doi:10.1006/jmbi.2000.3801 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2000) **299**, 1005–1017

JMB

Crystal Structure of *Pseudomonas aeruginosa* PAK Pilin Suggests a Main-chain-dominated Mode of Receptor Binding

Bart Hazes[1]*, Parimi A. Sastry[1], Koto Hayakawa[1], Randy J. Read[2] and Randall T. Irvin[1]

[1]*Department of Medical Microbiology and Immunology Canadian Bacterial Diseases Network, University of Alberta Edmonton, Alberta, T6G 2H7 Canada*

[2]*Division of Structural Medicine, Department of Haematology, University of Cambridge, Cambridge Institute for Medical Research Cambridge, CB2 2XY, UK*

**Corresponding author*

Fibers of pilin monomers (pili) form the dominant adhesin of *Pseudomonas aeruginosa*, and they play an important role in infections by this opportunistic bacterial pathogen. Blocking adhesion is therefore a target for vaccine development. The receptor-binding site is located in a C-terminal disulphide-bonded loop of each pilin monomer, but functional binding sites are displayed only at the tip of the pilus. A factor complicating vaccination is that different bacterial strains produce distinct, and sometimes highly divergent, pilin variants. It is surprising that all strains still appear to bind a common receptor, asialo-GM1. Here, we present the 1.63 Å crystal structure of pilin from *P. aeruginosa* strain PAK. The structure shows that the proposed receptor-binding site is formed by two β-turns that create a surface dominated by main-chain atoms. Receptor specificity could therefore be maintained, whilst allowing side-chain variation, if the main-chain conformation is conserved. The location of the binding site relative to the proposed packing of the pilus fiber raises new issues and suggests that the current fiber model may have to be reconsidered. Finally, the structure of the C-terminal disulphide-bonded loop will provide the template for the structure-based design of a consensus sequence vaccine.

Keywords: type IV pilin; *Pseudomonas aeruginosa*; crystal structure; carbohydrate binding; vaccine design

Introduction

Pseudomonas aeruginosa is an opportunistic pathogen that infects immunosuppressed individuals and compromised tissues such as burn wounds (Bang *et al*., 1998), damaged corneal tissue (Dart & Seal, 1988), and the trachea of intubated patients (Crouch Brewer *et al*., 1996) or patients with cystic fibrosis (Tummler *et al*., 1997). Treatment of patients is complicated by an innate resistance to antimicrobial agents (Alonso *et al*., 1999) and the presence of an alginate capsule (May *et al*., 1991). As a consequence, *P. aeruginosa* infections contribute significantly to morbidity and mortality in the groups at risk.

For *P. aeruginosa* and many other pathogenic organisms the ability to adhere to host tissues is essential to initiate an infection (Beachey, 1981). Adhesins are therefore a common constituent of the repertoire of virulence factors of pathogens. In *P. aeruginosa* several adhesins have been described; viz. the alginate capsule (Doig *et al*., 1987), exoenzyme S (Baker *et al*., 1991), flagella (Ramphal *et al*., 1996), and pili (Doig *et al*., 1988). For *P. aeruginosa*, pili have been found to be the dominant adhesins early in the infection process (for a review, see Hahn, 1997). The importance of pili-mediated adherence is further demonstrated by the loss of virulence of engineered *P. aeruginosa* strains that lack functional pili (Farinha *et al*., 1994; Tang *et al*., 1995) and by the protective effect of immunisation with purified pilins (Sheth *et al*., 1995; Cachia *et al*., 1998). Although immunisation was successful in animal models, the protection was found to be strain-specific. Determination of the amino acid sequences of pilins of several *P. aeruginosa* strains

This work is dedicated to Parimi Sastry, who passed away shortly after the PAK pilin structure was determined.

E-mail address of the corresponding author: bart.hazes@ualberta.ca

0022-2836/00/041005–13 $35.00/0

has indeed revealed a striking level of sequence divergence (Figure 1).

The pili of *P. aeruginosa* are polymers of non-covalently associated pilin monomers, forming a fiber that is approximately 6 nm wide and up to several micrometres long. They are polar, retractable, and are considered as prototypes of type IV pili. An assembly model of the type IV pilus has recently been proposed based on the crystal structure of *Neisseria gonorrhoeae* strain MS11 pilin (MS11 pilin) and other experimental data (Parge *et al*., 1995; Forest & Tainer, 1997). All type IV pili share a highly conserved N-terminal hydrophobic sequence (Figure 1) which is believed to be central to the assembly and integrity of the pilus fiber. The general features of the proposed fiber model for MS11 pili would be accordingly expected to be applicable to all type IV pili. Proteins with a very similar N-terminal hydrophobic α-helix also play a role in the general secretion pathway which is required for assembly of the type IV pilus and the bacterial secretion of numerous enzymes and toxins (Lu *et al*., 1997).

The adhesive properties of type IV pili are located at the tip of the pilus. In *P. aeruginosa* each monomer (named pilA) has a receptor binding site. However, only the binding sites of monomers at the pilus tip are functional, with the others apparently obscured during fiber assembly (Lee *et al*., 1994). In *N. gonorrhoeae* pili the monomer (pilE) forms a fiber that can bind red blood cells; however, binding to epithelial cells requires pilC (Scheuerpflug *et al*., 1999), a distinct and possibly tip-located protein (Rudel *et al*., 1995). The gene for a pilC homologue exists in *P. aeruginosa* (pilY1) (Alm *et al*., 1996), but its function is as yet unclear and indeed its presence in the pilus is uncertain.

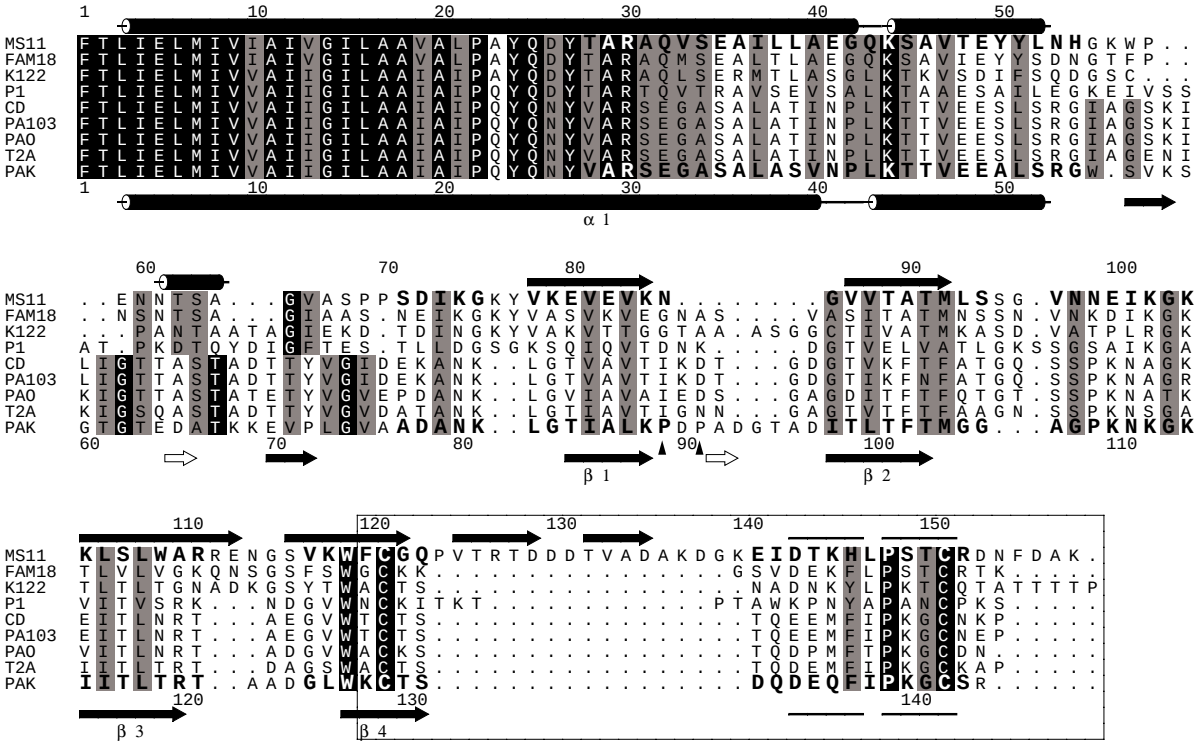

Figure 1. Multiple sequence alignment of type IV pilins. The observed secondary structural elements, as given by PROCHECK (Laskowski, 1993), for MS11 pilin and PAK pilin are given above and below each sequence block, respectively, using cylinders for α-helices, filled arrows for β-strands, and horizontal lines for the C-terminal β-turns. Residue numbers for these two sequences are also provided. Regions that are structurally equivalent in MS11 and PAK pilin are shown in bold and slightly larger font. Positions in these regions, or in the N-terminal hydrophobic helix, that are strictly conserved are printed in reverse video. Positions with moderate conservation are printed on a gray background. For positions 55 to 77 (PAK numbering), strict and moderate conservation is calculated separately for the top four and bottom five sequences, respectively. This reflects our belief that these two groups of sequences adopt distinct conformations in this region. Two very short β-strands in the PAK structure are indicated by open arrows, and two filled triangles highlight the positions of the two *cis*-proline residues. The C-terminal disulfide-bonded loop, which contains the proposed binding site and is the focus of peptide-based vaccine design, is boxed. The listed sequences are from; *N. gonorrhoeae* strain MS11 (gi|3212472), *N. meningitidis* strain FAM18 (gi|2228578), and *P. aeruginosa* strains K122-4 (gi|77636), P1 (gi|151472), CD (gi|120405), PA103 (gi|120436), PAO (gi|120440), T2A (gi|545132), and PAK (gi|120438). The alignment was prepared with CLUSTALW (Thompson *et al*., 1994) with minor manual editing, and the Figure was created with the program ALSCRIPT (Barton, 1993).

