, 45, 79, 91 and in the badly defined region B:52-59, which seems to be quite flexible in the native structure. This loop is involved in the recognition of the sugar

ligands and was shown to be much better defined after lactose binding (Sixma *et al.*, 1992a). The other deviating regions are involved in crystal contacts in at least one subunit (see Table 8) with the exception of the region around residue B:79. Apparently the crystal contacts cause some deviation of the symmetry. In contrast, the interaction with the A subunit does not cause very large deviations from the average C α atom positions of the B subunits. The region involved in the interaction, between residues B:63 and B:81, shows relatively small variability, except around residues B:79, where the deviation is somewhat larger, especially in subunit B#5, which has relatively little interaction with the A subunit (Tables 4 through 7). Apparently the AB₅ structure is obtained by only minimal changes in the structure of the individual B subunits.

4. Discussion

(a) Fivefold pores

Pores exist in various other multimer proteins, mostly in tetrameric (Miller, 1989) and pentameric structures. Comparison of the central pores of LT and of the B pentamer of verotoxin (Stein *et al.*, 1992) shows that these pores have a very different character. The central pore in verotoxin would be expected to have a similar function as in LT: providing the contact area with the A subunit. Surprisingly though, it is much shorter and completely uncharged (Stein *et al.*, 1992). A pore with much more similarity to the LT central pore, is found in heavy riboflavin synthase (Ladenstein *et al.*, 1988), an $\alpha_3\beta_{60}$ protein, in which the 60 β -subunits form a capsid structure with virus-like, icosahedral symmetry. The pentamers in this complex have a central helix pore of similar size and length as the LT central pore. The helices along this pore also contain a number of charged groups, although not as continuously as in the LT helices. This pore is thought to be important for the transport of substrate molecules through the protein shell to reach the α_3 proteins localized within the capsid. Another, somewhat less similar situation was found in muconolacton isomerase (Katti *et al.*, 1989), a protein with 10-fold symmetry and point group 52. The pentamers in this complex have a cavity in their centre with a diameter varying between 11 and 40 Å, much larger than the pore in LT and lined by an even larger number of charged residues (11 per monomer of 96 residues). This cavity is not lined by parallel helices. A tetrameric protein, which has a pore with a similar large number of charges as LT, was found in annexin V (Huber *et al.*, 1990a). This pore is thought to function as a selective voltage-gated transport channel for ions. In contrast to LT (Mosser & Brisson, 1991), annexin changes the surface tension upon membrane binding, indicating that it enters the hydrophobic core of the membrane. This membrane binding is mediated by calcium, with binding sites at the end of the charged pore (Huber *et al.*, 1990b). Two putative calcium

binding sites in LT (Sixma *et al.*, 1992b), are also found at the end of the charged pore, but on the opposite face of the pentamer, with respect to the membrane-binding site.

It is clear from this overview that there is great variation in function of the pores in all these proteins, although some of them share some features with the LT pore. In contrast, the verotoxin-I pore, which would be expected to have a similar function to the LT pore has very different characteristics.

(b) Membrane binding and translocation

There is much evidence that the B pentamer does not change very much in conformation after binding to the G_{M1} ganglioside (Surewicz *et al.*, 1990; Dwyer & Bloomfield, 1982; De Wolf *et al.*, 1981; Ludwig *et al.*, 1985; Goins & Freire, 1988), but instead becomes more stable. We have shown (Sixma *et al.*, 1991) that binding of lactose, which revealed the binding site of the terminal galactose of G_{M1}, resulted in only minor changes of the B5 structure, but introduced a lower flexibility for the loop B:54 to 60. The binding site for lactose is located on the "convoluted" side of the B pentamer, far from the A subunit. This is surprising, because the A subunit has to cross a membrane to reach the cytosolic side where the substrate G_s protein is localized.

The C-terminal fragment of A2, however, is very close to the membrane after binding, and interaction of this fragment with the membrane may play a role in the translocation process. The four C-terminal residues have the sequence RDEL (or KDEL in cholera toxin) (Leong *et al.*, 1985; Yamamoto *et al.*, 1987; Mekalanos *et al.*, 1983). This type of sequence is known as an endoplasmic reticulum retention signal (Pelham, 1989, 1991). Possibly, interaction with KDEL receptors may occur also in the membrane translocation process of LT, either aiding in the transport of the A subunit over the membrane or to release A from the B pentamer (Hirst, 1991; Sixma *et al.*, 1992a).

This translocation process can occur in a number of ways: (1) by direct insertion and translocation of the A subunit; (2) by protein-mediated internalization and translocation; (3) by a membrane fusion process. Any of these processes probably occurs in the endosomes after an endocytotic process (Janicot *et al.*, 1991).

Spontaneous insertion of A1 into membranes has been shown in model systems (Wisniewski *et al.*, 1979; Wisniewski & Bramhall, 1981; Tomasi & Montecucco, 1981). The surface of the A subunit is not particularly hydrophobic (Ward *et al.*, 1981, see results, section (f)), although the low number of bound water molecules in the structure may indicate a not too hydrophilic surface. Interaction with a detergent also indicates some hydrophobicity (De Wolf *et al.*, 1987). From our structure it appears that separation of the A1 and A2 fragments releases a slightly more hydrophobic A1 area (58% hydrophobic), covered previously by A2, but this increase in hydrophobicity is not localized, but spread out

over the exposed interface surface of the A1 fragment. The lack of a distinct hydrophobic surface of the A1 fragment makes the possibility of direct membrane insertion of the structure of A1, in the conformation as observed in our X-ray structures, somewhat unlikely. However, membrane translocation may depend on rather small hydrophobic surface patches, perhaps even being created after minor conformational changes. This point needs obviously further analysis and experiments. One example of an area, in the A subunit, of possible local conformational changes, is the hydrophobic region in the protein discussed in section (c), below. Two loops, close to the active site, are packed together almost entirely by hydrophobic interaction and two salt bridges. Conformational changes in such an interface could change the surface and behavior of the protein dramatically.

