The EMBO Journal Vol.16 No.12 pp.3426–3434, 1997

Aerolysin and pertussis toxin share a common receptor-binding domain

Jamie Rossjohn, J.Thomas Buckley[1], Bart Hazes[2], Alexey G.Murzin[3], Randy J.Read[2] and Michael W.Parker[4]

The Ian Potter Foundation Protein Crystallography Laboratory, St Vincent's Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria 3065, Australia, [1]Department of Biochemistry and Microbiology, University of Victoria, British Columbia V8W 3P6, [2]Department of Medical Microbiology and Immunology, 1–41 Medical Sciences Building, University of Alberta, Edmonton, Alberta T6G 2H7, Canada and [3]Centre for Protein Engineering, MRC Centre, Hills Road, Cambridge CB2 2QH, UK

[4]Corresponding author

We have discovered that the bacterial toxins aerolysin and pertussis toxin share a common domain. This is surprising because the two toxins affect cells in very different ways. The common domain, which we call the APT domain, consists of two three-stranded anti-parallel β-sheets that come together and wrap around a central pair of helices. The APT domain shares a common fold with the C-type lectins and Link modules, and there appears to be a divergent relationship among the three families. One surface region of the APT domain is highly conserved, raising the possibility that the domains have a common function in both proteins. Mutation of one of the conserved surface residues in aerolysin, Tyr61, results in reduced receptor binding and activity, thus providing evidence that the APT domain may be involved in interaction with the toxin's receptor. Structural and biochemical evidence suggests that the APT domain contains a carbohydrate-binding site that can direct the toxins to their target cells.

Keywords: aerolysin/carbohydrate binding/C-type lectins/pertussis toxin/toxins

Introduction

Protein toxins are produced by many organisms and they are directed towards an equally diverse range of target cells. Their lethal effects can be the consequence of a variety of activities, ranging from cell membrane disruption to inhibition of protein synthesis (Parker, 1996). In spite of their diversity, all these toxins share the ability to bind selectively to host cell surface receptors and to insert into or translocate across the plasma membrane. Here we describe an unexpected common structural domain in two very different toxins, aerolysin and pertussis toxin (PT). A similar, though more distantly related domain is also found in the C-type lectin (CTL) and Link module families. An analysis of the common structural and functional aspects of all these proteins provides new insights into the biological functions of both toxins.

The channel-forming toxin aerolysin plays a major role in the virulence of *Aeromonas hydrophila*, an important pathogen of humans and other species (Altwegg and Geiss, 1989). It is among the best characterized of all the protein toxins (Parker *et al.*, 1996). The bacteria secrete an inactive dimeric form of the toxin, proaerolysin, which is converted to aerolysin by proteolytic removal of ~45 amino acids from the C-terminus. Aerolysin binds to specific surface receptors on target cells and this promotes oligomerization, which converts the protein into an insertion-competent state. Insertion results in the generation of discrete membrane channels, causing osmotic swelling of erythrocytes and eventual lysis. The crystal structure of proaerolysin has been determined to 2.8 Å resolution (Parker *et al.*, 1994). The monomer contains two distinct lobes, a large elongated lobe which consists predominantly of β structure and a smaller, compact lobe formed by the first 83 residues of the protein (Figure 1A and B). From chemical modification and mutagenesis studies, functions in each of the above steps in channel formation have been ascribed to several regions of the large lobe. However, no function has been assigned to the smaller lobe, although we have proposed that it may help stabilize the dimeric form of the molecule (Parker *et al.*, 1996).

PT is a major virulence factor of *Bordetella pertussis*, the causative agent of whooping cough (Pittman, 1984). The holotoxin is a hexameric protein with an A–B architecture (Gill, 1978). Its structure (Figure 1C) has been determined at 2.9 Å resolution (Stein *et al.*, 1994a). The A subunit (or S1) is an ADP-ribosyltransferase which affects cellular signalling by modifying the α subunit of inhibitory trimeric G-proteins. It shares sequence and structural homology with the A subunits of other ADP-ribosylating toxins. The B-oligomer possesses the cellular recognition functionality, enabling it to interact with glycoproteins and glycolipids on many types of eukaryotic cells (Gill, 1978; Hausman and Burns, 1993). It consists of five subunits, S2, S3, S5 and two copies of S4, that together form a pseudo-symmetric pentamer which resembles the truly symmetric homopentamers of the cholera toxin family (Sixma *et al.*, 1991; Merritt *et al.*, 1994; Zhang *et al.*, 1995) and the Shiga toxin family (Stein *et al.*, 1992; Fraser *et al.*, 1994). Subunits S4, S5 and the C-terminal domains of S2 and S3 are structurally related. Surprisingly, only the primary structures of S2 and S3, which share 70% sequence identity with each other, reflect the structural similarities. These two subunits are thought to be responsible for the carbohydrate affinity of the B-oligomer, and protein crystallography has revealed a sialic acid-binding site in their C-terminal domains (Stein *et al.*, 1994b).

The CTLs form a family of proteins that usually bind carbohydrates in a calcium-dependent fashion (Drickamer, 1988). Crystal structures have been determined for mannose-binding protein (MBP) (Weis *et al.*, 1991, 1992) and,