P. aeruginosa pilin has been found to bind glycoconjugates which contain a β-D-GalNAc(1 → 4)-β-D-Gal moiety. Asialo-GM1 and asialo-GM2 glycolipids are believed to be important physiological receptors (Saiman & Prince, 1993; Gupta *et al*., 1994; de Bentzmann *et al*., 1996; Comolli *et al*., 1999), but binding to glycoproteins has also been described (Doig *et al*., 1989; Rudner *et al*., 1992; Hazlett *et al*., 1995; Wu *et al*., 1995). A variety of techniques including antibody mapping (Doig *et al*., 1990), mutagenesis (Farinha *et al*., 1994), studies with synthetic peptides (Doig *et al*., 1988; Irvin *et al*., 1989; Lee *et al*., 1989), and 2D-NMR (Campbell *et al*., 1997a) all indicate that the receptor-binding site is located in the C-terminal disulfide-bonded loop (boxed in Figure 1). Surprisingly, there is significant sequence variation in this region between different strains (Figure 1), even though all strains are believed to still bind the same receptors (Ramphal *et al*., 1984). As a consequence, immunisation with purified pili does not generally protect against subsequent challenge with strains other than that used for immunization.

The observation that strains have conserved binding specificity suggests that they still share common features in their receptor-binding site. This raises the possibility that a strain-independent therapeutic can be found in the form of a receptor mimetic or by raising a cross-reactive immune response. Indeed, cross-reactive monoclonal antibodies have been reported (Lee *et al*., 1990; Sheth *et al*., 1995) and a peptide-based vaccine that optimizes the generation of a cross-reactive immune response is in development (Wong *et al*., 1994; Cachia *et al*., 1998). Here, we present the 1.63 Å crystal structure of *P. aeruginosa* strain PAK (PAK pilin) and discuss its features and its use towards designing an improved peptide-based vaccine.

Results and Discussion

Structure determination and structure quality

Crystallization of full-length *P. aeruginosa* pilin has proved to be extremely difficult due to poor solubility and reproducibility (N.E.C. Duke and R.J.R., unpublished results). Based on the MS11 pilin structure it became clear that this was due to the fully exposed and highly hydrophobic N-terminal α-helix (Figures 1 and 2). Since receptor binding is mediated by the C-terminal domain, we have prepared a truncated recombinant pilin molecule encoding residues 29 to 144 of the mature PAK pilin (see Materials and Methods). Extensive functional studies on a similarly truncated strain K122-4 pilin show that the truncated product retains its binding properties (Irvin *et al*., unpublished results). These results, the lack of interactions between the N-terminal α-helix and the binding site, and preliminary binding data for truncated PAK pilin (not shown), all indicate that truncated PAK pilin also retains a functional receptor-binding site. The truncated PAK pilin is highly soluble and yielded excellent crystals. The resolution of the 1.63 Å data set used in this study is actually limited by the detector hardware rather than by the diffracting power of the crystal. A data set to better than 1.0 Å resolution has recently been collected at the SBC beamline of the APS synchrotron (Chicago, USA) which should allow us to obtain an atomic resolution model of PAK pilin in the future.

At this moment, the only type IV pilin of known structure is the 2.6 Å crystal structure of MS11 pilin (Parge *et al*., 1995; Forest *et al*., 1999). Residues 29 to 144 of PAK pilin, present in our truncated construct, have only 17.2 % sequence identity with MS11 pilin, and no convincing molecular replacement solution could be found using the AMoRe program (Navaza, 1994). Experimental phases were therefore obtained using the multiple isomorphous replacement technique (see Materials and Methods and Table 1). The quality of the starting phases, combined with the high resolution and completeness of the native data (only the F_{000} reflection is missing out to 1.64 Å), allowed wARP (Lamzin & Wilson, 1997) to auto-trace 116 out of the 120 residues (97 %) that are visible in the final model. The r.m.s.d. from the final refined model was 0.10 Å for all main-chain atoms including C^{β}. The three N-terminal residues, not visible in the final model, and the fourth residue (residue 25) were not modeled at all. The fifth residue (residue 26) was present in wARP's dummy atom model, but was not auto-traced. Residues Pro89 and Asp90 were accurately modeled by dummy atoms, but were not auto-traced because Pro89 and Pro91 are *cis*-proline residues (Figure 3 and below).

The final refined model has R and R_{free} values of 15.3 % and 18.1 %, respectively. PROCHECK analysis (Laskowski *et al*., 1993) shows that 96 % of all residues are in the most favoured regions of the Ramachandran plot, and the remaining 4 % in the additionally allowed regions. The r.m.s.d. of bond lengths and angles are 0.012 Å and 1.0°, respectively, and the r.m.s.d. for *B*-factors of bonded main-chain and side-chain atoms are 2.6 and 4.4 Å^2, respectively. As may be expected, the electron density is excellent for most of the model as is clear from Figure 3. Only the first three residues, all from the OmpA fusion, could not be built due to lack of electron density. Several residues (Arg30, Lys68, Lys88, Lys110, Lys128, and Glu135), all solvent-exposed, did not show appreciable electron density for part of their side-chains and the atoms without clear experimental density have not been included in the model. Residues Glu49, Gln136 and Arg144 had very weak side-chain density, and for these residues the occupancy of the side-chain atoms was set to 0.5. Alternative side-chain conformations could be modeled for several residues (Ser31, Ser52, Val57, Thr84, Lys88, Thr99, Met104, and Lys112). The final structure of PAK pilin contains 864 non-hydrogen protein atoms in 120 residues and 131 water molecules.

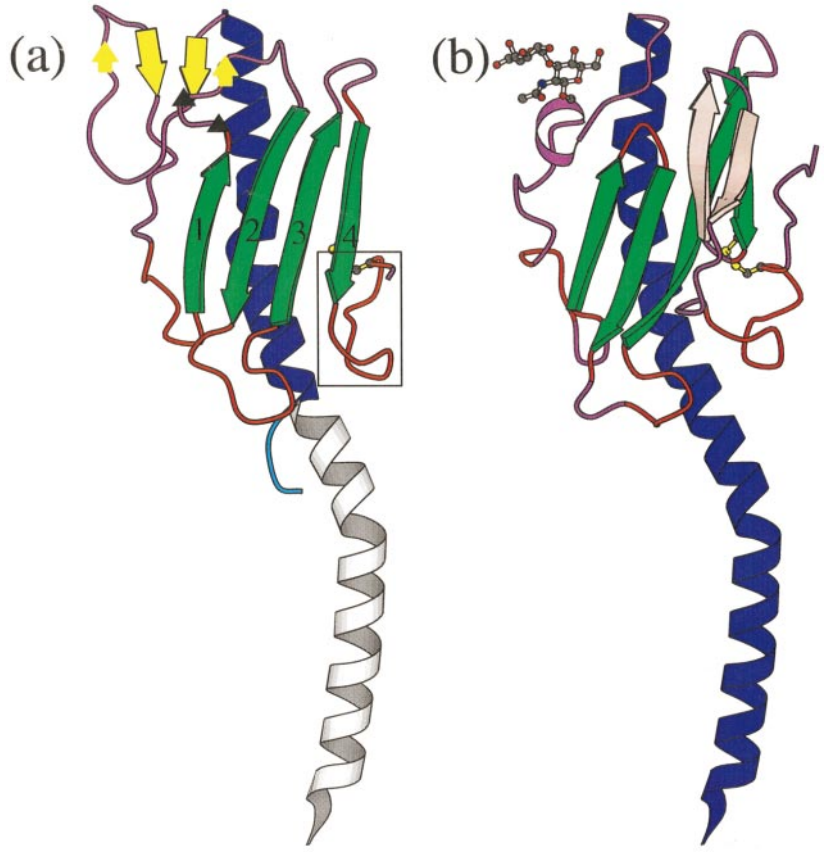

Figure 2. Diagrams of PAK and MS11 pilins. Structurally conserved residues are colored green in β-strands, blue in α-helices, and red in coil regions. Other coil regions are shown in purple. (a) N-terminally truncated PAK pilin. The truncated hydrophobic N-terminal helix (gray) was modeled on the basis of the MS11 model. N-terminal residues derived from the OmpA fusion construct are colored cyan. The minor β-sheet is shown in yellow. The outside strands are drawn as smaller arrows to indicate the fact that only a single residue is involved in proper β-sheet hydrogen bonding. Positions of the *cis*-proline residues are shown by filled triangles. The C-terminal disulfide-bonded loop, which contains the proposed binding site and is the focus of peptide-based vaccine design is boxed and the disulfide bond is shown as ball-and-stick model with yellow bonds. The boxed area is shown in more detail in Figure 4(a). (b) MS11 pilin (Parge *et al.*, 1995). The color scheme is as above with the addition of using pink for the C-terminal β-hairpin. The glycosylation at position 63 is shown as a ball-and-stick model. This Figure, as well as Figures 4 and 5, were created with the program MOLSCRIPT (Kraulis, 1991).