A different type of direct insertion would involve the unfolding and subsequent translocation of the A subunit through the central pore of the B pentamer. This is a conceptually attractive idea (Gill, 1976), but an explanation will have to be found how the A subunit is transported through: (1) the narrow, very polar, central pore of B₅; followed by (2) the apolar, hydrophobic core of the membrane, because there is extensive evidence that the B subunits do not enter into the hydrophobic part of a membrane themselves (Wisniewski & Bramhall, 1981; Tomasi & Montecucco, 1981; Mosser & Brisson, 1991). Moreover, interactions of AB₅ with membrane receptors G_{M1} in test systems did not indicate major conformational changes of the A subunit (Surewicz et al., 1990; Van Heyningen 1982; De Wolf et al., 1981).

Interaction with other proteins (option (2)) in the translocation process has also been suggested (Hagmann & Fishman, 1981). No specific protein has yet been found with this function but, given the C-terminal "RDEL" sequence of the A2 fragment, a "KDEL"-receptor-like protein might possibly be involved (Lewis & Pelham, 1990). So far no evidence of fusion processes (option (3)) with a role in the CT-LT translocation process has been found.

The mode of insertion and translocation of the A1 fragment of LT-CT into the target cell remains an intriguing question, and it will have to be studied further.

(c) *The active site*

The activity of LT and CT is an *N*-glycosylation reaction of β -NAD⁺, linking the α -ADP-ribose to the guanidinium group of an arginine in the α -subunit of the protein G_S, and releasing nicotinamide (Moss & Vaughan, 1977; Oppenheimer, 1978; Moss et al., 1981). For activity the A subunit has to be divided into the A1 and A2 fragment by reduction of the intra-chain disulfide bond between residues A1:187 and A2:199 and proteolytic cleavage approximately after residue A1:192. It is not clear whether this process is necessary for the release of the A1 subunit during translocation, for

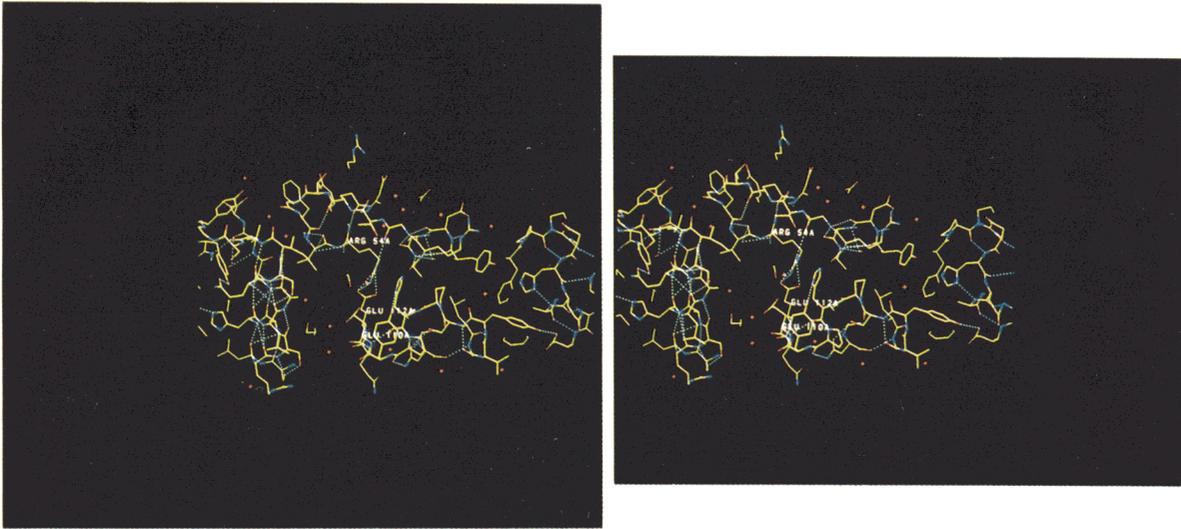
the opening up of binding sites, or for conformational changes necessary for binding or catalysis. The active site has to accommodate both NAD⁺ and an arginine side-chain of a protein molecule. In addition, interaction with small GTP-binding proteins (ARFs) has been shown to stimulate the *in vitro* activity of LT and CT (Lee et al., 1991). This effect is thought to be mediated by direct binding to A1, and was shown to lower the binding constants for the two substrates (Noda et al., 1990). Not much is known about the precise site on LT responsible for binding of ARF.

The crystal structure of LT, described in this paper, is the unnicked, unreduced form and, therefore, inactive. However, two residues with importance for activity: Glu A1:112 and Ser A1:61 (Tsujii et al., 1990, 1991; Harford et al., 1989) are spatially close in this structure. Their location in the model of the A subunit is shown in Figure 12. Comparison with the exotoxin A structure has shown that Glu A1:112 in LT has an analogous location to the catalytic residue of ETA, Glu553. It seems therefore quite likely that Glu112 performs an essential function in the catalytic centre of LT.

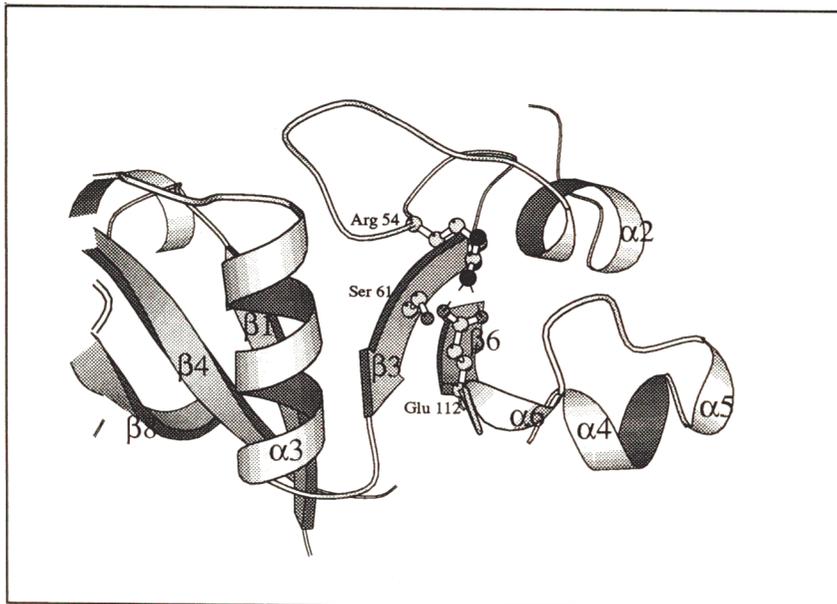
A large crevice runs across the molecule between strand β 1 and β 3 (i.e. between residues 4 to 9 and residues 59 to 63. Fig. 15 (a)(b)). It is at the interface between the two small sheets, as mentioned in the description of the A1 subunit. In the X-ray structure this crevice contains a number of solvent molecules (Fig. 15(a)) and it is lined by helix α 3 formed by residues A1:66 to 76. Surprisingly, the sequence of this helix is not conserved in CT: Ala72 is a Leu in cholera toxin, which will make the crevice narrower (see Fig. 15(a)). Glu112 is located in the middle of this crevice, on the other side from the helix (Fig. 15(a)(b)). It has been suggested that this crevice is involved in, for example, NAD⁺ binding (Sixma et al., 1991). The hydrophobic part of arginine A1:7, another residue with importance for activity, is part of the lining of this crevice, but the guanidinium group is shielded behind Val53 in the current structure (Fig. 15(c)).