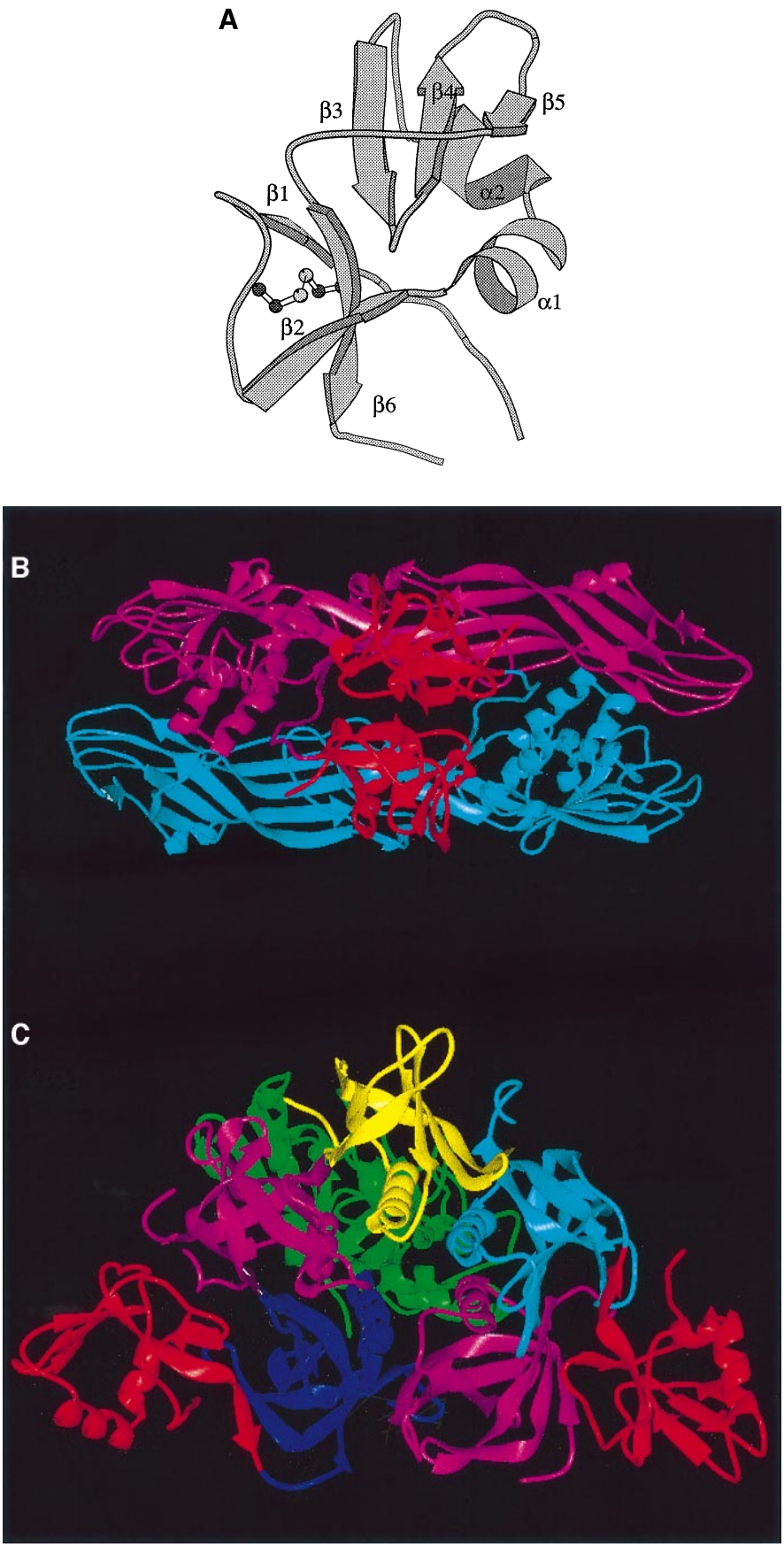

Fig. 1. (**A**) A schematic representation of the APT fold, (**B**) the position of the APT domain in the proaerolysin dimer and (**C**) the position of the APT domain in PT. The APT domain in each toxin is coloured in red. The pertussis S1, S4, S5 and the C-terminal domains of S2 and S3 subunits are coloured green, purple, yellow, light and dark blue respectively. These figures were generated using MidasPlus (Ferrin *et al.*, 1988).

more recently, for E-selectin (Graves *et al.*, 1994). The structure of an MBP–oligosaccharide complex showed that the calcium ion contributes to carbohydrate binding through direct interactions with the 3′ and 4′ hydroxyl groups of a terminal mannose residue (Weis *et al.*, 1992). A calcium ion is also present at the same site in E-selectin, and biochemical data indicate that it shares the MBP carbohydrate-binding site. However, mutagenesis results and modelling studies suggest that E-selectin has a multivalent carbohydrate-binding site, which may interact with up to three sugar residues of its natural target, the Lewis-X antigen on leukocytes (Graves *et al.*, 1994). Interestingly, a sequence similarity has been reported between PT and the CTLs (Saukkonen *et al.*, 1992). However, although the PT S2 and S3 subunits do contain a domain with structural homology to CTLs, they do not contain the calcium-dependent carbohydrate-binding site (Stein *et al.*, 1994a). A CTL fold has also been found in the crystal structure of lithostathine (Bertrand *et al.*, 1996) and in the NMR structure of Link module (Kohda *et al.*, 1996). Lithostathine prevents pancreatic stone formation. It does not contain the characteristic CTL calcium-binding site and nor does it bind carbohydrates. Nevertheless, lithostathine differs in structure from MBP and E-selectin no more than they do between themselves, and it does appear to belong to the CTL family (Bertrand *et al.*, 1996). Link module is involved in the assembly of the extracellular matrix, as well as in cell adhesion and migration. It also lacks the typical CTL calcium-binding site but it nevertheless binds carbohydrate (hyaluronan) (Brissett and Perkins, 1996).

Recently, a striking structural similarity was observed between the N-terminal domain of aerolysin and the N-terminal domains of PT subunits S2 and S3. This discovery was made independently by two of us during a search for structural homologues of aerolysin using the program DALI (Holm and Sander, 1993) and 'by eye' when updating the SCOP database (Murzin *et al.*, 1995). Based on the structural homology, conserved sequence features and the biology of these toxins, we propose that this domain serves as a lectin which facilitates host cell attachment through a carbohydrate-based receptor. New insights into the structure and function of the CTLs add support for this hypothesis and suggest that the toxin domains may have been derived from an animal host as a result of horizontal gene transfer.

Results

Common fold of aerolysin and pertussis toxin

The common fold of aerolysin and PT consists of two three-stranded antiparallel β-sheets and two α-helices (Figure 1A). The sheets are in an unusual non-co-planar arrangement, forming a pocket that accommodates both α-helices. The helices themselves are adjacent in the sequence and are joined to each other at an approximate right angle. There is a conserved disulfide bridge connecting a long surface-exposed loop to the C-terminal β-strand.

The N-terminal domains of PT subunits S2 and S3 share 63% sequence identity and they superimpose with a root-mean-square (r.m.s.) deviation on α-carbon atoms of 0.7 Å (Table I and Figure 2A). In view of the strong similarity between the S2 and S3 subunits, we will not discuss them separately in our comparisons below, but rather refer to them as S2/3. The structure of the N-terminal domain of aerolysin (~80 residues) can be superimposed almost entirely onto those of the PT domains, with r.m.s. deviations of 1.5–1.8 Å (Table I and Figure 2A). There are two insertions in the PT subunits compared with aerolysin. One of them affects the secondary structure of the common fold by forming an extra turn of helix at the beginning of helix α2 (Figure 3A). The largest insertion of six residues occurs in the long loop between strands β5 and β6. There is one insertion in aerolysin compared with both PT subunits. This insertion of one residue is located in a surface loop (Figure 3A).

Table I. Root-mean-square deviations between APT domains and C-type lectin folds

	A	P2	P3	M	E
Aerolysin domain 1 (A)	–	78	73	50	58
PT subunit S2 (P2)	1.8	–	91	54	50
PT subunit S3 (P3)	1.5	0.7	–	52	55
MBP (M)	1.9	2.4	2.3	–	99
E-selectin (E)	2.1	2.4	2.5	1.4	–

R.m.s. deviations (Å) calculated on α-carbon atoms upon superposition of the crystal structures are shown below the diagonal. The number of superimposed atoms is shown above the diagonal. Superpositions were performed with the LSQMAN option in the O-suite of programs (Jones *et al.*, 1991).