Structure description

Figure 2 shows a representation of (a) the N-terminally truncated PAK pilin structure, next to (b) full-length MS11 pilin. These structures share a basic fold consisting of a four-stranded antiparallel β-sheet (the major β-sheet, green) packed onto an α-helix (blue). In total 83 C^{α} atoms of MS11 pilin could be superimposed on the PAK pilin structure, giving an r.m.s.d. of 1.9 Å. These C^{α} atoms were distributed over seven segments (printed in bold font in Figure 1). Although MS11 and PAK pilin are clearly related, significant differences exist as would be expected from the high sequence divergence. A major difference occurs immediately after the α-helix. In MS11 pilin, the α-helix is followed by an irregular loop that is glycosylated at Ser63 (Figure 2(b)) and phosphorylated at Ser68 (Parge *et al.*, 1995; Forest *et al.*, 1999). In PAK pilin, this region forms three short irregular β-strands (the minor β-sheet, yellow in Figure 2(a)). No post-translational modification was observed, as expected from the fact that the protein was expressed in *Escherichia coli*. However, even in its natural host no MS11-like post-translational modification is believed to occur in PAK pilin, as Ser63 and Ser68 are replaced by Ala and Gly, respectively. MS11 pilin can also have α-glycerophosphate attached to residue Ser94 (Stimson *et al.*, 1996; Forest *et al.*, 1999), but again the corresponding residue has been replaced by glycine in PAK pilin. Furthermore, biochemical data have indicated that PAK and PAO pilin are not glycosylated or phosphorylated (Paranchych *et al.*, 1979). Therefore our structure should reflect the mature protein except for the N-terminal truncation.

It has been noted previously that *P. aeruginosa* pilin sequences for strains CD, PA103, PAO, T2A, and PAK are more closely related to each other than to the pilins of strains K122-4 and P1 (Pasloske *et al.*, 1988; Castric & Deal, 1994). When the sequence data are combined with the new structural knowledge it appears that the minor β-sheet of PAK pilin is also present in strains CD, PA103, PAO, and T2A. In contrast, the corresponding loop structure of MS11 pilin appears to be shared by the pilins of the *N. meningitidis* FAM18 strain and the *P. aeruginosa* K122-4 and P1 strains (Figure 1). The closer structural relationship of some *P. aeruginosa* pilin strains with gonococcal homologues is further supported by the fact that K122-4 pilin is more closely related to the *N. meningitidis* FAM18 pilin (45.6 % identity for the globular domain) than to the pilins of the other *P. aeruginosa* strains (best match is 23.3 % identity with strain PAK). The division of pilin sequences

into two groups may also be relevant to glycosylation as P1 pilin, like the *Neisseria* pilins, has been reported to be glycosylated (Castric, 1995). It is interesting that a gene required for glycosylation was detected just downstream from the pilin gene in strain P1 and its close homologues, but not in the same location in the PAK-like pilin strains (Castric, 1995). Glycosylation of P1 pilin could take place at the threonine residue that corresponds to Ser63 in MS11 (Figure 1). Threonine is also conserved in K122-4 pilin but the presence of glycosylation or a pilO-related gene has not been studied for this strain.

A second large structural deviation between MS11 and PAK pilin is located near the C terminus of the molecule where a two-stranded β-hairpin in MS11 pilin (pink in Figure 2(b)) is missing from the PAK pilin structure. The β-hairpin in MS11 pilin corresponds to the hypervariable region and sequence analysis indicates that this β-hairpin is found only in *N. gonorrhoeae* and *N. meningitidis* pilins. It is interesting that the β-hairpin is not present in all *Neisseria* pilin sequences, as *N. meningitidis* strain FAM18 pilin, referred to as a class II pilin (Aho *et al.*, 1997), lacks the β-hairpin (Figure 1). Both the FAM18 and K122-4 pilins appear to have a chimeric nature with an MS11-like irregular loop following the N-terminal α-helix and a PAK-like C-terminal disulfide-linked loop. We anticipate being able to verify this observation in the near future, as crystals of a truncated K122-4 pilin have been obtained that diffract to 1.8 Å (unpublished results).

Both PAK and MS11 pilin have a distortion in their α-helix around residue 42. This position is often, but not always, occupied by either a proline residue (as in PAK pilin) or a glycine residue (as in MS11 pilin). The curve introduced into the helix may be required for proper packing of monomers in the fiber. Another structural feature of interest in PAK pilin is the double *cis*-proline loop which terminates β-strand 1 (Figure 3). The double *cis*-proline loop brings aspartate 90 into a position where it can make favourable hydrogen bonds with the side-chain hydroxyl group and main-chain nitrogen atom of Ser59. In addition, it brings Ala92 into a position where it can form β-sheet-type hydrogen bonds with valine 57, thereby bridging the minor and major β-sheets (shown as a small yellow arrow in Figure 2(a)). These interactions may be important for PAK pilin but, since the proline residues are not conserved in any other pilin, the interactions cannot be critical for function. Two *cis*-proline residues separated by only a single residue are rare but not unique in the structural database. A scan of all PDB entries with the program WHAT IF (Vriend, 1990) revealed five other occurrences in proteins refined to 3.0 Å or higher resolution. These are (PDB codes in parentheses); Fab heavy chain (12E8 and many others), arylsulfatase A (1AUK), methionyl-tRNA fmet formyltransferase (1FMT), Fe-only hydrogenase (1FEH), and N-cadherin (1NCG). No common role for the *cis*-proline residues is apparent from these structures.

Receptor-binding loop

Extensive studies using antibody mapping (Doig *et al.*, 1990), synthetic peptides (Doig *et al.*, 1988;

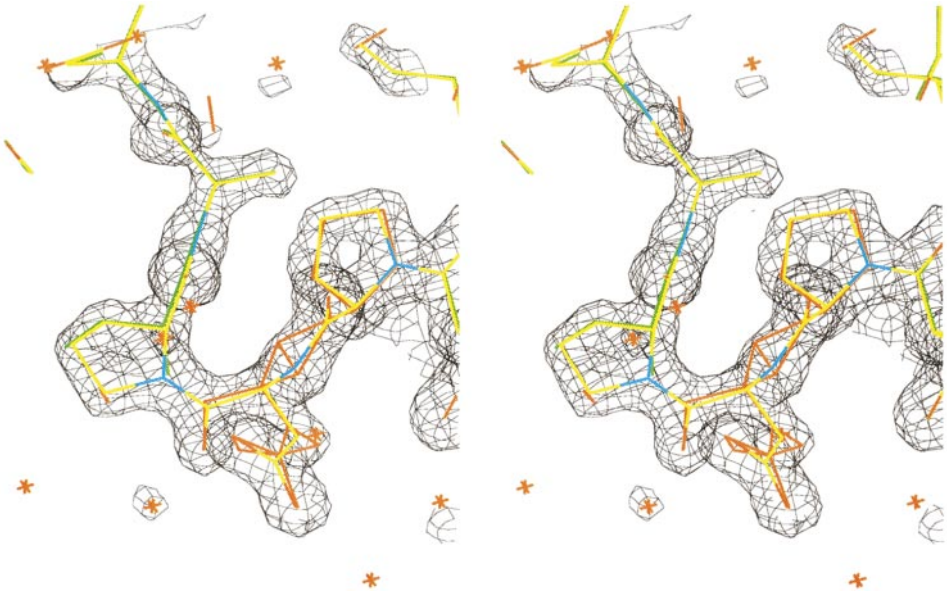

Figure 3. Electron density diagram around the double *cis*-proline loop. The Figure shows a SIGMAA-weighted $2\,mF_o - DF_c$ electron density map, contoured at 2σ, based on the final refined model. The final model itself is included and colored by atom type, carbon (yellow) nitrogen (blue) and oxygen (red) atoms. The auto-traced wARP model is shown with green lines. Red lines connect wARP dummy atoms that were not part of the tracing, whereas red crosses represent non-bonded dummy atoms. This Figure was created with Xfit (McRee, 1999).