The importance of arginine A1:7 for activity was shown by mutagenesis studies in both LT (Lobet et al., 1991) and CT (Burnette et al., 1991). Even a lysine at this site is not tolerated. The arginine makes a salt bridge with Asp A1:9, with its N^ε and N^η atoms, while the N^η atom is involved in good hydrogen bonds with the carbonyl groups of valine A1:53, arginine A1:54 and serine A1:61 (Fig 15(c)). Substitution by lysine would probably not disturb the salt bridge, but the hydrogen bonds on the other side of the guanidinium group would no longer be formed. The positions of Ser61 and Arg54 could very well change by loss of this backbone stabilization.

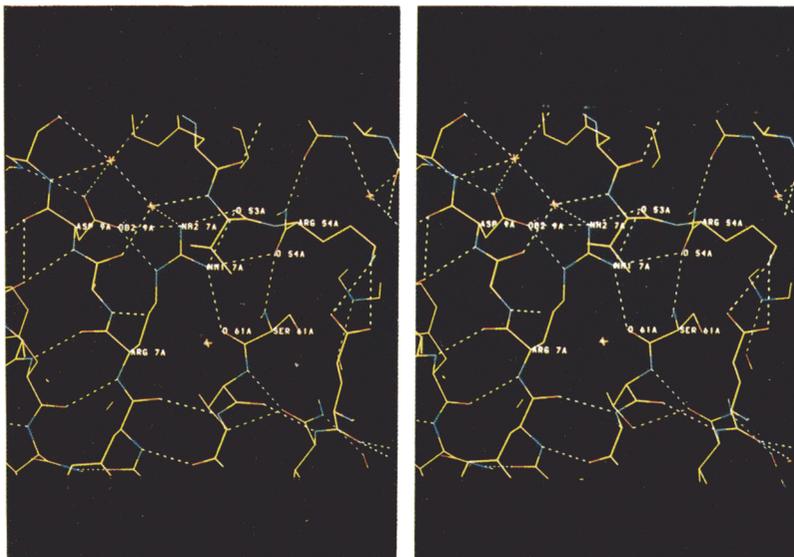
The mutations of Arg7 have been suggested by analogy to mutants in pertussis toxin (Burnette et al., 1988; Pizza et al., 1988). The S1 sequence of pertussis toxin (PT) has high sequence identity to LT and CT in this region (Nicosia et al., 1986; Lochter et al., 1986) and a mutation of the equivalent Arg9



(a)



(b)



(c)

to lysine in pertussis toxin also caused a loss in activity. It seems likely that the function of this arginine is similar in the different toxins. Analysis of the possible hydrogen bonding ligands in pertussis toxin shows that Ser61 is in another region with high homology between PT and LT, but Arg54 is not conserved. Therefore the hydrogen bond connecting Arg7 and the backbone carbonyl atom of Ser61 (and thus strand $\beta 1$ and $\beta 3$) is probably important for maintaining an active structure in LT and PT.

Although there is no sequence homology, arginine 7 in LT is in a region with structural equivalence to ETA, and by analogy also to diphtheria toxin (Sixma *et al.*, 1991; Choe *et al.*, 1992). These proteins have respectively His440 and His21 at this position. Chemical modifications of His21 in diphtheria toxin caused the loss of NAD⁺ binding (Papini *et al.*, 1989), which fits very well with the conservation of this sensitivity of this site. In the exotoxin A structure His440 makes a hydrogen bond to the equivalent carbonyl atom in strand $\beta 3$ (Tyr470), similar to that which occurs between Arg7 and Ser61 in LT. This hydrogen bond is thus very well conserved across the ADP-ribosylating toxin family. There is, however, no equivalent of the two hydrogen bonds of the Arg7 to the carbonyl atoms of residues 53 and 54 in ETA, since the entire loop between 49 and 55 has no equivalence in the ETA structure: it is a solvent-accessible area instead. The presence of this loop in LT may be the explanation of the much higher K_m for NAD⁺ in LT and CT with values of 4 mM (Galloway & van Heyningen, 1987) or 5.6 mM (Larew *et al.*, 1991) than in the other ADP-ribosylating toxins: 2.5 μ M for ETA (Lory & Collier, 1980), 9 μ M for diphtheria toxin (Lory *et al.*, 1980), 12.5 μ M for pertussis toxin (Cieplak *et al.*, 1990). The postulated importance of this loop could be tested by mutations in which residues 41 to 56 are deleted or replaced by a single residue.

In the structure, Glu A1:110 and Glu A1:112 make a salt bridge to Arg A1:54. As shown in Figure 15, these salt bridges connect the regions A1:100–110 and A1:40–54. These two regions are packed nicely against each other by exclusively hydrophobic interactions except for the two salt bridges mentioned above. This interface is at right angles to the elongated crevice mentioned above. The activity of LT and CT is known to be higher in the presence of detergents (Tsuji *et al.*, 1991; Gill &

Woolkalis, 1988). It was shown with model compounds that an acceptor compound binds better if a hydrophobic region is present next to the guanidinium group (Tait & Nassau, 1984; Soman *et al.*, 1986). A very speculative explanation could be that the hydrophobic region next to the essential Glu112 needs to open up somewhat to allow binding of the guanidinium containing receptor (e.g. G_{sz}). Analysis of solvent-inaccessible cavities in the A subunit showed a cavity of more than 35 Å³ near this hydrophobic interface, in a more hydrophilic area. This cavity is partially occupied by two water molecules and is likely to be part of the substrate binding site. Binding studies will obviously be required for confirmation of these ideas and for further analysis of the catalytic mechanism.