Conserved residues in the interior and on the surface

For ease of discussion, we henceforth refer to the common domain as the APT (aerolysin pertussis toxin) domain, and, unless otherwise stated, we use aerolysin sequence numbering. Figure 3A can be used to obtain the corresponding residue numbers in the PT subunits. In the structure-based alignment of the three APT domain sequences (Figure 3A), 16 of 82 residues (20%) are identical. There are 24 common residues (29%) if conservative substitutions are taken into account (Figure 3A). Given the structural homology, the level of sequence identity makes it almost certain that the APT domain in each toxin diverged from a common ancestor.

When the 16 strictly conserved residues are mapped onto the APT structure, eight are completely buried and presumably play an important role in stabilizing the overall fold. Four conserved surface residues also serve a structural role since they form a disulfide bond (Cys19–Cys75) and a salt bridge (Arg24–Glu30). The remaining four residues, Trp45, Ile47, Tyr61 and Lys66, cluster on the surface in a manner which fully exposes the lysine and the three large hydrophobic residues to the solvent (see Figure 3B). In addition, an exposed hydrophobic residue is conserved at position 57 (Met in aerolysin and Leu in S2/3 of PT). All these residues superimpose well, with Trp45, Ile47 and Tyr61 having virtually identical side chain conformations.

Search for other APT domains in the sequence database

Having concluded that the APT domain forms a compact folding unit that appears to have been added independently to aerolysin and the S2/S3 subunits of PT, we became

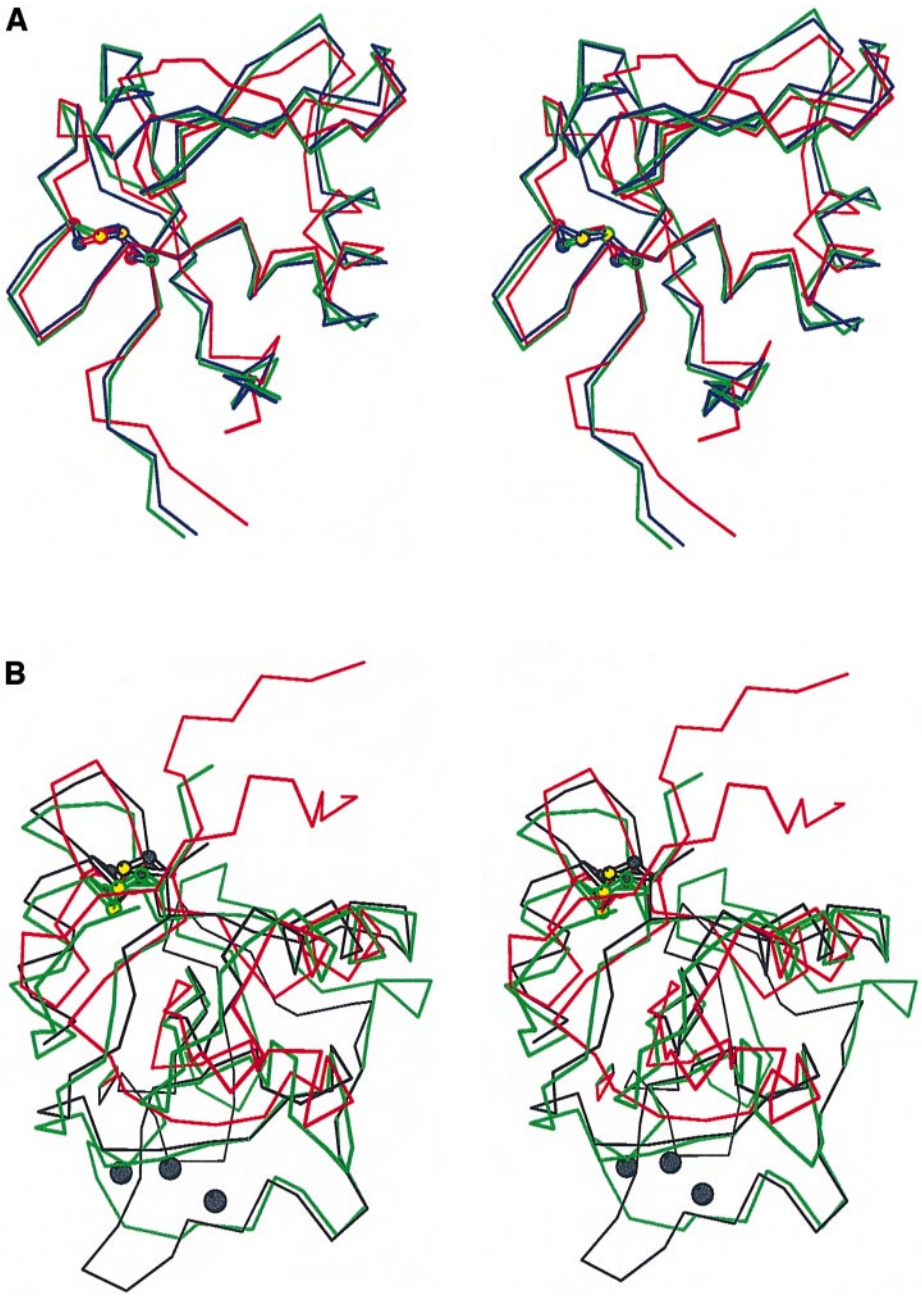

Fig. 2. (**A**) Stereo diagram of the superimposed α-carbon atoms of the APT domains. The N-terminal domain of aerolysin is coloured red and the N-terminal domains of the S2 and S3 subunits of PT are coloured green and blue respectively. The conserved disulfide bridge is shown in ball-and-stick. This figure was generated using MOLSCRIPT (Kraulis, 1991). (**B**) A stereo α-carbon superposition of the APT domain of aerolysin (red) and CTL domains of rat MBP (Weis *et al.*, 1991; Weis *et al.*, 1992) (black) and E-selectin (Graves *et al.*, 1994) (green). The calcium ions in MBP are shown as spheres. This figure was generated with MOLSCRIPT (Kraulis, 1991).

interested in determining if it occurs elsewhere. The sequences of APT domains from the different toxins have diverged to the point where their similarity is not recognized by simple searches, but the significant fraction of conserved structural and probable functional residues allows a more sensitive motif search (Bairoch *et al.*, 1996). Several different sequence motifs derived from the alignment of the three known structures (Figure 3A) were scanned against the non-redundant GenBank sequence database, but none of the scans detected any other sequences that might belong to the APT family. (The searches were performed using the ISREC PatternFind Server at Lausanne. The URL address is http://ulrec3.unil.ch/software/PATFND_mailform.html). Therefore, we concluded that so far the APT family consists of only three known members.

