Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the S pilus periplasmic chaperone SfaE at 2.2 Å resolution

Stefan D. Knight,[a]* Devapriya Choudhury,[a]† Scott Hultgren,[b] Jerome Pinkner,[b] Vivian Stojanoff[c]‡ and Andrew Thompson[d]

[a]Department of Molecular Biology, Uppsala Biomedical Centre, Swedish University of Agricultural Sciences, Box 590, S-753 24 Uppsala, Sweden, [b]Department of Molecular Microbiology, Washington University School of Medicine, St Louis, Missouri 63110, USA, [c]ESRF, Avenue des Martyrs, 38400 Grenoble, France, and [d]EMBL Grenoble Outstation, c/o Avenue des Martyrs, BP 156X, 38042 Grenoble, France

† Present address: Centre for Biotechnology, Jawaharlal Nehru University, New Delhi 110067, India.
‡ Present address: Brookhaven National Laboratory, National Synchrotron Light Source, Building 725D, Upton NY 11973, USA.

Correspondence e-mail: stefan@xray.bmc.uu.se

Received 8 February 2002
Accepted 2 April 2002

PDB Reference: SfaE, 1l4i, r1l4isf.

S pili are sialic acid binding hair-like appendages expressed by pathogenic strains of *Escherichia coli*. The presence of S pili has been implicated as a virulence factor in both urinary-tract infections and new-born meningitis. Assembly of S pili proceeds *via* the ubiquitous chaperone/usher pathway. Previously, structures of the homologous chaperones PapD and FimC involved in assembly of P and type-1 pili, respectively, have been solved. Here, the 2.2 Å X-ray structure of the S pilus chaperone SfaE is reported. SfaE has the same overall L-shaped structure as PapD and FimC, with two immunoglobulin-like domains oriented at about a 90° angle to each other. Conserved residues in the subunit-binding cleft known to be critical for chaperone function occupy essentially identical positions in SfaE, FimC and PapD. As in free PapD and FimC, the long *F*1–*G*1 loop connecting the two last strands of the N-terminal domain is disordered. SfaE crystallizes as a dimer with an extensive dimer interface involving the subunit-binding surfaces of the chaperone. Dimerization *via* these regions has previously been observed for PapD and might be a general side effect arising from the subunit-binding properties of periplasmic chaperones. The domain interface contains an extended hydrogen-bond network involving three invariant charged residues and two structurally conserved water molecules. It is suggested that disruption of the domain interactions may destabilize the N-terminal domain through exposure of three conserved hydrophobic residues, thereby promoting release of pilus subunits during pilus assembly.

1. Introduction

Early steps in bacterial pathogenesis usually require establishment of intimate host–pathogen interactions. This is often accomplished by specific recognition of carbohydrate structures on the host cell by dedicated attachment organelles incorporating lectin-like proteins. Pathogenic *E. coli* can express a variety of carbohydrate-binding pili allowing attachment to and colonization of various host tissues. For example, Galα(1–4)Gal-binding P pili, a common virulence factor in pyelonephritis-causing *E. coli*, allow colonization of the upper urinary tract. Binding to mannose-containing uroplakin receptors in the bladder uroepithelium *via* type-1 pili allow uropathogenic *E. coli* to colonize the bladder and cause cystitis. S pili have been implicated both in urinary-tract infections and in neonatal meningitis (Hacker, 1992). They bind to sialic acid containing oligosaccharides (Korhonen *et al.*, 1984), giving the bacteria the ability to recognize a variety of tissue surfaces including epithelial cells of the urinary tract (Virkola *et al.*, 1988) and the kidneys (Korhonen *et al.*, 1986) and brain endothelial cells (Prasadarao *et al.*, 1993, 1997). S pili

also mediate binding to a variety of other targets including erythrocytes, the cellular form of human fibronectin (Saren *et al.*, 1999) and the cellular matrix protein laminin (Virkola *et al.*, 1993).

S, P and type-1 pili are bipartite structures consisting of a 7 nm thick rod and a thin tip fibrillum which is relatively long and flexible in S and P pili, but short and stubby in type 1 pili (Jones *et al.*, 1996). The rod is a polymer made from the major subunit (SfaA in S pili, PapA in P pili and FimA in type-1 pili). The tip fibrillum is a heteropolymer which incorporates the carbohydrate-binding adhesin as well as structural minor subunits. The sialic acid binding function of S pili has been attributed to SfaS (Khan *et al.*, 2000; Morschhauser *et al.*, 1993), although a subunit (SfaH) which is very similar in primary sequence to the mannose-binding type 1 adhesin FimH is incorporated into the tip fibrillum. Biogenesis of S, P, type 1 and many other adhesive pili requires a specialized periplasmic chaperone protein which sequesters subunits as they enter the periplasm and protects them from premature association (Knight *et al.*, 2000; Sauer, Knight *et al.*, 2000; Sauer, Mulvey *et al.*, 2000). The chaperone–subunit complexes are subsequently presented to an outer membrane protein called the usher, where the chaperone dissociates and its place is taken up by the next subunit in the pilus. This process is repeated; with each successive addition of a subunit, the nascent pilus extrudes out of the cell surface.