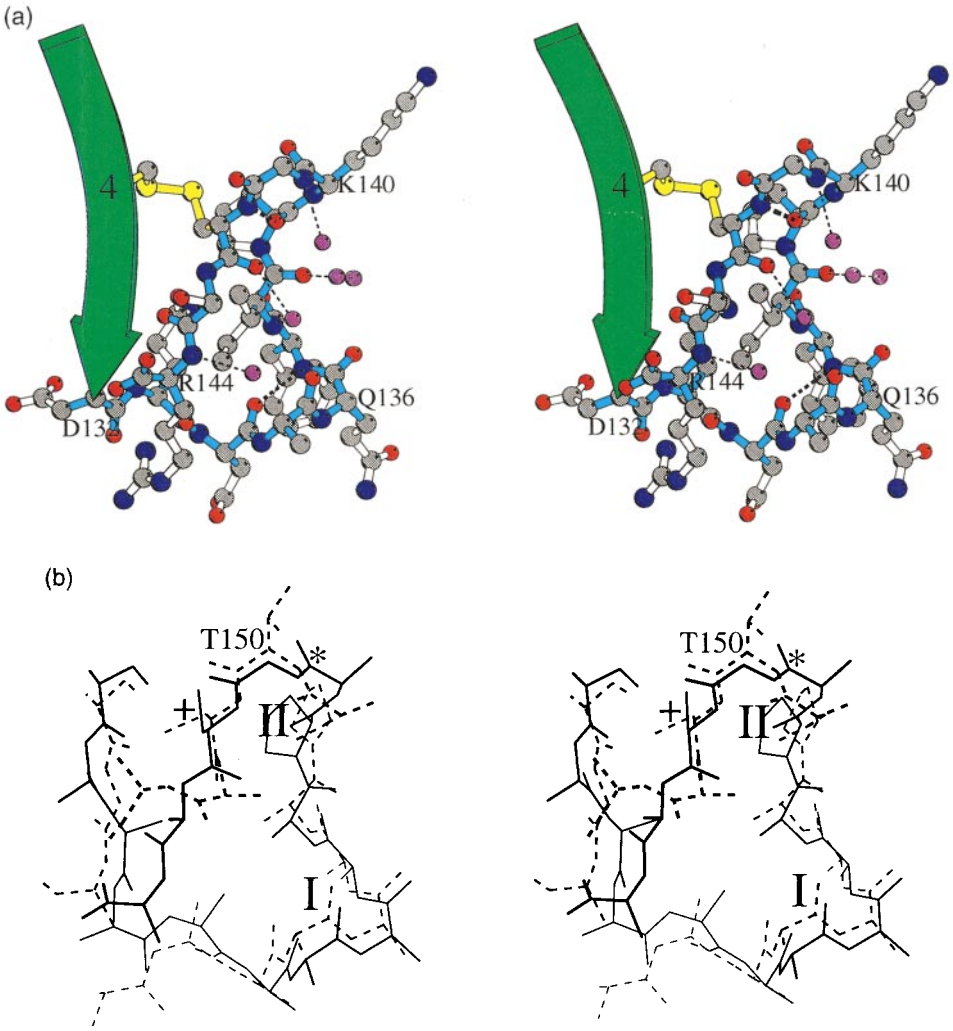

Figure 4. Stereo images of the receptor-binding loop. (a) Enlarged view of the boxed area of Figure 2(a) with residues 132 to 144 shown in ball-and-stick representation. Bonds that connect main-chain atoms are shown in cyan to highlight the main-chain conformation. Every fourth residue, starting at residue 132, has been numbered. The hydrogen bonds between residues i and $i+3$ of the β-turns are shown as bold broken lines. Note that residues 135 to 144 form a surface that is dominated by main-chain atoms. Five water molecules located in the proposed binding pocket (purple spheres) that are connected to their hydrogen-bonding partners are shown with broken lines. Glu135 has a disordered side-chain and the side-chain atoms beyond C^{β} are therefore not depicted. (b) Superposition of the receptor-binding loop of PAK pilin (continuous line) and the corresponding region in MS11 pilin (broken line). The main-chain atoms, including C^{β}, are shown plus the side-chains of the strictly conserved cysteine and proline residues, and Thr150 of MS11 pilin. Residues 124 to 140 of MS11 pilin, which form the extra β-hairpin, are not shown. The first β-turn (labeled I) is a type I turn in both structures. However, the second β-turn (labeled II) is a type II turn in PAK and a type I turn in MS11. This difference results from a peptide flip (labeled *), most likely due to Thr150 in MS11 pilin (see the text). Another difference of interest is the change in the chi-1 dihedral angle of one of the strictly conserved cysteine residues (labeled +).

Irvin *et al*., 1989; Lee *et al*., 1989), mutagenesis (Farinha *et al*., 1994), and 2D-NMR (Campbell *et al*., 1997a) all point to the presence of a receptor-binding site in the C-terminal disulfide-bonded loop. This loop is boxed in Figures 1 and 2(a) and is shown in more detail in Figure 4(a). Previous 2D NMR studies on peptides corresponding to the receptor-binding loop in *P. aeruginosa* strains PAK, PAO, KB7, and P1 revealed the conservation of a type I β-turn followed by a type II β-turn (Asp134 to Phe137 and Pro139 to Cys142, respectively in PAK) (McInnes *et al*. 1993, Campbell *et al*., 1997a,b). The PAK crystal structure contains β-turns of the predicted types in the same locations. However, the orientation of the β-turns with respect to each other differs, mostly due to a 120° difference in the psi main-chain dihedral angle of Phe137. NMR data have indicated that Pro139 occurs in both the *cis* and *trans* configuration in a peptide corresponding to residues 128-144 of PAK pilin. Our structure shows that in the presence of the complete globular domain, Pro139 adopts a

trans conformation as had been suggested based on antibody-induced conformational changes of the *cis* but not the *trans* conformer (Campbell *et al*., 1997a).

A detailed view of the disulphide-bonded loop is given in Figure 4(a). This shows that the β-turns are oriented such that their main-chain-dominated faces are solvent-accessible and pointing towards each other. The result is a V-shaped groove lined by main-chain atoms. This pattern is further extended by the main-chain atoms of residues Ile138, Ser143, and the C-terminal Arg144. Cys129, Phe137, Pro 139 and Cys142 are the most conserved amino acid residues in the receptor-binding loop and might therefore have been expected to take part in receptor binding. Interestingly, the structure shows that the side-chains of these residues are not solvent-accessible. Instead, they are located in the interface between the receptor-binding loop and the body of the protein and their main function appears to be to stabilize the conformation of the loop. The picture which emerges is that the protein-receptor interactions may be formed exclusively, or predominantly, by main-chain rather than side-chain atoms. Conservation of the main-chain conformation rather than primary sequence would therefore be critical for function. The NMR observation that peptides corresponding to the disulphide-bonded loop of four distinct *P. aeruginosa* pilin strains all conserve the two β-turns, and that antibody or receptor binding does not induce large conformational changes, fits in this picture of a conserved main-chain conformation.

Wong *et al*. (1995) have used single alanine replacement studies of a synthetic peptide corresponding to residues 128 to 144 of a PAK pilin variant (the variant has lysine as residue 144 rather than arginine as in our structure) to determine the residues important for binding to A549 cells. Residues Ser131, Gln136, Ile138, Pro139, Gly141, and Lys144 were found to be important for binding. Ile138 occupies a central position in the loop (Figure 4(a)) and therefore can be expected to contribute significantly to structural stability. Pro139 is completely conserved and the restricted phi-psi space of proline residues may be relevant to stabilize the structure of the loop. Gly141 has main-chain phi/psi angles of $99^\circ/-14^\circ$. These are unfavourable for non-glycine residues and an alanine replacement could therefore be expected to affect the main-chain structure. Ser131 is oriented such that its side-chain hydroxyl group makes a well-defined hydrogen bond to the main-chain nitrogen atom of Gln133. So again, a replacement by an alanine residue would likely affect the main-chain conformation or stability. Gln136 and Lys144 (Arg144 in our structure) do not appear to affect the main-chain conformation as their side-chains are disordered and do not interact with other protein atoms. These side-chains could, in principle, interact directly with the receptor or the cell membrane, but a more general effect such as peptide solubility cannot be excluded.

Gly141 is conserved in most *P. aeruginosa* pilins, but has been replaced by asparagine and threonine in strains P1 and K122-4, respectively. Asparagine may still be able to adopt phi/psi angles close to $99^\circ/-14^\circ$; however, threonine is much more restricted and it is highly unlikely to adopt such main-chain dihedral angles. Accordingly, we predict a distinctly different main-chain conformation around this residue in the K122-4 strain. It is interesting that the MS11 structure also contains threonine at the corresponding position and a peptide flip is indeed observed in the structure (labeled with * in Figure 4(b)). This changes the type II β-turn in PAK pilin into a type I β-turn in MS11 pilin. In spite of this, the main-chain conformation of the receptor-binding loops in PAK pilin and MS11 pilin remain very similar (Figure 4(b)) and residues 129, 130 and 133 to 142 of PAK pilin can be superimposed on the corresponding atoms in MS11 pilin with an r.m.s.d. of 1.0 Å for 12 C^α atoms. We predict that K122-4 pilin will have two type I β-turns, like MS11 pilin. Interestingly, binding studies demonstrate that truncated K122-4 pilin can still bind asialo-GM1 (R.T.I., unpublished results), indicating that receptor specificity has not been lost.