Relationship to the C-type lectin and the Link families

A structural homology has been reported previously between S2/3 and MBP (Stein *et al.*, 1994a), which was not detected by our sequence database search. In light of new insights into the CTLs and related proteins that have become available since that study, it was relevant to

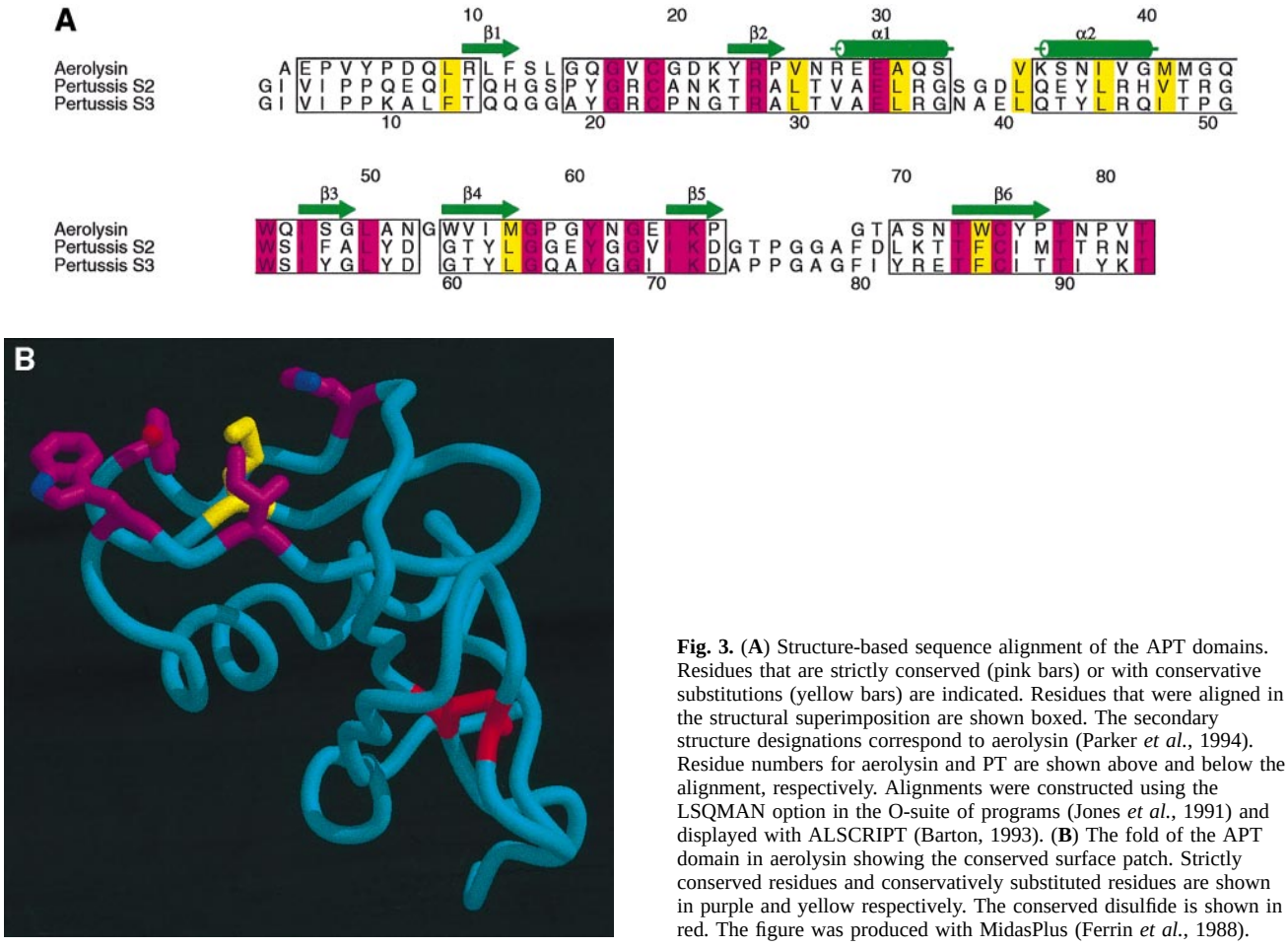

Fig. 3. (**A**) Structure-based sequence alignment of the APT domains. Residues that are strictly conserved (pink bars) or with conservative substitutions (yellow bars) are indicated. Residues that were aligned in the structural superimposition are shown boxed. The secondary structure designations correspond to aerolysin (Parker *et al.*, 1994). Residue numbers for aerolysin and PT are shown above and below the alignment, respectively. Alignments were constructed using the LSQMAN option in the O-suite of programs (Jones *et al.*, 1991) and displayed with ALSCRIPT (Barton, 1993). (**B**) The fold of the APT domain in aerolysin showing the conserved surface patch. Strictly conserved residues and conservatively substituted residues are shown in purple and yellow respectively. The conserved disulfide is shown in red. The figure was produced with MidasPlus (Ferrin *et al.*, 1988).

reconsider a possible evolutionary and functional relationship with the APT domain.

The common fold of the APT, CTL and Link families includes all six strands of the two β-sheets and the second of the two helices in the APT fold. Superposition of the APT and CTL structures yields a common core of ~52 residues with an r.m.s. deviation in α-carbon positions ranging from 1.9 to 2.5 Å (Figure 2B and Table I). Comparison with the Link module gives essentially the same common core but a higher r.m.s. deviation ranging from 2.9 to 3.1 Å (not shown). The CTL and Link folds share one more α-helix, the first in each fold, which is topologically equivalent to an extended loop in the APT fold. Conversely, the first helix in the APT fold corresponds to an extended loop in both the CTL and Link folds. Outside the common fold, the CTL and Link structures each have individual long loops. The CTLs have a characteristic loop between strands β3 and β4 which provides most of the calcium/carbohydrate-binding site. The Link modules have loops inserted after strand β1 and between strands β4 and β5. Both loops have conserved sequences that have been implicated in the binding of hyaluronan (Kohda *et al.*, 1996).

Altering one of the conserved surface residues of the APT domain reduces binding and activity

In order to understand the functional significance of the conserved surface of the APT domain, we targeted this domain for site-directed mutagenesis studies. Based on analogy with the CTLs, we suspected that the APT domain could be involved in receptor binding. Because the identity of the receptor for the APT domain of PT is not yet known, we concentrated our studies on aerolysin, for which receptor candidates have been identified (see below).