5. Conclusions

The structure of LT has been determined at a resolution of 1.95 Å and analyzed, with special attention for the interactions between the A subunit and the B pentamer and for the active site in the A1 fragment. The asymmetry caused by the presence of the A subunit does not cause major deviations in the structure of the individual B subunits. The A–B interface contains a mixture of hydrophobic and hydrophilic interactions, primarily mediated by the A2 fragment, which runs through the central pore of the B pentamer and emerges on the other end as a small helix. This region, with a C-terminal RDEL sequence, invisible in the electron density, is expected to interact with the membrane upon G_{M1} binding, and may be instrumental for the membrane translocation process. A putative NAD⁺-binding crevice is located near the proposed catalytic residue Glu A1:112. In this crevice the side-chain of Arg7 makes a hydrogen bond which is conserved across the ADP-ribosylating toxin family, indicating an important role for this residue. A salt bridge between Glu A1:112 and Arg A1:54 joins two loops *via* an exclusively hydrophobic interface, which also may have some function in catalysis or membrane transport.

We thank past and present members of the Groningen protein crystallography group for their support and help, in Groningen and at the synchrotron. Special thanks are due to Sylvia E. Pronk for her tireless work in the initial stages of the project. We thank Fred M. D. Vellieux for

Figure 15. The putative active site in LT. (a) Colour stereo Fig. showing that a long crevice runs from top to bottom in the center of this Fig. The proposed catalytic residue Glu A1:112 is at the side of this water-filled crevice, in the middle of the figure, making a salt bridge to Arg54. On the right hand side of Glu A1:112 is a cavity, filled with 2 more water molecules. Note a region, in this Fig. on the right hand side of Glu A1:112, where the "loops" A1:40–54 and A1:100–110 pack together *via* only hydrophobic interactions, except a salt bridge of Arg A1:54 to Glu A1:110 and Glu A1:112. The image was made using FRODO. Half-bonds are coloured according to atom type: oxygen: red; nitrogen: blue; carbon: yellow. Hydrogen bonds are indicated by broken white lines. (b) Schematic diagram (Kraulis, 1991) showing the secondary structure environment of the proposed active site. The salt bridge between Glu A1:112 and Arg A1:54 is indicated, as well as Ser A1:61, and the loops mentioned under (a). (c) Colour stereo Fig. showing the hydrogen bonding of Arg7 in LT, a residue with importance for activity. The arginine side-chain makes hydrogen bonds both to Asp9 and to the main-chain of Arg54 and Val53, Ser61, where the latter keeps together strand $\beta 1$ and $\beta 3$, which explains why even a mutation to lysine has a large effect.

his suite of averaging programs and Hillie Groendijk and Ellen S. Wartna for crystallization and heavy-atom searches in the early stages of the structure determination. We thank Mario Domenighini and Tim R. Hirst for useful discussions. This project was supported by the Netherlands Foundation for Chemical Research (SON) and the Netherlands Organization for Scientific Research (NWO). Co-ordinates described in this article have been deposited at the Brookhaven Protein Data Bank, and will be released with code 1LTS.