Carbohydrate-binding sites typically contain an aromatic residue which acts as a hydrophobic stacking surface. This is especially true for the binding of β-galactosyl moieties which have a distinct hydrophobic face (Rini, 1995). Furthermore, carbohydrate-binding sites typically form hydrogen bonds with the sugar hydroxyl groups using side-chain atoms or a mixture of side-chain and main-chain atoms. In our proposed binding cleft there is a notable lack of side-chains to provide either hydrophobic or hydrogen bonding functions. However, Ile138 in the center of the proposed binding surface, and the C^α and C^β atoms of surrounding residues could provide hydrophobic interactions for carbohydrate binding and the main-chain carbonyl oxygen atoms and amide nitrogen atoms could hydrogen bond to the carbohydrate ligand. In the crystal structure, five water molecules are found in the proposed receptor-binding pocket (shown as purple spheres in Figure 4(a)). These water molecules interact with main-chain carbonyl oxygen atoms of residues Ile138 and Cys142, and with the main-chain nitrogen atoms of residues Gly141 and Arg144. These interactions could mimic those of hydroxyl groups of the receptor. We have explored potential receptor-binding modes by molecular docking calculations of a GalNAc-Gal disaccharide with the program package DockVision (Hart *et al*., 1997). Several receptor dockings were obtained, but none was significantly more plausible than the others. It is clear that an experimental structure of a pilin molecule in complex with a receptor analogue is required to firmly establish the location and mode of receptor binding.

Binding site relative to the fiber model

High-affinity binding of lectins to carbohydrates is achieved by multi-valent binding (Rini, 1995). This can be accomplished either by distinct binding sites on a single molecule, by identical binding sites on a multimeric complex, or by a combination of the above. For *P. aeruginosa* pilins there is evidence for only a single binding site per monomer. Therefore, to achieve multi-valency, multiple monomers at the tip of the pilus should contribute to binding. One class of important physiological receptors for *P. aeruginosa* pili are believed to be the glycolipids asialo-GM1 and asialo-GM2 (Saiman & Prince, 1993; Gupta *et al*., 1994; de Bentzmann *et al*., 1996; Comolli *et al*., 1999). Since glycolipids are presented on a flat cell surface, multi-valent binding in this case requires that the binding sites are presented on one face of the protein so as to be able to simultaneously engage the receptors. For pili, the presentation of the binding sites can only be appreciated in the context of the assembled fiber. Figure 5 shows the PAK fiber model based on the assembly parameters described by Parge *et al*. (1995). The model is sterically plausible and its distribution of buried and exposed surfaces fits expectations. However, the binding sites are ill-positioned for their function if the tip of the pilus is located at the C-terminal end of the main α-helix, as is currently assumed (Forest & Tainer, 1997). First, access to the binding site is blocked by the minor β-sheet of a neighbouring subunit in all monomers, including the terminal monomer (Figure 5). Adjustments of the pilus assembly parameters may expose the binding site of the terminal monomer, but the binding sites of the other monomers at the tip of the pilus would remain occluded. Furthermore, even if all binding sites of the terminal five monomers were exposed, they would clearly not be presented on one face of the pilus (Figure 5).

In the current fiber-assembly model, the monomers are related by a 5-fold screw operation along the fiber axis. Since the crystal symmetry includes a 4-fold screw axis, it is conceivable that the crystal packing reflects some aspects of the pilus assembly. However, visual inspection of the crystal packing arrangement suggests that this is not the case. The X-ray fiber diffraction pattern calculated for the proposed model of MS11 pili does not fit the observed fiber diffraction data from type IV pili (Folkhard *et al*., 1981) (J.H. Davies, L.C. Welsh & D.A. Marvin, unpublished results). The true pilus structure could be substantially different. It is even conceivable to have a fiber formed from discs of pentamers rather than a single-start helical arrangement. This would place all binding sites at the tip of the pilus in an identical environment and equidistant from the membrane, not unlike the situation in bacterial toxin B-pentamers (Merritt *et al*., 1994; Ling *et al*., 1998). Although attractive from a functional perspective, the binding sites would still be too far from the tip of the pilus to make binding to glycolipids feasible. Given the current data we cannot conceive of a pilus model that would provide a multi-valent membrane-proximal binding surface unless the monomers at the tip of the pilus undergo a conformational

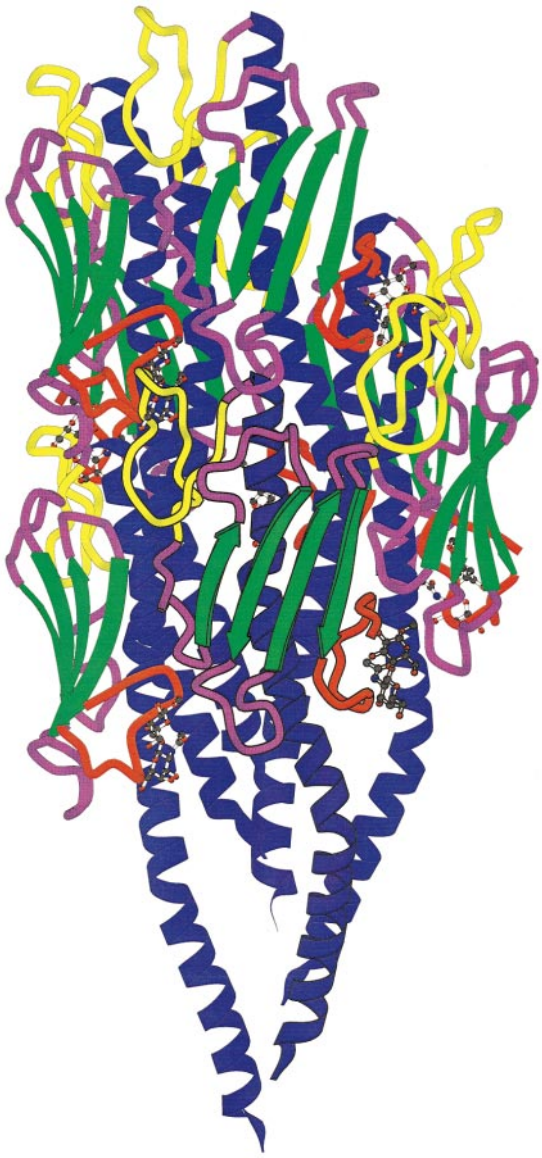

Figure 5. PAK pilin fiber model based on the assembly parameters described by Parge *et al*. (1995). The α-helix (blue) and major β-sheet (green) are shown. The minor β-sheet region (yellow coils), the C-terminal disulphide-bonded loop (red coils) and the remainder of the residues (purple coils) are also shown. The graphical elements of the second monomer from the bottom are enclosed in bold lines to highlight the boundary of a pilin monomer. The receptor analogue β-D-GalNAc(1 → 4)-β-D-Gal, represented as a ball-and-stick model, is shown in a position suggested by molecular docking studies. It has been included to illustrate the location of the binding site and the size of the disaccharide relative to the pilus. However, the depicted binding mode is not expected to be correct in its details.

rearrangement that brings the binding sites closer to the membrane, or if we assume an inverted polarity of the pilus, e.g. with the N-terminal hydrophobic helix directed towards the pilus tip. A structural rearrangement of a pilin monomer upon fiber assembly is not without precedent, as there is structural evidence for chaperone-catalyzed β-strand swapping in the assembly of the P pilus (Sauer *et al*., 1999). A similar function of chaperones might play a role in type IV pilus assembly and could explain the energy requirement of this process and the inability to spontaneously form type IV fibers from its monomers (Watts *et al*., 1982).

In a model where the N terminus of the α-helix is exposed at the tip of the pilus, the receptor-binding sites are occluded for all monomers except those at the tip of the pilus (Figure 5). This fits the biochemical observations. Proposing an inverted pilus polarity is not unreasonable, as the currently assumed polarity is not based on experimental observations (Forest & Tainer, 1997). Instead, it is based on the fact that it seems more plausible to have the hydrophobic N-terminal α-helices buried in the bacterial membrane, rather than exposed to the solvent. The latter issue would be resolved if the hydrophobic helices were proteolytically removed from the monomers located at the pilus tip. In this respect it is important to note that N-terminally truncated pilin (S-pilin), in which the hydrophobic α-helix has been deleted, is secreted by certain *N. gonorrhoeae* variants (Haas *et al*., 1987). However, the actual existence of exposed hydrophobic helices at the pilus tip does not have to be excluded. Indeed, it may even be relevant for function as initial membrane adherence through carbohydrate-binding may be followed by insertion of the hydrophobic helices, not unlike the sequence of events in hemolysins (Rossjohn *et al*., 1998).