We prepared and purified four proaerolysin mutants with single amino acid replacements on different surfaces of the APT domain. Based on the above comparison, we expected that three of the mutant proteins, Ser13Cys, Glu29Cys and Met41Cys, would bind much like wild-type since they are located on surfaces other than the conserved surface, but that the fourth, Tyr61Cys, which is one of the conserved APT residues, might behave quite differently. The assay we used to compare binding was developed during recent studies on aerolysin-binding proteins. We have identified a 47 kDa glycoprotein on rat erythrocyes that binds aerolysin with high affinity (Gruber *et al.*, 1994). The protein is attached to the cell surface with a glycosylphosphatidylinositol (GPI) anchor (Parker *et al.*, 1996; S.Cowell, H.J.Gruber, W.Aschauer and J.T.Buckley, in preparation). We have found that several other proteins bind aerolysin with similar affinity. These include Thy-1 on T lymphocytes and neurones and the variable surface glycoprotein (VSG) of trypanosomes (K.Nelson and J.T.Buckley, in preparation). These proteins appear to be unrelated, with the notable exception that they all contain GPI anchors, and we have preliminary evidence that at least part of the proaerolysin recognition determinant is the glycosyl portion of the anchor. We

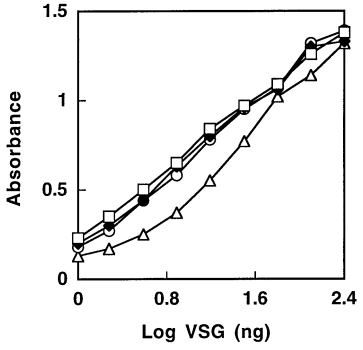

Fig. 4. Changing Tyr61 in the aerolysin APT domain affects binding. The binding of aerolysin mutants bound to variable surface glycoprotein was monitored by measuring absorbance on coated plates. Each point is the mean of duplicate determinations. See text for details. Key to the figure: wild-type is □; Ser13Cys is ◆; Met41Cys is ○; Tyr61Cys is △.

have developed an assay for aerolysin binding to VSG, described in Materials and methods, in which the purified glycoprotein is serially diluted in a microtitre plate. Using this method with wild-type proaerolysin, we can detect <1 ng of VSG. The results in Figure 4 show that Ser13Cys and Met41Cys were indistinguishable from wild-type in this assay (Glu29Cys, which is not shown, also behaved like wild-type), whereas Tyr61Cys bound less well to the receptor. Similar results were obtained in the haemolytic titre assay, where the Tyr61Cys mutant was 70–100 times less active than wild-type, while Ser13Cys and Glu29Cys had activities much closer to wild-type (data not shown). (Met41Cys cannot be compared usefully in this way because an interchain disulfide bridge is formed, inactivating the protein; Hardie *et al.*, 1995). Taken together, the *in vitro* VSG binding results and the haemolytic titres support the view that the APT domain is involved in receptor binding by aerolysin and that Tyr61 is an important amino acid in the interaction.

Discussion

The APT domain is a receptor-binding domain

Aerolysin and PT are bacterial toxins with unrelated toxic mechanisms. However, both toxins utilize cell surface receptors for initial binding. This led us to believe that the APT domain might play a role in binding to receptors. This belief is supported by the finding of a patch of conserved surface residues, including fully exposed aromatic residues. Fully solvent-exposed tryptophans and, to a lesser extent, tyrosines are uncommon in soluble proteins, but are typical in carbohydrate-binding sites (Vyas, 1991). Accordingly, the conserved surface region is a likely candidate for a carbohydrate-binding site. Interestingly, the CTL and Link module families, both of which are specialized in carbohydrate binding, also resemble the APT fold and have known or proposed binding sites that correspond topologically to the conserved surface region of the APT domain, as discussed below.

An independent indication that the APT domain contributes to receptor binding comes from a comparison of aerolysin and clostridial α-toxin. Both toxins have similar toxic mechanisms and they share 27% sequence identity (Ballard *et al.*, 1995; Parker *et al.*, 1996). However, the clostridial α-toxin lacks the APT domain and this may explain why much higher concentrations are required for channel-forming activity (~100 nM versus 1 nM for aerolysin). This observation is consistent with the proposal that the APT domain contributes to the receptor-binding affinity of aerolysin. More direct evidence in support of the idea that the APT domain provides aerolysin with high affinity binding ability is our observation that mutation of Tyr61, one of the strictly conserved surface residues in the aerolysin APT domain, reduces binding of the toxin to receptor (Figure 4). One possible orientation of the aerolysin dimer on the membrane surface brings the APT receptor-binding region close to the surface where it could bind to the GPI anchor of the receptor (Figure 5A).

Several biochemical studies have shown that PT binds to carbohydrates with terminal sialic acid residues (Armstrong *et al.*, 1988; Brennan *et al.*, 1988; Witvliet *et al.*, 1989), and protein crystallography has revealed a sialic acid-binding site in the C-terminal domains of S2/3 (Stein *et al.*, 1994b). PT and members of the cholera and Shiga toxin families have all been proposed to bind with the flat surface of their B-pentamers, opposite their catalytic subunits, in contact with the membrane (Figure 5B). This orientation is supported by lipid labelling studies on PT (Montecucco *et al.*, 1986). The orientation of the sialic acid-binding site and its distance from the flat base of the toxin indicate that this must be a glycoprotein-binding site. However, there is experimental evidence that PT also binds to various glycolipids (Saukkonen *et al.*, 1992; Hausman and Burns, 1993), implying that an additional binding site may exist. The conserved surface region of the APT domain is near the flat base of the toxin (Figure 5B), making it a prime candidate for the reported glycolipid affinity. Experimental results also suggest that the glycolipid-binding site resides in the APT domain. Loosmore *et al.* (1993) reported a holotoxin mutant in which Tyr82 (PT numbering) in the S3 subunit is replaced by an alanine. This mutant had <10% of wild-type toxicity, as measured in a Chinese hamster ovary cell clustering assay. Tyr82 is an exposed residue at the beginning of strand 6. It is not conserved in the APT domains but is adjacent to the conserved surface patch, with its side chain pointing toward the patch. It makes van der Waals interactions with the conserved isoleucine residue, suggesting it may contribute to receptor binding. The corresponding residue in the S2 subunit is a leucine, which also interacts with the conserved isoleucine residue. Other studies that point to a role for the APT domain in carbohydrate binding have used synthetic peptides or mutagenesis of the isolated S2 and S3 domains (Saukkonen *et al.*, 1992; Rozdzinski *et al.*, 1993; Tallett *et al.*, 1993; Heerze *et al.*, 1995).