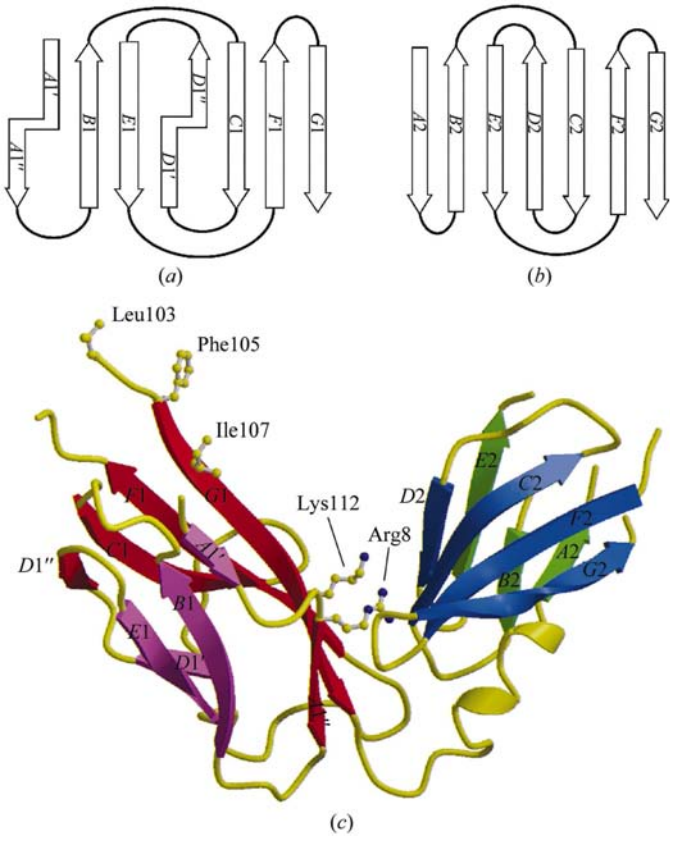

Figure 1
Topology diagrams of the N-terminal (*a*) and C-terminal (*b*) domain of periplasmic chaperones. (*c*) Ribbon diagram of SfaE. The three invariantly hydrophobic residues forming the *G*1 donor-strand motif (Leu103, Phe105, Ile107) are shown as ball-and-stick, as are the two invariant positively charged residues (Arg8 and Lys112) at the base of the subunit-binding cleft.

Periplasmic chaperones consist of two domains with immunoglobulin-like (Ig-like) topology joined at approximately right angles with a large cleft between the two domains (Figs. 1*a* and 1*b*; Choudhury *et al.*, 1999; Holmgren & Branden, 1989; Holmgren *et al.*, 1992; Pellecchia *et al.*, 1998; Sauer *et al.*, 1999). The two last β-strands of the N-terminal domain (*F*1 and *G*1) are connected by a long flexible loop that protrudes away from the body of the domain. The *F*1–*G*1 loop and the beginning of the *G*1 β-strand contains a conserved β-zipper motif of alternating hydrophobic and hydrophilic residues that is critical for chaperone function (Hung *et al.*, 1996). The hydrophobic side chains form a surface-exposed hydrophobic ridge in the free chaperone that is part of an essential subunit-binding surface (Choudhury *et al.*, 1999; Sauer *et al.*, 1999; Soto *et al.*, 1998). Two basic residues (Arg8 and Lys112 in PapD) at the bottom of the cleft between the two chaperone domains are also indispensable for chaperone function and are conserved throughout the periplasmic chaperone family (Hung *et al.*, 1996; Kuehn *et al.*, 1993).