Implications for vaccine design

A long-standing goal of the research into *P. aeruginosa* pilins has been to develop a suitable vaccine against this pathogen. Studies with antiserum generated by immunisation with various pilin-derived peptides has shown that only antiserum directed at the C-terminal receptor-binding loop inhibits binding of pili to target cells (Lee *et al*., 1989). It is interesting that the titers of this anti-peptide antiserum against the peptide itself were approximately the same as the titers against the PAK protein, unlike all other tested anti-peptide antisera which reacted much more weakly with the whole pilin molecule (Lee *et al*., 1989). This may be due to the fact that the receptor-binding loop peptide retains a native-like pair of β-turns in solution (McInnes *et al*. 1993, Campbell *et al*., 1997a). The receptor-binding loop peptide has also been found to induce a cross-reactive immune response. For optimal cross-reactivity, the disulphide bond had to be formed. This again suggests the importance of structure of the peptide for its immunogenic properties. One of the design goals for the peptide-based vaccine is therefore to stabilize a native-like conformation of the peptide. The crystal structure now provides us with a detailed picture of the native conformation for PAK pilin and the natural set of interactions that contributes to its stability. Modified peptides can now be designed to further stabilize the observed structure. The structure also identifies residues that do not appear to contribute to structural stability. Alterations at these positions can be considered to improve cross-reactivity, solubility or other desirable properties. For instance, mutation of Glu135 in PAK pilin to an Ala residue was found previously to significantly enhance cross-reactivity (Wong *et al*., 1994; Cachia *et al*., 1998).

Conclusion

The crystal structure of PAK pilin presented here provides us with the most accurate model of a type IV pilin currently available. The disulphide-bonded loop structure reveals the epitope to which a protective immune response should be generated by a peptide-based vaccine. In addition, the structure of the disulphide-bonded loop, its packing onto the remainder of the protein, and the existing biochemical data narrow down the location of the receptor-binding site to a V-shaped groove flanked by two β-turns. The position of the binding site is rather surprising given the current model for the pilus fiber and may require a re-thinking of some of our assumptions on type IV pilus fiber structure. The observation that *P. aeruginosa* pilins from several strains bind the same receptor in spite of significant sequence variation, may be linked to the fact that the proposed binding surface is dominated by main-chain atoms. Therefore, conservation of binding affinity and specificity can be achieved without strict sequence conservation as long as the main-chain conformation is retained.

Materials and Methods

The production of truncated PAK pilin follows the same procedures used to prepare a truncated strain K122-4 pilin (Irvin *et al*., unpublished results). In brief, DNA encoding PAK pilin residues 29 to 144 was obtained by PCR from the full-length PAK pilin gene (Pasloske *et al*., 1988) using the oligonucleotides SP-5 (5′-GGAATTCGCTCGTTCGGAAGGCGCA-3′) and SP-6 (5′-CCCAAGCTTCTTACAAAATTACCTAGA-3′). The amplified DNA fragment was purified on an 8% (w/v) polyacrylamide gel, digested with *Eco*RI and *Hin*dIII, purified by phenol-extraction, and ligated in-frame behind the OmpA leader sequence of plasmid pRLD (Tripet *et al*. 1996). The construct was sequenced in both directions and was found to contain no errors. *E. coli* BL21 harbouring the expression plasmid was grown at 37 °C, with shaking, in LB broth containing carbenicillin (100 μg/ml) to an absorbance of A_{550} 0.5-0.7. Expression of the truncated protein was induced by the addition of isopropyl β-D-thiogalactopyranoside to a final concentration of 1 mM. The bacteria were grown for an

additional six hours at 37 °C before the expressed periplasmic protein was extracted by osmotic shock. The protein was purified by sequential column chromatography using DEAE-cellulose (15 cm × 0.5 cm, DE-32, Whatman Inc., Clifton, NJ), carboxymethyl-cellulose (30 cm × 2 cm, CM-52, Whatman Inc., Clifton, NJ), and Sephadex G-75 (48 cm × 1 cm, Superfine, Pharmacia Fine Chemicals, Uppsala, Sweden) columns, respectively. The purified protein appeared as a single band on SDS-PAGE and its identity was confirmed by N-terminal sequencing and immunoblotting using rabbit polyclonal anti-PAK pilus antisera. The final product contained seven residues derived from the expression vector (Ala-Leu-Glu-Gly-Thr-Glu-Phe), followed by residues 29 to 144 of mature PAK pilin. Crystals were grown by mixing 3 µl of a 10 mg/ml protein solution in doubly distilled water with 3 µl of precipitant (60 % $(NH_4)_2SO_4$, 0.1 M Hepes (pH 8.2)) on a cover slip which was then suspended over a well with 1 ml of precipitant solution. Crystals of irregular shape grew normally within ten days at room temperature and were crushed to serve as nuclei for microseeding. To this end hanging drops were set up as described above but after 6 to 12 hours of pre-equilibration some micro seeds were added. After growth, well-shaped small crystals (<0.05 mm) were selected, washed in 58 % (v/v) $(NH_4)_2SO_4$, 0.1 M Hepes (pH 8.2), and added to fresh pre-equilibrated hanging drops. Crystals typically grew to a size of 0.3 mm × 0.3 mm × 0.2 mm and had spacegroup $P4_12_12$ with cell dimensions of $a = b = 38.1$ Å and $c = 149.8$ Å.

The native and gold cyanide derivative datasets were each collected from a single crystal at room temperature. CuK^{α} radiation, generated by an Elliot GX-13 rotating anode generator (0.1 mm cathode operated at 40 kV, 60 mA), was focussed by a Supper mirror and data were collected on a MAR300 image plate detector. The gold cyanide dataset was collected from a crystal that had been soaked for 22 hours in precipitant solution containing 20 mM $KAu(CN)_2$. The uranyl acetate dataset was collected at room temperature on beamline 7.2 at the SRS (Daresbury, UK) using a wavelength of 1.448 Å. A full dataset was collected from a single crystal that had been soaked for 24 hours in 60 % $(NH_4)_2SO_4$, 0.1 M Hepes (pH 7.5), 40 mM uranyl acetate. All data were integrated with MOSFLM (Leslie, 1992) and scaled by SCALA (Evans, 1993). Scaling of the derivatives used the native data as a reference and allowed variation of the scale factor over the detector surface. Further data processing was carried out with the programs from the CCP4 package (Collaborative Computing Project, Number 4, 1994). Localisation of the heavy atoms and subsequent phasing was carried out by SOLVE (Terwilliger & Berendzen, 1999) using isomorphous and anomalous differences. Both $P4_12_12$ and $P4_32_12$ space groups were tested, and the results for $P4_12_12$ were clearly superior. The resulting density map was finally subjected to solvent flattening, histogram mapping and multi-resolution modification using DM (Cowtan, 1994). This improved the figure of merit from 0.56 to 0.80 for all data to 2.5 Å. Data processing and phasing statistics are given in Table 1.

Building of the main chain was carried out by wARP (Lamzin & Wilson, 1997) using the map coefficients from DM and all available native data. Side-chains were built using the Xfit program from the XtalView package (McRee, 1999). The model was refined in four rounds using the program REFMAC (Murshudov *et al*., 1997) and standard protocols. An explicit bulk solvent correction was calculated with CNS (Brünger *et al*., 1998) and scaled by 0.85 prior to use in REFMAC as this was found to give a better agreement with the observed structure factors based on structure factor correlation analysis. All data were used for refinement with 750 reflections set aside for R_{free} calculation and σ_A estimation. Each refinement round was followed by model building with Xfit using cross-validated SIGMAA-weighted density maps (Read, 1997). Water molecules were added using wARP. Several residues were modeled with alternative side-chain conformations. In these cases the occupancy was fixed at 0.5 for both conformations while the *B*-factors were refined. Individual anisotropic *B*-values were refined for two cycles at the very end of

Table 1. Diffraction data and phasing statistics

Data set	Native	$KAu(CN)_2$	UO_2Ac
Cell *a*, *c* (Å)	38.11, 149.78	38.08, 149.36	38.41, 151.60
Resolution range (Å)	1.63-37.44	2.44-25.36	2.25-36.94
Observations	110,628	19,342	71,853
Unique data	14,500	4512	5747
Completeness (%)	99.5 (96.4)[a]	96.7 (87.1)[b]	99.6 (97.2)[c]
$I/\sigma I$	26.7 (10.1)[a]	21.12 (13.5)[b]	61.9 (16.7)[c]
R_{sym} (%)	4.9 (19.0)[a]	4.3 (8.7)[b]	6.5 (8.0)[c]
R^{d}_{meas} (%)	5.4 (22.7)[a]	5.3 (11.1)[b]	7.1 (8.7)[c]
R^{e}_{anom} (%)	1.7 (7.9)[a]	2.7 (6.0)[b]	5.6 (6.9)[c]
R^{f}_{native} (%)	-	10.3 (11.6)[b]	33.1 (38.1)[c]
Norm. prob.[g]	-	5.2/− 0.1	22.3/− 1.2
Figure of merit	-	0.38	0.31
Centric phasing power	-	0.9	0.94
Acentric phasing power	-	1.09	1.06
R_{Cullis}	-	0.63	0.62
No. of sites	-	4	4

[a] Between 1.63 and 1.72 Å resolution.
[b] Between 2.53 and 2.44 Å resolution.
[c] Between 2.25 and 2.32 Å resolution.
[d] Multiplicity corrected *R*-factor (Diederichs & Karplus, 1997), relative to mean I^+ and I^- for derivatives.
[e] $\Sigma\ |\langle I^+\rangle \rightarrow \langle I^-\rangle| / \Sigma\ (\langle I^+\rangle + \langle I^-\rangle)$.
[f] *R*-factor between native and derivative after scaling with SCALEIT.
[g] Normal probability analysis (Howell & Smith, 1992), as calculated by SCALEIT (gradient/intercept).

refinement using sphericity restraints of 2.0. The validity of anisotropic *B*-value refinement at the given resolution was supported by the R_{free} statistic which dropped from 18.6 to 18.1 %.