Topological conservation of the carbohydrate-binding sites

The CTL and Link module domains are known to recognize cell surface oligosaccharide structures (Drickamer, 1988; Kohda *et al.*, 1996). If the structural similarity between these proteins and the APT domain is accompanied by a functional similarity then it would be reasonable to expect that their carbohydrate-binding sites are also structurally related. In an earlier study, the PT S2/3 subunits have been compared with MBP (Stein *et al.*, 1994a). It was

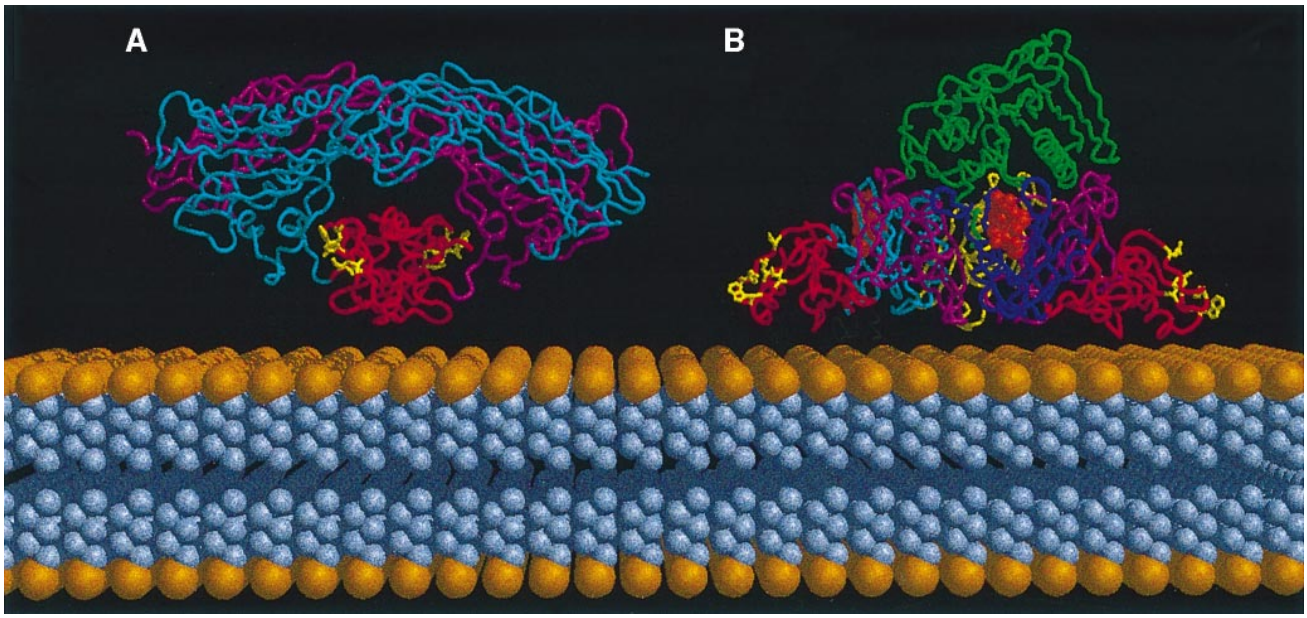

Fig. 5. Toxin interaction with target cell membranes. The APT domains are coloured red and the conserved surface patch denoted by a yellow ball-and-stick. (**A**) The proaerolysin dimer (Parker *et al.*, 1994). (**B**) Pertussis toxin (Stein *et al.*, 1994a). Sugar molecules bound to the OB-folds as seen in the crystal structure of the complex (Stein *et al.*, 1994b) are shown in orange CPK. These figures were produced using the program MidasPlus (Ferrin *et al.*, 1988).

found that the crucial loop that binds both the calcium ion and the carbohydrate in the CTL family was absent in S2/3. However, a more recent structural determination of E-selectin combined with biochemical data suggests that the MBP carbohydrate-binding site may be just one subsite of a larger, multivalent binding site (Graves *et al.*, 1994). A modelling study revealed that the natural target of E-selectin, the sialyl Lewis-X antigen, is recognized through a region that superimposes with the conserved surface patch on the APT domain in addition to the calcium-binding loop (Graves *et al.*, 1994). The hyaluronan-binding site in the Link module, which does not possess the calcium-binding loop, has been postulated to be located on a solvent-exposed hydrophobic patch of aromatic residues which is surrounded by a number of basic residues (Kohda *et al.*, 1996). Not only does the amino acid composition of this site resemble the conserved surface on the APT domain (and the E-selectin sialyl Lewis-X-binding site) but the sites are also located in equivalent positions.

The APT domain is a module

The APT domain appears to be a distinct folding unit because, in aerolysin and in S2/3, it is connected to the rest of the toxin by a long loop. The observation that the C-terminal domains of aerolysin and S2/3 are completely unrelated to each other further suggests that the APT domain was added to each of these proteins by independent gene fusion events. It follows from this reasoning that there has been an APT-less ancestral toxin, and this conclusion is supported by the existence of proteins related to both aerolysin and PT that lack the APT domain. Thus *Clostridium septicum* α-toxin shares 27% sequence identity with the C-terminal domains of aerolysin, but does not have an APT domain (Ballard *et al.*, 1995) as we pointed out earlier. The PT subunits S4 and S5 are structurally related to the C-terminal domains of S2/3 but they too lack the APT domain. Structurally related subunits without an APT domain are also found in the cholera and Shiga toxin families where they form homopentamers (Sixma *et al.*, 1991; Stein *et al.*, 1992; Fraser *et al.*, 1994; Merritt *et al.*, 1994; Zhang *et al.*, 1995). These observations are most consistent with a model in which a common ancestor of the B-pentamers in these toxins was a single domain subunit forming a homopentameric structure. In PT this subunit was initially duplicated to three distinct genes. In view of the extremely low sequence identity between S4, S5 and the C-terminal domain of S2/3, this must have occurred very long ago. The APT domain was then fused to the ancestor of the S2 and S3 subunits which, based on the 70% sequence identity between them, originated from a much more recent gene duplication.

Evolutionary origin of the APT domain

In the common fold, there is no significant pairwise sequence identity between the APT, CTL and Link module families. The only invariant sequence feature is a disulfide bridge between the first helix in the CTL and Link folds, or its topological equivalent in the APT fold, and the C-terminal β-strand (Figure 2B). This disulfide is very probably a crucial determinant of this common fold. However, in spite of the lack of sequence similarity, it seems likely that the folds in the three families have diverged from a common ancestor. Indeed, each family is rather divergent within itself, with the sequence similarity between distant members being very modest and due mainly to functionally conserved residues. The conservation of a common fold and a common location of the receptor-binding site is most easily explained by a common ancestry. The rather complex architecture and topology of the APT fold, which is so distinct from simple folds seen in most α+β proteins, make such a chance similarity unlikely. A similar argument has been made for the divergent relationship of the catalytic domains of the

prokaryotic ADP-ribosylating toxins and the functionally related eukaryotic poly(ADP-ribose) polymerase. These proteins share a complex $\alpha+\beta$ fold in the absence of significant sequence similarity (Ruf *et al.*, 1996).