References

- Agarwal, R. C. (1978). A new least-squares refinement technique based on the Fast Fourier transform algorithm. *Acta Crystallogr. sect. A*, **34**, 791–809.
- Allured, V. S., Collier, R. J., Carroll, S. F. & McKay, D. B. (1986). Structure of exotoxin A of *Pseudomonas aeruginosa* at 3.0 Ångstrom resolution. *Proc. Nat. Acad. Sci., U.S.A.* **83**, 1320–1324.
- Arnone, A., Bier, C. J., Cotton, F. A., Day, V. W., Hazen, E. E. Jr, Richardson, D. C., Richardson, J. S. & Yonath, A. (1971). A high resolution structure of an inhibitor complex of the extracellular nuclease of *Staphylococcus aureus*. *J. Biol. Chem.* **245**, 2302–2316.
- Barlow, D. J. & Thornton, J. M. (1983). Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867–885.
- Barlow, D. J. & Thornton, J. M. (1988). Helix geometry in proteins. *J. Mol. Biol.* **201**, 601–619.
- Black, R. E. (1986). The epidemiology of cholera and enterotoxigenic *E. coli* diarrheal disease. In *Developments of Vaccines and Drugs against Diarrhea*. 11th Nobel Conference, Stockholm 1985 (Holmgren, J., Lindberg, A. & Möllby, R. eds), pp. 23–32. Studentlitteratur, Lund, Sweden.
- Bricogne, G. (1976). Methods and programs for direct space exploitation of geometric redundancies. *Acta Crystallogr. sect. A*, **32**, 832–847.
- Brunton, J. L. (1990). The shiga toxin family: molecular nature and possible role in disease. In *The Bacteria* (Iglewski, B. & Clark, U., eds), vol. 11, pp. 377–397. Academic Press, New York.
- Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic *R*-factor refinement by molecular dynamics. *Science*, **235**, 458–460.
- Brünger, A. T., Krukowski, A. & Erickson, J. W. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr. sect. A*, **46**, 585–593.
- Burnette, W. N., Cieplak, W., Mar, V. L., Kaljot, K. T., Sato, H. & Keith, J. M. (1988). Pertussis toxin S1 mutant with reduced enzyme activity and a conserved protective epitope. *Science*, **242**, 72–74.
- Burnette, W. N., Mar, V. L., Platler, B. W., Schlotterbeck, J. D., McGinley, M. D., Stoney, K. S., Rohde, M. F. & Kaslow, H. R. (1991). Site-specific mutagenesis of the catalytic subunit of cholera toxin: substituting lysine for arginine 7 causes loss of activity. *Infect. Immun.* **59**, 4266–4270.
- Carroll, S. F. & Collier, R. J. (1984). NAD⁺-binding site of diphtheria toxin: identification of a residue within the nicotinamide subsite by photochemical modification with NAD. *Proc. Nat. Acad. Sci., U.S.A.* **81**, 3307–3311.
- Carroll, S. F. & Collier, R. J. (1987). Active site of *Pseudomonas aeruginosa* exotoxin A. *J. Biol. Chem.* **262**, 8707–8711.
- Cassel, D. & Pfeuffer, T. (1978). Mechanism of cholera toxin action: covalent modification of the guanyl nucleotide-binding protein of the adenylate cyclase system. *Proc. Nat. Acad. Sci., U.S.A.* **75**, 2669–2673.
- Choe, S., Bennett, M. J., Fujii, G., Curmi, P. M. G., Kantardjieff, K. A., Collier, R. J. & Eisenberg, D. (1992). The crystal structure of diphtheria toxin. *Nature (London)*, **357**, 216–222.
- Cieplak, W., Jr, Loch, C., Mar, V. L., Burnette, W. N. & Keith, J. M. (1990). Photolabelling of mutant forms of the S1 subunit of pertussis toxin with NAD⁺. *Biochem. J.* **268**, 547–551.
- Connolly, M. L. (1983). Analytical molecular surface. *J. Appl. Crystallogr.* **16**, 548–558.
- Dallas, W. S. & Falkow, S. (1980). Amino acid sequence homology between cholera toxin and *Escherichia coli* heat-labile toxin. *Nature (London)*, **288**, 499–501.
- Dallas, W. S., Gill, D. M., & Falkow, S. (1979). Clusters encoding *Escherichia coli* heat-labile toxin. *J. Bacteriol.* **139**, 850–858.
- De Wolf, M. J. S., Fridkin, M. & Kohn, L. D. (1981). Tryptophan residues of cholera toxin and its A and B protomers; intrinsic fluorescence and solute quenching upon interacting with the ganglioside G_{M1}, oligo-G_{M1} or dansylated oligo-G_{M1}. *J. Biol. Chem.* **256**, 5489–5496.
- De Wolf, M. J. S., Van Dessel, G. A. F., Lagrou, A. R., Hilderson, H. J. J. & Dierick, W. S. H. (1987). pH-induced transitions in cholera toxin conformation: a fluorescence study. *Biochemistry*, **26**, 3799–3806.
- Douglas, C. M. & Collier, R. J. (1987). Exotoxin A of *Pseudomonas aeruginosa*: substitution of glutamic acid 553 with aspartic acid drastically reduces toxicity and enzymatic activity. *J. Bacteriol.* **169**, 4967–4971.
- Dwyer, J. D. & Bloomfield, V. A. (1982). Subunit arrangement of cholera toxin in solution and bound to receptor-containing model membranes. *Biochemistry*, **21**, 3227–3231.
- Dykes, C. W., Halliday, I. J., Hobden, A. N., Read, M. J. & Harford, S. (1985). A comparison of the nucleotide sequence of the A subunit of heat-labile enterotoxin and cholera toxin. *FEMS Letters*, **26**, 171–174.
- Finkelstein, R. A. (1988). Cholera, the cholera enterotoxins, and the cholera enterotoxin-related enterotoxin family. In *Immunochemical and Molecular Genetic Analysis of Bacterial Pathogens* (Owen, P. & Foster, T. J., eds), pp. 85–102, Elsevier Science Publishers.
- Fujinaga, M., Gros, P. & van Gunsteren, W. F. (1989). Testing the method of crystallographic refinement using molecular dynamics. *J. Appl. Crystallogr.* **22**, 1–8.
- Galloway, T. S. & van Heyningen, S. (1987). Binding of NAD⁺ by cholera toxin. *Biochem. J.* **244**, 225–230.
- Gill, D. M. (1975). Involvement of nicotinamide adenine dinucleotide in the action of cholera toxin *in vitro*. *Proc. Nat. Acad. Sci., U.S.A.* **72**, 2064–2068.
- Gill, D. M. (1976). The arrangement of subunits in cholera toxin. *Biochemistry*, **15**, 1242–1248.
- Gill, D. M. & Richardson, S. H. (1980). Adenosine diphosphate-ribosylation of adenylate cyclase catalyzed by heat-labile enterotoxin of *Escherichia coli*: comparison with cholera toxin. *J. Infect. Dis.* **141**, 64–70.
- Gill, D. M. & Woolkalis, M. (1988). [³²P]ADP-ribosylation of proteins catalyzed by cholera toxin and related heat-labile enterotoxins. *Methods Enzymol.* **165**, 235–245.
- Goins, B. & Freire, E. (1988). Thermal stability and intersubunit interactions of cholera toxin in solution