Protein Data Bank deposition codes

Coordinates and structure factor data for native and derivative amplitudes have been deposited at the RCSB Protein Data Bank with deposition codes 1DZO and r1DZOsf, respectively.

Acknowledgments

The authors thank Dr G. Vriend for searching the structural database for proteins with a *cis*Pro-X-*cis*Pro motif, the Alberta Peptide Institute for N-terminal sequencing of the PAK construct, and Dr D. Bundle for providing the coordinates of the receptor analogue. The Laboratory of Molecular Biology (Cambridge, UK) is acknowledged for generous access to their X-ray diffraction equipment. The technical assistance of Dr C. Grant and M. Kaplan is gratefully acknowledged. This work was supported by grants from the Howard Hughes Medical Institute and the Wellcome Trust to R.J.R., and Canadian Bacterial Diseases Network funding to R.T.I.

References

Aho, E. L., Botten, J. W., Hall, R. J., Larson, M. K. & Ness, J. K. (1997). Characterization of a class II pilin expression locus from *Neisseria meningitidis*: evidence for increased diversity among pilin genes in pathogenic *Neisseria* species. *Infect. Immun.* **65**, 2613-2620.

Alm, R. A., Hallinan, J. P., Watson, A. A. & Mattick, J. S. (1996). Fimbrial biogenesis genes of *Pseudomonas aeruginosa*: pilW and pilX increase the similarity of type IV fimbriae to the GSP protein-secretion systems and pilY1 encodes a gonococcal PilC homologue. *Mol. Microbiol.* **22**, 161-173.

Alonso, A., Campanario, E. & Martinez, J. L. (1999). Emergence of multidrug-resistant mutants is increased under antibiotic selective pressure in *Pseudomonas aeruginosa*. *Microbiology,* **145**, 2857-2862.

Baker, N. R., Minor, V., Deal, C., Shahrabadi, M. S., Simpson, D. A. & Woods, D. E. (1991). *Pseudomonas aeruginosa* exoenzyme S is an adhesion. *Infect. Immun.* **59**, 2859-2863.

Bang, R. L., Gang, R. K., Sanyal, S. C., Mokaddas, E. & Ebrahim, M. K. (1998). Burn septicaemia: an analysis of 79 patients. *Burns,* **24**, 354-361.

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

Beachey, E. H. (1981). Bacterial adherence: adhesin-receptor interactions mediating the attachtment of bacteria to mucosal surfaces. *J. Infect. Dis.* **143**, 325-345.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography and NMR System: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.

Cachia, P. J., Glasier, L. M., Hodgins, R. R. W., Wong, W. Y., Irvin, R. T. & Hodges, R. S. (1998). The use of synthetic peptides in the design of a consensus sequence vaccine for *Pseudomonas aeruginosa*. *J. Pept. Res.* **52**, 289-299.

Campbell, A. P., Wong, W. Y., Houston, M., Jr, Schweizer, F., Cachia, P. J., Irvin, R. T., Hindsgaul, O., Hodges, R. S. & Sykes, B. D. (1997a). Interaction of the receptor binding domains of *Pseudomonas aeruginosa* pili strains PAK, PAO, KB7 and P1 to a cross-reactive antibody and receptor analog: implications for synthetic vaccine design. *J. Mol. Biol.* **267**, 382-402.

Campbell, A. P., Bautista, D. L., Tripet, B., Wong, W. Y., Irvin, R. T., Hodges, R. S. & Sykes, B. D. (1997b). Solution secondary structure of a bacterially expressed peptide from the receptor binding domain of *Pseudomonas aeruginosa* pili strain PAK: a heteronuclear multidimensional NMR study. *Biochemistry,* **36**, 12791-12801.

Castric, P. A. (1995). pilO, a gene required for glycosylation of *Pseudomonas aeruginosa* 1244 pilin. *Microbiology,* **141**, 1247-1254.

Castric, P. A. & Deal, C. D. (1994). Differentiation of *Pseudomonas aeruginosa* pili based on sequence and B-cell epitope analyses. *Infect. Immun.* **62**, 371-376.

Collaborative Computational Project No. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.

Comolli, J. C., Waite, L. L., Mostov, K. E. & Engel, J. N. (1999). Pili binding to asialo-GM1 on epithelial cells can mediate cytotoxicity or bacterial internalization by *Pseudomonas aeruginosa*. *Infect. Immun.* **67**, 3207-3214.

Cowtan, K. (1994). DM: an automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM News. Protein Crystallog.* **31**, 34-48.

Crouch Brewer, S., Wunderink, R. G., Jones, C. B. & Leeper, K. V., Jr (1996). Ventilator-associated pneumonia due to *Pseudomonas aeruginosa*. *Chest,* **109**, 1019-1029.

Dart, J. K. & Seal, D. V. (1988). Pathogenesis and therapy of *Pseudomonas aeruginosa* keratitis. *Eye,* **2**, S46-S55.

de Bentzmann, S., Roger, P., Dupuit, F., Bajolet-Laudinat, O., Fuchey, C., Plotkowski, M. C. & Puchelle, E. (1996). Asialo GM1 is a receptor for *Pseudomonas aeruginosa* adherence to regenerating respiratory epithelial cells. *Infect. Immun.* **64**, 1582-1588.

Diederichs, K. & Karplus, P. A. (1997). Improved *R*-factors for diffraction data analysis in macromolecular crystallography. *Nature Struct. Biol.* **4**, 269-275.

Doig, P., Smith, N. R., Todd, T. & Irvin, R. T. (1987). Characterization of the binding of *Pseudomonas aeruginosa* alginate to human epithelial cells. *Infect. Immun.* **55**, 1517-1522.

Doig, P., Todd, T., Sastry, P. A., Lee, K. K., Hodges, R. S., Paranchych, W. & Irvin, R. T. (1988). Role of pili in adhesion of *Pseudomonas aeruginosa* to human respiratory epithelial cells. *Infect. Immun.* **56**, 1641-1646.

Doig, P., Paranchych, W., Sastry, P. A. & Irvin, R. T. (1989). Human buccal epithelial cell receptors of *Pseudomonas aeruginosa*: identification of glycoproteins with pilus binding activity. *Can. J. Microbiol.* **35**, 1141-1145.

Doig, P., Sastry, P. A., Hodges, R. S., Lee, K. K., Paranchych, W. & Irvin, R. T. (1990). Inhibition of pilus-mediated adhesion of *Pseudomonas aeruginosa* to human buccal epithelial cells by monoclonal antibodies directed against pili. *Infect. Immun.* **58**, 124-130.

Evans, P. R. (1993). *SCALA version 3.3*, MRC Laboratory of Molecular Biology, Cambridge, UK.

Farinha, M. A., Conway, B. D., Glasier, L. M., Ellert, N. W., Irvin, R. T., Sherburne, R. & Paranchych, W. (1994). Alteration of the pilin adhesin of *Pseudomonas aeruginosa* PAO results in normal pilus biogenesis but a loss of adherence to human pneumocyte cells and decreased virulence in mice. *Infect. Immun.* **62**, 4118-4123.

Folkhard, W., Marvin, D. A., Watts, T. H. & Paranchych, W. (1981). Structure of polar pili from *Pseudomonas aeruginosa* strains K and O. *J. Mol. Biol.* **149**, 79-93.

Forest, K. T. & Tainer, J. A. (1997). Type-4 pilus-structure: outside to inside and top to bottom: a minireview. *Gene,* **192**, 165-169.

Forest, K. T., Dunham, S. A., Koomey, M. & Tainer, J. A. (1999). Crystallographic structure reveals phosphorylated pilin from *Neisseria*: phosphoserine sites modify type IV pilus surface chemistry and fibre morphology. *Mol. Microbiol.* **31**, 743-752.

Gupta, S. K., Berk, R. S., Masinick, S. & Hazlett, L. D. (1994). Pili and lipopolysaccharide of *Pseudomonas aeruginosa* bind to the glycolipid asialo GM1. *Infect. Immun.* **62**, 4572-4579.

Haas, R., Schwarz, H. & Meyer, T. F. (1987). Release of soluble pilin antigen coupled with gene conversion in *Neisseria gonorrhoeae*. *Proc. Natl Acad. Sci. USA,* **84**, 9079-9083.

Hahn, H. P. (1997). The type-4 pilus is the major virulence-associated adhesin of *Pseudomonas aeruginosa*: a review. *Gene,* **192**, 99-108.

Hart, T. N., Ness, S. R. & Read, R. J. (1997). Critical evaluation of the research docking program for the CASP2 challenge. *Proteins: Struct. Funct. Genet.* **Suppl 1**, 205-209.

Hazlett, L., Rudner, X., Masinick, S., Ireland, M. & Gupta, S. (1995). In the immature mouse, *Pseudomonas aeruginosa* pili bind a 57-kd (α2-6) sialylated corneal epithelial cell surface protein: a first step in infection. *Invest. Ophthalmol. Vis. Sci.* **36**, 634-643.