If the APT, CTL and Link families are evolutionarily related, then what was their common ancestor? The prokaryotic APT family has no obvious eukaryotic relatives, whereas so far both CTL and Link modules have been identified or predicted only in metazoan proteins. No plant, fungal or protozoan proteins are known or predicted to have the CTL-like fold. The CTL and Link families appear to be large and diverse (Drickamer, 1988; Brissett and Perkins, 1996; Kohda *et al.*, 1996), suggesting that they appeared during an early stage of animal evolution and that in animals there may well be other CTL-like families. In contrast, we could identify only three members of the APT family in spite of the fact that the sequence database contains several completed or nearly completed bacterial genomes. These observations suggest an animal origin for the APT domain. The bacteria could have acquired the hypothetical ancestral APT domain from an animal host protein that is yet to be discovered. By analogy with the CTL and Link functions, this APT domain ancestor would most likely have recognized cell surface carbohydrates. Incorporation of receptor binding into their toxins would have given the bacteria an advantage by increasing the sensitivity of the host cells.

Conclusions

We have identified a common structural domain in aerolysin and the PT S2 and S3 subunits. A similar, though more divergent, domain also exists in CTLs and Link module. All these proteins are known or expected to bind carbohydrates. Although the details of the proposed binding sites differ significantly, they all occupy the same area of the protein surface. The conservation of structure as well as function most likely results from a divergent evolution. Since binding to host cells is the initial step in toxic activity, the localization of binding sites in aerolysin and the PT S2 and S3 subunits will offer one more target in the design of drugs and vaccines directed against these toxins.

Materials and methods

Materials

Purified VSG was generously provided by Dr T.W.Pearson, University of Victoria. Rabbit polyclonal anti-aerolysin was produced in one of our laboratories (J.T.B.). Goat anti-rabbit secondary antibody was obtained from Amersham and *p*-nitrophenylphosphate was purchased from Sigma (Sigma Chemical Co., St Louis, MO). Outdated human blood was supplied by the Canadian Red Cross.

Site-directed mutagenesis

Site-directed mutagenesis was performed and sequences were confirmed using procedures we have described previously (Green and Buckley, 1990; Hardie *et al.*, 1995). Wild-type and all the mutant proteins were purified by hydroxyapatite and ion exchange chromatography as described by Buckley (1990).

Binding assays

All of the steps were carried out at 37°C. VSG was dissolved in water and serially diluted in the wells of a microtitre plate. The plate was dried overnight and the wells were blocked by adding 200 µl of 3% bovine serum albumin (BSA) in 10 mM phosphate, 0.15 M NaCl, pH 7.4 (PBS). The albumin solution was removed after 30 min by washing three times with PBS containing 0.05% (w/v) Tween, then 100 µl of 20 nM proaerolysin in the same buffer containing 1% BSA was added to each well. After a further 30 min, the proaerolysin solution was removed, the wells were washed three times with the PBS, Tween, BSA buffer, and 100 µl of rabbit anti-aerolysin antiserum (diluted 1 to 4000 in the same buffer) was added. After 30 min and three washes with the buffer, this was followed by a 30 min incubation with 100 µl of goat anti-rabbit alkaline phosphatase secondary antibody (2 µg/ml). Finally, after three more washes, 200 µl of 1 mg/ml *p*-nitrophenylphosphate in 100 mM diethanolamine, 0.5 mM $MgCl_2$, pH 9.8 was added, and the plates were read at 405 nm after 1 h. The results are from one of at least two separate experiments.

Haemolysis assay

The activities of the aerolysin mutants were compared with wild-type as we have described previously (Howard and Buckley, 1982; Hardie *et al.*, 1995). Briefly, 100 µl of each purified proaerolysin was activated by treatment with 2 µg/ml trypsin and then 2-fold serially diluted in the wells of a microtitre plate. An equal volume of 0.8% (v/v) washed human erythrocytes was added to each well and the plates were incubated at 37°C for 1 h. We have shown recently that the human erythrocyte contains a previously undetected homologue of the rat erythrocyte receptor (Nelson *et al.*, in preparation). The well containing the highest dilution of each sample that caused 100% haemolysis was recorded.

Acknowledgements

We are grateful to Dr A.J.Day for providing us with the coordinates of Link module prior to their public release. We thank E.Pleshak for her skilled techinal assistance and Dr T.W.Pearson for providing us with purified VSG. J.T.B. acknowledges support from the National Science and Engineering Council of Canada and the British Columbia Health Sciences Research Foundation. R.J.R. is a Scholar of the Alberta Heritage Foundation for Medical Research and acknowledges support from the Protein Engineering Network of Centres of Excellence and Connaught Laboratories. B.H. acknowledges financial support from the Medical Research Council of Canada. A.G.M. is an MRC Senior Research Fellow. M.W.P. is a Wellcome Australian Senior Research Fellow and acknowledges support from the Australian National Health and Medical Research Council.

References

Altwegg,M. and Geiss,H.K. (1989) *Aeromonas* as a human pathogen. *CRC Crit. Rev. Microbiol.*, **16**, 253–286.

Armstrong,G.D., Howard,L.A. and Peppler,M.S. (1988) Use of glycosyltransferases to restore pertussis toxin receptor activity to asialogalactofetuin. *J. Biol. Chem.*, **263**, 8677–8684.

Bairoch,A., Bucher,P. and Hofmann,K. (1996) The PROSITE database, its status in 1995. *Nucleic Acid Res.*, **24**, 189–196.

Ballard,J., Crabtree,J., Roe,B.A. and Tweten,R.K. (1995) Activation and mechanism of *Clostridium septicum* alpha toxin. *Infect. Immun.*, **63**, 340–344.

Barton,G.J. (1993) ALSCRIPT—a tool to format multiple sequence alignments. *Protein Eng.*, **6**, 37–40.

Bertrand,J.A., Pignol,D., Bernard,J.-P., Verdier,J.-M., Dagorn,J.-C. and Fontecilla-Camps,J.C. (1996) Crystal structure of human lithostathine, the pancreatic inhibitor of stone formation. *EMBO J.*, **15**, 2678–2684.

Brennan,M.J., David,J.L., Kenimer,J.G. and Manclark,C.R. (1988) Lectin-like binding of pertussis toxin to a 165-kilodalton Chinese hamster ovary cell glycoprotein. *J. Biol. Chem.*, **263**, 4895–4899.

Brissett,N.C. and Perkins,S.J. (1996) The protein fold of the hyaluronate-binding proteoglycan tandem repeat domain of the link protein, aggrecan and CD44 is similar to that of the C-type lectin superfamily. *FEBS Lett.*, **388**, 211–216.

Buckley,J.T. (1990) Purification of cloned proaerolysin released by a low protease mutant of *Aeromonas salmonicida*. *Biochem. Cell Biol.*, **68**, 221–224.

Drickamer,K. (1988) Carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.*, **263**, 9557–9560.

Ferrin,T.E., Huang,C.C., Jarvis,L.E. and Langridge,R. (1988) The MIDAS display system. *J. Mol. Graphics*, **6**, 13–27.

Fraser,M.E., Chernaia,M.M., Kozlov,Y.V. and James,M.N.G. (1994) Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5 Å resolution. *Nature Struct. Biol.*, **1**, 59–64.