- and in association with its cell-surface receptor ganglioside G_{M1}. *Biochemistry*, **27**, 2046–2052.
- Gros, P., Fujinaga, M., Dijkstra, B. W., Kalk, K. H. & Hol, W. G. J. (1989a). Crystallographic refinement by incorporation of molecular dynamics: thermostable serine protease thermolysin complexed with eglin c. *Acta Crystallogr. Sect. B*, **45**, 488–499.
- Gros, P., Betzel, Ch., Dauter, Z., Wilson, K. S. & Hol, W. G. J. (1989b). Molecular dynamics refinement of a thermolysin-eglin-c complex at 1.98 Å resolution and comparison of two crystal forms that differ in calcium content. *J. Mol. Biol.* **210**, 347–367.
- Gyr, K. & Meier, R. (1987). Acute infectious diarrhea. *Trop. Med. Parasit.* **38**, 236–238.
- Hagmann, J. & Fishman, P. H. (1981). Inhibitors of protein synthesis block action of cholera toxin. *Biochem. Biophys. Res. Commun.* **98**, 677–684.
- Hansson, H.-A., Holmgren, J. & Svennerholm, L. (1977). Ultrastructural localization of cell membrane G_{M1} ganglioside by cholera toxin. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 3782–3786.
- Hardy, S. J. S., Holmgren, J., Johansson, S., Sanchez, J. & Hirst, T. R. (1988). Coordinated assembly of multi-subunit proteins: oligomerization of bacterial enterotoxins *in vivo* and *in vitro*. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 7109–7113.
- Harford, S., Dykes, C. W., Hobden, A. N., Read, M. J. & Halliday, I. J. (1989). Inactivation of the *Escherichia coli* heat-labile enterotoxin by *in vitro* mutagenesis of the A-subunit gene. *Eur. J. Biochem.* **183**, 311–316.
- Hendrickson, W. A. & Lattman, E. E. (1970). Representation of phase probability distributions for simplified combination of independent phase information. *Acta Crystallogr. Sect. B*, **26**, 136–143.
- Hirst, T. R. (1991). Assembly and secretion of oligomeric toxins by gram negative bacteria. In *Sourcebook of Bacterial Protein Toxins* (Alouf, J. & Freer, J., eds), pp. 124–132, Academic Press, New York.
- Hirst, T. R., Sanchez, J., Kaper, J. B., Hardy, S. J. S. & Holmgren, J. (1984). Mechanism of toxin secretion by *Vibrio cholerae* investigated in strains harboring plasmids that encode heat-labile enterotoxins of *Escherichia coli*. *Proc. Nat. Acad. Sci., U.S.A.*, **81**, 7752–7756.
- Hofstra, H. & Witholt, B. (1984). Kinetics of synthesis, processing and membrane transport of heat-labile enterotoxin, a periplasmic protein in *Escherichia coli*. *J. Biol. Chem.* **259**, 15182–15187.
- Hofstra, H. & Witholt, B. (1985). Heat-labile enterotoxin in *Escherichia coli*. Kinetics of association of subunits into periplasmic holotoxin. *J. Biol. Chem.* **260**, 16037–16044.
- Huber, R., Römisch, J. & Paques, E.-P. (1990a). The crystal and molecular structure of human annexin V, an anticoagulant protein that binds to calcium and membranes. *EMBO J.* **9**, 3867–3874.
- Huber, R., Schneider, M., Mayr, I., Römisch, J. & Paques, E.-P. (1990b). The calcium binding sites in human annexin V by crystal structure analysis at 2.0 Å resolution. Implications for membrane binding and calcium channel activity. *FEBS Letters*, **275**, 15–21.
- Hunt, P. D. & Hardy, S. J. S. (1991). Heat-labile enterotoxin can be released from *Escherichia coli* cells by host intestinal factors. *Infect. Immun.* **59**, 168–171.
- Hutchinson, E. G. & Thornton, J. M. (1990) HERA—a program to draw schematic diagrams of protein secondary structures. *Proteins*, **8**, 203–212.
- Iida, T., Tsuji, T., Honda, T., Miwatani, T., Wakabayashi, S., Wada, K. & Matsubara, H. (1989). A single amino acid substitution in B subunit of *Escherichia coli* enterotoxin affects its oligomer formation. *J. Biol. Chem.* **264**, 14065–14070.
- Janicot, M. & Desbuquois, B. (1987). Fate of injected ¹²⁵I-labelled cholera toxin taken up by rat liver *in vivo*. *Eur. J. Biochem.* **163**, 433–442.
- Janicot, M., Fouque, F. & Desbuquois, B. (1991). Activation of rat liver adenylate cyclase by cholera toxin requires toxin internalization and processing in endosomes. *J. Biol. Chem.* **266**, 12858–12865.
- Jobling, M. G. & Holmes, R. K. (1991). Analysis of structure and function of the B subunit of cholera toxin by the use of site-directed mutagenesis. *Mol. Microbiol.* **5**, 1755–1767.
- Jones, T. A. (1978). Graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–272.
- Jones, T. A. & Thirup, S. (1986). Using known substructures in protein model building and crystallography. *EMBO J.* **5**, 819–822.
- Joseph, K. C., Kim, S. U., Stieber, A. & Gonatas, N. K. (1978). Endocytosis of cholera toxin into neuronal GERL. *Proc. Nat. Acad. Sci., U.S.A.* **75**, 2815–2819.
- Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A*, **32**, 922–923.
- Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
- Kahn, R. A. & Gilman, A. G. (1984). Purification of a protein cofactor required for ADP-ribosylation of the stimulatory regulatory component of adenylate cyclase by cholera toxin. *J. Biol. Chem.* **259**, 6228–6234.
- Katti, S. K., Katz, B. A. & Wyckoff, H. W. (1989). Crystal structure of muconolactone isomerase at 3.3 Å resolution. *J. Mol. Biol.* **205**, 557–571.
- Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. *J. Appl. Crystallogr.* **24**, 946–950.
- Kunkel, S. L. & Robertson, D. C. (1979). Purification and chemical characterization of the heat-labile enterotoxin produced by enterotoxigenic *Escherichia coli*. *Infect. Immun.* **25**, 586–596.
- Ladenstein, R., Schneider, M., Huber, R., Bartunik, H.-D., Wilson, K., Schott, K. & Bacher, A. (1988). Heavy riboflavin synthase from *Bacillus subtilis*. *J. Mol. Biol.* **203**, 1045–1070.
- Larew, J. S.-A., Peterson, J. E. & Graves, D. J. (1991). Determination of the kinetic mechanism of arginine-specific ADP-ribosyltransferases using a high performance liquid chromatographic assay. *J. Biol. Chem.* **266**, 52–57.
- Lee, C.-M., Chang, P. P., Tsai, S.-C., Adamik, R., Price, S. R., Kunz, B. C., Moss, J., Twiddy, E. M. & Holmes, R. K. (1991). Activation of *Escherichia coli* heat-labile enterotoxins by native and recombinant adenosine diphosphate-ribosylation factors, 20-kD guanine nucleotide-binding proteins. *J. Clin. Invest.* **87**, 1780–1786.
- Leong, J., Vinal, A. C. & Dallas, W. S. (1985). Nucleotide sequence comparison between heat-labile toxin B-subunit cistrons from *Escherichia coli* of human and porcine origin. *Infect. Immun.* **48**, 73–77.
- Leslie, A. G. W. (1987). A reciprocal-space method for calculating a molecular envelope using the algorithm of B. C. Wang. *Acta Crystallogr. Sect. A*, **43**, 134–136.