Howell, P. L. & Smith, G. D. (1992). Identification of heavy-atom derivatives by normal probability methods. *J. Appl. Crystallog.* **25**, 81-86.

Irvin, R. T., Doig, P., Lee, K. K., Sastry, P. A., Paranchych, W., Todd, T. & Hodges, R. S. (1989). Characterization of the *Pseudomonas aeruginosa* pilus adhesin: confirmation that the pilin structural protein subunit contains a human epithelial cell-binding domain. *Infect. Immun.* **57**, 3720-3726.

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

Lamzin, V. S. & Wilson, K. S. (1997). Automated refinement for protein crystallography. *Methods Enzymol.* **277**, 269-305.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.

Lee, K. K., Doig, P., Irvin, R. T., Paranchych, W. & Hodges, R. S. (1989). Mapping the surface regions of *Pseudomonas aeruginosa* PAK pilin: the importance of the C-terminal region for adherence to human buccal epithelial cells. *Mol. Microbiol.* **3**, 1493-1499.

Lee, K. K., Paranchych, W. & Hodges, R. S. (1990). Cross-reactive and strain-specific antipeptide antibodies to *Pseudomonas aeruginosa* PAK and PAO pili. *Infect. Immun.* **58**, 2727-2732.

Lee, K. K., Sheth, H. B., Wong, W. Y., Sherburne, R., Paranchych, W., Hodges, R. S., Lingwood, C. A., Krivan, H. & Irvin, R. T. (1994). The binding of *Pseudomonas aeruginosa* pili to glycosphingolipids is a tip-associated event involving the C-terminal region of the structural pilin subunit. *Mol. Microbiol.* **11**, 705-713.

Leslie, A. G. W. (1992). *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography.* vol. 26, Daresbury Laboratory, Warrington, UK.

Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L. & Read, R. J. (1998). Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry,* **37**, 1777-1788.

Lu, H. M., Motley, S. T. & Lory, S. (1997). Interactions of the components of the general secretion pathway: role of *Pseudomonas aeruginosa* type IV pilin subunits in complex formation and extracellular protein secretion. *Mol. Microbiol.* **25**, 247-259.

May, T. B., Shinabarger, D., Maharaj, R., Kato, J., Chu, L., DeVault, J. D., Roychoudhury, S., Zielinski, N. A., Berry, A., Rothmel, R. K., Misra, T. K. & Chakrabarty, A. M. (1991). Alginate synthesis by *Pseudomona aeruginosa*: a key pathogenic factor in chronic pulmonary infections of cystic fibrosis patients. *Clin. Microbiol. Rev.* **4**, 191-206.

McInnes, C., Sonnichsen, F. D., Kay, C. M., Hodges, R. S. & Sykes, B. D. (1993). NMR solution structure and flexibility of a peptide antigen representing the receptor binding domain of *Pseudomonas aeruginosa*. *Biochemistry,* **32**, 13432-13440.

McRee, D. E. (1999). XtalView/Xfit: a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156-165.

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

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D,* **53**, 240-255.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A,* **50**, 157-163.

Paranchych, W., Sastry, P. A., Frost, L. S., Carpenter, M., Armstrong, G. D. & Watts, T. H. (1979). Biochemical studies on pili isolated from *Pseudomonas aeruginosa* strain PAO. *Can. J. Microbiol.* **25**, 1175-1181.

Parge, H. E., Forest, K. T., Hickey, M. J., Christensen, D. A., Getzoff, E. D. & Tainer, J. A. (1995). Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature,* **378**, 32-38.

Pasloske, B. L., Sastry, P. A., Finlay, B. B. & Paranchych, W. (1988). Two unusual pilin sequences from different isolates of *Pseudomonas aeruginosa*. *J. Bacteriol.* **170**, 3738-3741.

Ramphal, R., Sadoff, J. C., Pyle, M. & Silipigni, J. D. (1984). Role of pili in the adherence of *Pseudomonas aeruginosa* to injured tracheal epithelium. *Infect. Immun.* **44**, 38-40.

Ramphal, R., Arora, S. K. & Ritchings, B. W. (1996). Recognition of mucin by the adhesin-flagellar system of *Pseudomonas aeruginosa*. *Am. J. Respir. Crit. Care Med.* **154**, S170-S174.

Read, R. J. (1997). Model phases: probabilities and bias. *Methods Enzymol.* **277**, 110-128.

Rini, J. M. (1995). Lectin structure. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551-577.

Rossjohn, J., Feil, S. C., McKinstry, W. J., Tsernoglou, D., van der Goot, G., Buckley, J. T. & Parker, M. W. (1998). Aerolysin - a paradigm for membrane insertion of beta-sheet protein toxins? *J. Struct. Biol.* **121**, 92-100.

Rudel, T., Scheuerpflug, I. & Meyer, T. F. (1995). *Neisseria* PilC protein identified as type-4 pilus tip-located adhesin. *Nature,* **373**, 357-359.

Rudner, X. L., Zheng, Z., Berk, R. S., Irvin, R. T. & Hazlett, L. D. (1992). Corneal epithelial glycoproteins exhibit *Pseudomonas aeruginosa* pilus binding activity. *Invest. Ophthalmol. Vis. Sci.* **33**, 2185-2193.

Saiman, L. & Prince, A. (1993). *Pseudomonas aeruginosa* pili bind to asialoGM1 which is increased on the surface of cystic fibrosis epithelial cells. *J. Clin. Invest.* **92**, 1875-1880.

Sauer, F. G., Futterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J. & Waksman, G. (1999). Structural basis of chaperone function and pilus biogenesis. *Science,* **285**, 1058-1061.

Scheuerpflug, I., Rudel, T., Ryll, R., Pandit, J. & Meyer, T. F. (1999). Roles of PilC and PilE proteins in pilus-mediated adherence of *Neisseria gonorrhoeae* and *Neisseria meningitidis* to human erythrocytes and endothelial and epithelial cells. *Infect. Immun.* **67**, 834-843.

Sheth, H. B., Glasier, L. M., Ellert, N. W., Cachia, P., Kohn, W., Lee, K. K., Paranchych, W., Hodges, R. S. & Irvin, R. T. (1995). Development of an anti-adhesive vaccine for *Pseudomonas aeruginosa* targeting the C-terminal region of the pilin structural protein. *Biomed. Pept. Proteins Nucl. Acids,* **1**, 141-148.

Stimson, E., Virji, M., Barker, S., Panico, M., Blench, I., Saunders, J., Payne, G., Moxon, E. R., Dell, A. & Morris, H. R. (1996). Discovery of a novel protein modification: alpha-glycerophosphate is a substituent of meningococcal pilin. *Biochem. J.* **15**, 29-33.

Tang, H., Kays, M. & Prince, A. (1995). Role of *Pseudomonas aeruginosa* pili in acute pulmonary infection. *Infect. Immun.* **63**, 1278-1285.

Terwilliger, T. C. & Berendzen, J. (1999). Automated MIR and MAD structure solution. *Acta Crystallog. sect. D,* **55**, 849-861.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). Improved sensitivity of profile searches through the use of sequence weights and gap excision. *Comput. Appl. Biosci.* **10**, 19-29.

Tripet, B., Yu, L., Bautista, D., Wong, W. Y., Irvin, R. T. & Hodges, R. S. (1996). Engineering a de novo designed coiled-coil dimerization domain for rapid detection, purification and characterization of recombinant peptides. *Protein Eng.* **9**, 1029-1042.

Tummler, B., Bosshammer, J., Breitenstein, S., Brockhausen, I., Gudowius, P., Herrmann, C., Herrmann, S., Heuer, T., Kubesch, P., Mekus, F., Romling, U., Schmidt, K. D., Spangenberg, C. & Walter, S. (1997). Infections with *Pseudomonas aeruginosa* in patients with cystic fibrosis. *Behring Inst. Mitt.* **98**, 249-255.

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52-56.

Watts, T. H., Scraba, D. G. & Paranchych, W. (1982). Formation of 9-nm filaments from pilin monomers obtained by octyl-glucoside dissociation of *Pseudomonas aeruginosa* pili. *J. Bacteriol.* **151**, 1508-1513.

Wong, W. Y., Sheth, H. B., Irvin, R. T. & Hodges, R. S. (1994). Enhancing strain cross-reactivity of a synthetic peptide vaccine for *Pseudomonas aeruginosa*. *Pediatr. Pulmonol.* **157-158**, S10.

Wong, W. Y., Campbell, A. P., McInnes, C., Sykes, B. D., Paranchych, W., Irvin, R. T. & Hodges, R. S. (1995). Structure-function analysis of the adherence-binding domain on the pilin of *Pseudomonas aeruginosa* strains PAK and KB7. *Biochemistry,* **34**, 12963-12972.

Wu, X., Gupta, S. K. & Hazlett, L. D. (1995). Characterization of *P. aeruginosa* pili binding human corneal epithelial proteins. *Curr. Eye Res.* **14**, 969-977.

Edited by R. Huber

(Received 8 February 2000; received in revised form 18 April 2000; accepted 18 April 2000)