Gill,D.M. (1978) Seven toxic peptides that cross cell membranes. In Jeljaszewica,J. and Wadstrom,T. (eds), *Bacterial Toxins and Cell Membranes*. Academic Press, New York, pp. 291–332.

Graves,B.J. *et al*. (1994) Insight into E-selectin ligand interaction from the crystal structure and mutagenesis of the lec EGF domains. *Nature*, **367**, 532–538.

Green,M.J. and Buckley,J.T. (1990) Site-directed mutagenesis of the hole-forming toxin aerolysin: studies on the role of histidines in receptor binding and oligomerization of the monomer. *Biochemistry*, **29**, 2177–2180.

Gruber,H.J., Wilmsen,H.-U., Cowell,S., Schindler,H. and Buckley,J.T. (1994) Partial purification of the rat erythrocyte receptor for the channel-forming toxin aerolysin and reconstitution into planar lipid bilayers. *Mol. Microbiol.*, **14**, 1093–1011.

Hardie,K.R., Schulze,A., Parker,M.W. and Buckley,J.T. (1995) *Vibrio* spp. secrete proaerolysin as a folded dimer without the need for disulphide bond formation. *Mol. Microbiol.*, **17**, 1035–1044.

Hausman,S.Z. and Burns,D.L. (1993) Binding of pertussis toxin to lipid vesicles containing glycolipids. *Infect. Immun.*, **61**, 335–337.

Heerze,L.D., Smith,R.H., Wang,N. and Armstrong,G.D. (1995) Utilization of sialic acid-binding synthetic peptide sequences derived from pertussis toxin as novel anti-inflammatory agents. *Glycobiology*, **5**, 427–433.

Holm,L. and Sander,C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.

Howard,S.P. and Buckley,J.T. (1982) Membrane glycoprotein receptor and hole-forming properties of a cytolytic protein. *Biochemistry*, **21**, 1662–1667.

Jones,T.A., Zhou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.

Kraulis,J.P. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Kohda,D., Morton,C.J., Parkar,A.A., Hatanaka,H., Inagaki,F.M., Campbell,I.D. and Day,A.J. (1996) Solution structure of the Link module: a hyaluronan-binding domain involved in extracellular matrix stability and cell migration. *Cell*, **86**, 767–775.

Loosmore,S., Zealey,G., Cockle,S., Boux,H., Chong,P., Yacoob,R. and Klein,M. (1993) Characterization of pertussis toxin analogs containing mutations in B-oligomer subunits. *Infect. Immun.*, **61**, 2316–2324.

Merritt,E.A., Sarfaty,S., van den Akker,F., L'hoir,C., Martial,J.A. and Hol,W.G.J. (1994) Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide. *Protein Sci.*, **3**, 166–175.

Montecucco,C., Tomasi,M., Schiavo,G. and Rappuoli,R. (1986) Hydrophobic photolabelling of pertussis toxin subunits interacting with lipids. *FEBS Lett.*, **194**, 301–304.

Murzin,A.G, Brenner,S.E., Hubbard,T. and Chothia,C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.*, **247**, 536–540.

Parker,M.W., ed. (1996) *Protein Toxin Structure*. R.G.Landes Company, Austin, TX.

Parker,M.W., Buckley,J.T., Postma,J.P.M., Tucker,A.D., Leonard,K., Pattus,F. and Tsernoglou,D. (1994) Structure of the *Aeromonas* toxin proaerolysin in its water-soluble and membrane-channel states. *Nature*, **367**, 292–295.

Parker,M.W., van der Goot,F.G. and Buckley,J.T. (1996) Aerolysin—the ins and outs of a model channel-forming toxin. *Mol. Microbiol.*, **19**, 205–212.

Pittman,M. (1984) The concept of pertussis as a toxin-mediated disease. *Pediatr. Infect. Dis.*, **3**, 467–486.

Rozdzinski,E., Jones,T., Burnette,W.N., Burroughs,M. and Tuomanen,E. (1993) Antiinflammatory effects in experimental meningitis of prokaryotic peptides that mimic selectins. *J. Infect. Dis.*, **168**, 1422–1428.

Ruf,A., de Murcia,J.M., de Murcia,G.M. and Schulz,G.E. (1996) Structure of the catalytic fragment of poly (ADP-ribose) polymerase from chicken. *Proc. Natl Acad. Sci. USA*, **93**, 7481–7485.

Saukkonen,K., Burnette,W.N., Mar,V.L., Masure,H.R. and Tuomanen,E.I. (1992) Pertussis toxin has eukaryotic-like carbohydrate recognition domains. *Proc. Natl Acad. Sci. USA*, **89**, 118–122.

Sixma,T.K., Pronk,S.E., Kalk,K.H., Wartna,E.S., van Zanten,B.A.M., Witholt,B. and Hol,W.G.J. (1991) Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E.coli*. *Nature*, **351**, 371–377.

Stein,P.E., Boodhoo,A., Tyrrell,G.J., Brunton,J.L. and Read,R.J. (1992) Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E.coli*. *Nature*, **355**, 748–750.

Stein,P.E., Boodhoo,A., Armstrong,G.D., Cockle,S.A., Klein,M.H. and Read,R.J. (1994a) The crystal structure of pertussis toxin. *Structure*, **2**, 45–57.

Stein,P.E., Boodhoo,A., Armstrong,G.D., Heerze,L.D., Cockle,S.A., Klein,M.H. and Read,R.J. (1994b) Structure of a pertussis toxin–sugar complex as a model for receptor binding. *Nature Struct. Biol.*, **1**, 591–596.

Tallet,A., Seabrook,R.N., Irons,L.I., Robinson,A., van Heyningen,S. and Atkinson,T. (1993) Localisation of a receptor-binding domain on the S3 subunit of pertussis toxin by peptide mapping. *Eur. J. Biochem.*, **211**, 743–748.

Vyas,N.K. (1991) Atomic features of protein–carbohydrate interactions. *Curr. Opin. Struct. Biol.*, **1**, 732–740.

Weis,W.I., Kahn,R., Fourme,R., Drickamer,K. and Hendrickson,W.A. (1991) Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science*, **254**, 1608–1615.

Weis,W.I., Drickamer,K. and Hendrickson,W.A. (1992) Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature*, **360**, 127–134.

Witvliet,M.H., Burns,D.L., Brennan,M.J., Poolman,J.T. and Manclark,C.R. (1989) Binding of pertussis toxin to eukaryotic cells and glycoproteins. *Infect. Immun.*, **57**, 3324–3330.

Zhang,R.-G., Scott,D.L., Westbrook,M.L., Nance,S., Spangler,B.D., Shipley,G.G. and Westbrook,E.M. (1995) The three-dimensional structure of cholera toxin. *J. Mol. Biol.*, **251**, 563–573.

Received on December 10, 1996; revised on February 7, 1997