- Lewis, M. J. & Pelham H. R. B. (1990). A human homologue of the yeast HDEL receptor. *Nature (London)*, **348**, 162–163.
- Lobet, Y., Cluff, C. W. & Cieplak, W. Jr (1991). Effect of site-directed mutagenic alterations on ADP-ribosyltransferase activity of the A subunit of *Escherichia coli* heat-labile enterotoxin. *Infect. Immun.* **59**, 2870–2879.
- Locht, C. & Keith, J. M. (1986). Pertussis toxin gene: nucleotide sequence and genetic organization. *Science*, **232**, 1258–1264.
- Lory, S. & Collier, R. J. (1980). Expression of enzymic activity by exotoxin A from *Pseudomonas aeruginosa*. *Infect. Immun.* **14**, 1077–1086.
- Lory, S., Carroll, S. F., Bernard, P. D. & Collier, R. J. (1980). Ligand interactions of diphtheria toxin II. Ligand interactions between the NAD site and the P site. *J. Biol. Chem.* **255**, 12016–12019.
- Ludwig, D. S., Holmes, R. K. & Schoolnik, G. K. (1985). Chemical and immunochemical studies on the receptor binding domain of cholera toxin B subunit. *J. Biol. Chem.* **260**, 12528–12534.
- Machin, P. A., Wonacott, A. J. & Moss, D. (1983). *Daresbury Lab. News*, **10**, 3–9.
- Mekalanos, J. J., Collier, R. J. & Romig, W. R. (1979a). Enzymic action of cholera toxin. I. New method of assay and the mechanism of ADP-ribosyl transfer. *J. Biol. Chem.* **254**, 5849–5854.
- Mekalanos, J. J., Collier, R. J. & Romig, W. R. (1979b). Enzymic action of cholera toxin. II. Relationships to proteolytic processing, disulfide bond reduction, and subunit composition. *J. Biol. Chem.* **254**, 5855–5861.
- Mekalanos, J. J., Swartz, D. J., Pearson, D. N., Harford, N., Groyne, F. & de Wilde, M. (1983). Cholera toxin genes: nucleotide sequence, deletion analysis and vaccine development. *Nature (London)*, **306**, 551–557.
- Miller, S. (1989). The structure of interfaces between subunits of dimeric and tetrameric proteins. *Protein Eng.* **3**, 77–83.
- Miller, S., Janin, J., Lesk, A. M. & Chothia, C. (1987). Interior and surface of monomeric proteins. *J. Mol. Biol.* **161**, 641–656.
- Montesano, R., Roth, J., Robert, A. & Orci, L. (1982). Non-coated membrane invaginations are involved in binding and internalization of cholera and tetanus toxins. *Nature (London)*, **296**, 651–653.
- Moss, J. & Vaughan, M. (1977). Mechanism of action of cholera toxin. Evidence for ADP-ribosyltransferase activity with arginine as an acceptor. *J. Biol. Chem.* **252**, 2455–2457.
- Moss, J. & Vaughan M. (1988). ADP-ribosylation of guanyl nucleotide-binding regulatory proteins by bacterial toxins. *Advan. Enzymol.* **61**, 303–379.
- Moss, J. & Vaughan, M. (1991). Activation of cholera toxin and *Escherichia coli* heat-labile enterotoxins by ADP-ribosylation factors, a family of 20 kDa guanine nucleotide-binding proteins. *Mol. Microbiol.* **5**, 2621–2627.
- Moss, J., Garrison, S., Oppenheimer, N. J. & Richardson, S. H. (1979). NAD-dependent ADP-ribosylation of arginine and proteins by *Escherichia coli* heat-labile enterotoxin. *J. Biol. Chem.* **254**, 6270–6272.
- Moss, J., Osborne, J. C., Jr, Fishman, P. H., Nakaya, S. & Robertson D. C. (1981). *Escherichia coli* heat-labile enterotoxin. Ganglioside specificity and ADP-ribosyltransferase activity. *J. Biol. Chem.* **256**, 12861–12865.
- Mosser, G. & Brisson, A. (1991). Conditions of two-dimensional crystallization of cholera toxin B-subunit on lipid films containing ganglioside GM1. *J. Struct. Biol.* **106**, 191–198.
- Nicosia, A., Perugini, M., Franzini, C., Casagli, M. C., Borri, M. G., Antoni, G., Almoni, M., Neri, P., Ratti, P. & Rappuoli, R. (1986). Cloning and sequencing of the pertussis toxin genes: operon structure and gene duplication. *Proc. Nat. Acad. Sci., U.S.A.* **83**, 4631–4635.
- Noda, M., Tsai, S.-C., Adamik, R., Moss, J. & Vaughan, M. (1990). Mechanism of cholera toxin activation by a guanine nucleotide-dependent 19 kDa protein. *Biochim. Biophys. Acta*, **1034**, 195–199.
- Oppenheimer, N. J. (1978). Structural determination and stereospecificity of the cholera toxin catalyzed reaction of NAD⁺ with guanidines. *J. Biol. Chem.* **253**, 4907–4910.
- Papini, E., Schiavo, G., Sandona, D., Rappuoli, R. & Montecucco, C. (1989). Histidine 21 is at the NAD⁺ binding site of diphtheria toxin. *J. Biol. Chem.* **264**, 12385–12388.
- Pelham, H. R. B. (1989). Control of protein exit from the endoplasmic reticulum. *Annu. Rev. Cell. Biol.* **5**, 1–23.
- Pelham, H. R. B. (1991). Recycling of proteins between the endoplasmic reticulum and Golgi complex. *Curr. Opin. Cell Biol.* **3**, 585–591.
- Pizza, M.-G., Bartoloni, A., Prugnola, A., Silvestri, S. & Rappuoli, R. (1988). Subunit S1 of pertussis toxin: mapping of the regions essential for ADP-ribosyltransferase activity. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 7521–7525.
- Priestle, J. P. (1988). A stereo cartoon drawing program for proteins. *J. Appl. Crystallogr.* **21**, 515–537.
- Pronk, S. E., Hofstra, H., Groendijk, H., Kingma, J., Swarte, M. B. A., Dorner, F., Drenth, J., Hol, W. G. J. & Witholt, B. (1985). Heat-labile enterotoxin of *Escherichia coli*. Characterization of different crystal forms. *J. Biol. Chem.* **260**, 13580–13584.
- Ramachandran, G. N. & Sasisekharan (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **23**, 283–438.
- Rao, S. T. & Rossmann, M. G. (1973). Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **76**, 241–256.
- Rayment, I. (1983). Molecular replacement method at low resolution: optimum strategy and intrinsic limitations as determined by calculations on icosahedral virus models. *Acta Crystallogr. sect. A*, **39**, 102–116.
- Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. sect. A*, **42**, 140–149.
- Ribi, H. O., Ludwig, D. S., Mercer, K. L., Schoolnik, G. K. & Kornberg, R. D. (1988). Three-dimensional structure of cholera toxin penetrating a lipid membrane. *Science*, **239**, 1272–1276.
- Rossmann, M. G. (1988). Determination of virus structures by the use of molecular replacement density averaging. In *Daresbury Proceedings* (Bailey, S., Dodson, E. & Phillips, S., eds), vol. 4, pp. 49–56, SERC Daresbury Laboratory.
- Rossmann, M. G. & Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr.* **15**, 24–31.
- Sack, R. B. (1975). Human diarrheal disease caused by enterotoxigenic *Escherichia coli*. *Annu. Rev. Microbiol.* **29**, 333–353.
- Sandkvist, M., Hirst, T. R. & Bagdasarian, M. (1987). Alterations at the carboxyl terminus change

- assembly and secretion properties of the B subunit of *Escherichia coli* heat-labile enterotoxin. *J. Bacteriol.* **169**, 4570–4576.
- Sandkvist, M., Hirst, T. R. & Bagdasarian, M. (1990). Minimal deletion of amino acids from the carboxyl terminus of the B subunit of heat-labile enterotoxin causes defects in its assembly and release from the cytoplasmic membrane of *Escherichia coli*. *J. Biol. Chem.* **265**, 15239–15244.
- Sibanda, B. L., Blundell, T. L. & Thornton, J. M. (1989). Conformation of β -hairpins in protein structures. *J. Mol. Biol.* **206**, 759–777.
- Sim, G. A. (1959). The distribution of phase angles for structures containing heavy atoms. II. A modification of the normal heavy-atom method for non-centrosymmetrical structures. *Acta Crystallogr.* **12**, 813–815.
- Sim, G. A. (1960). A note on the heavy-atom method. *Acta Crystallogr.* **13**, 511–512.
- Sixma, T. K., Pronk, S. E., Kalk, K. H., Wartna, E. S., van Zanten, B. A. M., Witholt, B. & Hol, W. G. J. (1991). Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli*. *Nature (London)*, **351**, 371–378.
- Sixma, T. K., Pronk, S. E., Kalk, K. H., van Zanten, B. A. M., Berghuis, A. M. & Hol, W. G. J. (1992a). Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography. *Nature (London)*, **355**, 561–564.
- Sixma, T. K., Terwisscha van Scheltinga, A. C., Kalk, K. H., Zhou K., Wartna, E. S. & Hol, W. G. J. (1992b). X-ray studies reveal lanthanide binding sites at the A/B₅ interface of *E. coli* heat labile enterotoxin. *FEBS Letters*, **297**, 179–182.
- Sixma, T. K., Aguirre, A., Terwisscha van Scheltinga, A. C., Wartna, E. S., Kalk, K. H. & Hol, W. G. J. (1992c). Heat-labile enterotoxin crystal forms with variable A/B₅ orientation. Analysis of conformational flexibility. *FEBS Letters*, **305**, 81–85.
- Sixma, T. K., Stein, P. E., Hol, W. G. J. & Read, R. J. (1993). Comparison of the β -pentamers of heat-labile enterotoxin and verotoxin-1: two structures with remarkable similarity and dissimilarity. *Biochemistry*, **32**, 191–198.
- Soman, G., Narayanan, J., Martin, B. L. & Graves, D. J. (1986). Use of substituted (benzylideneamino) guanidines in the study of guanidino group specific ADP-ribosyl transferase. *Biochemistry*, **25**, 4113–4119.
- Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-I from *E. coli*. *Nature (London)*, **355**, 748–750.
- Surewicz, W. K., Leddy, J. J. & Mantsch, H. H. (1990). Structure, stability, and receptor interaction of cholera toxin as studied by Fourier-transform infrared spectroscopy. *Biochemistry*, **29**, 8106–8111.
- Tait, M. & Nassau, P. M. (1984). Artificial low-molecular-mass substrates of cholera toxin. *Eur. J. Biochem.* **143**, 213–219.
- Tomasi, M. & Montecucco, C. (1981). Lipid insertion of cholera toxin after binding to G_{M1}-containing liposomes. *J. Biol. Chem.* **256**, 11177–11181.
- Tran, D., Carpentier, J.-L., Sawano, F., Gorden, P. & Orci, L. (1987). Ligands internalized through coated or noncoated invaginations follow a common intracellular pathway. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 7957–7961.
- Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr. sect. A*, **43**, 489–501.
- Tsuji, T., Inoue, T., Miyama, A., Okamoto, K., Honda, T. & Miwatani, T. (1990). A single amino acid substitution in the A subunit of *Escherichia coli* enterotoxin results in a loss of its toxic activity. *J. Biol. Chem.* **265**, 22520–22525.
- Tsuji, T., Inoue, T., Miyama, A. & Noda, M. (1991). Glutamic acid-112 of the A subunit of heat-labile enterotoxin from enterotoxigenic *Escherichia coli* is important for ADP-ribosyltransferase activity. *FEBS Letters*, **291**, 319–321.
- Van Heyningen, S. (1982). Conformational changes in subunit A of cholera toxin following the binding of ganglioside to subunit B. *Eur. J. Biochem.* **122**, 333–337.
- Vellieux, F. M. D., Kalk, K. H., Drenth, J. & Hol, W. G. J. (1990). Structure determination of quinoprotein methylamine dehydrogenase from *Thiobacillus versutus*. *Acta Crystallogr. sect. B*, **46**, 806–823.
- Ward, W. H. J., Britton, P. & van Heyningen, S. (1981). The hydrophobicities of cholera toxin, tetanus toxin and their components. *Biochem. J.* **199**, 457–460.
- Wisnieski, B. J. & Bramhall, J. S. (1981). Photolabelling of cholera toxin subunits during membrane penetration. *Nature (London)*, **289**, 319–321.
- Wisnieski, B. J., Shiflett, M. A., Mekalanos, J. & Bramhall, J. S. (1979). Analysis of transmembrane dynamics of cholera toxin using photoreactive probes. *J. Supramol. Struct.* **10**, 191–197.
- Yamamoto T., Gojobori, T. & Yokota, T. (1987). Evolutionary origin of pathogenic determinants in enterotoxigenic *Escherichia coli* and *Vibrio cholerae* O1. *J. Bacteriol.* **169**, 1352–1357.