The crystal structures of the type 1 pilus FimC–FimH chaperone–adhesin complex (Choudhury *et al.*, 1999) and the P pilus PapD–PapK chaperone–adapter subunit complex (Sauer *et al.*, 1999) have revealed that pilin subunits, just like the chaperones, have Ig-like folds. However, the final (seventh) β-strand of the fold is missing in the pilin subunits, which creates a cleft on the surface of the subunit where part of the hydrophobic core is exposed. Periplasmic chaperones bind to pilin subunits by inserting the G1 donor β-strand into this cleft in a process called donor-strand complementation. The chaperone G1 β-strand is inserted into the pilin-acceptor cleft with extensive main-chain-to-main-chain hydrogen bonding between the G1 donor strand and the two pilin edge strands that define the perimeters of the acceptor cleft. Alternating hydrophobic side chains from the conserved β-zipper motif in the G1 strand bind in subpockets within the acceptor cleft and complete the hydrophobic core of the pilin domain.

The periplasmic chaperone SfaE is responsible for the transport of S-fimbrial subunits across the periplasm before their assembly into the pilus. In this paper, we present the crystal structure of the free form of SfaE at 2.2 Å resolution and analyse it in the context of a number of recently determined crystal structures of periplasmic chaperones, both in complex with pilus subunits and in the free form.

2. Materials and methods

2.1. Fermentation and periplasmic preparation

The plasmid pJP12 encoding SfaE in pTrc99A was expressed in C600. Overnight cultures were diluted 1:20 into Luria broth and grown to an OD_{600} of 2.5 followed by induction with 0.1 m*M* IPTG for 1 h. Periplasms were

prepared as previously described (Jones *et al.*, 1993; Slonim *et al.*, 1992).

2.2. Protein purification

SfaE was purified from the periplasm of C600/pJP12. Periplasmic extracts were dialysed against 20 m*M* MES pH 6.0, injected onto a Source 15S column (Pharmacia) and SfaE was eluted at 45 m*M* NaCl. The eluate was injected onto a Source 15PHE column (Pharmacia) in 1.2 *M* $(NH_4)_2SO_4$/20 m*M* MES pH 6.0 and SfaE eluted at 0.7 *M* $(NH_4)_2SO_4$. SfaE was then dialysed against 20 m*M* MES pH 6.8.

2.3. Crystallization and data collection

Single crystals of SfaE suitable for X-ray diffraction studies could easily be grown under a number of conditions using either polyethylene glycol or ammonium sulfate as precipitant and ranging in pH from 4.5 to 8.5. Crystals were grown by hanging-drop vapour diffusion by placing a 2 µl drop of a 10 mg ml^{-1} solution of SfaE in 20 m*M* HEPES on a coverslip, mixing with a 2 µl drop of the reservoir solution and equilibrating for about one week. Regardless of the crystallization conditions, large $P2_12_12_1$ crystals with similar unit-cell parameters $a = 53.6$, $b = 83.9$, $c = 97.3$ Å were always obtained. The rod-shaped crystals grew to maximum dimensions of 0.2 × 0.2 × 0.4 mm in about two weeks. Initial native data were collected from a crystal grown using a reservoir solution consisting of 30% PEG 4000, 0.1 *M* Tris–HCl buffer pH 8.5 and 0.2 *M* sodium acetate. Heavy-atom derivatives were prepared by soaking native SfaE crystals in the crystallization solution with the addition of $Pt(NH_4)(NO_2)_4$, $K_2Pt(SCN)_4$ or K_2PtCl_4. The initial native and heavy-atom derivative data were collected at room temperature in the laboratory on a Xentronics MWPC detector using Cu $K\alpha$ radiation from a rotating anode, processed using *XDS* (Kabsch, 1988) and scaled with *AGROVATA*/*ROTAVATA* (Collaborative Computational Project, Number 4, 1994). Single-wavelength data for two selenium-substituted crystals grown under the same low-pH condition as native SfaE were collected on an R-AXIS IIC imaging-plate detector in the laboratory and used as additional heavy-atom derivatives. Native data to 2.2 Å Bragg spacing from a crystal grown at low pH (reservoir solution: 30% PEG 4000, 0.1 *M* sodium citrate buffer pH 5.6, 0.2 *M* ammonium acetate) were subsequently collected at 288 K using a MAR Research imaging plate at beamline X31 (DESY, Hamburg) and processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). For data-collection statistics see Table 1.

2.4. Structure solution and refinement

Crystals soaked in heavy-atom solutions were radiation sensitive and only diffracted to low resolution (Table 1); the resulting difference Patterson maps were therefore quite noisy. Nonetheless, visual inspection of Harker sections from the three Pt derivatives revealed a common pattern of peaks that could be explained by the presence of two common Pt sites in each of the derivatives. *SHARP* (Fortelle & Bricogne, 1997) refinement of these two sites using the initial native data and the K_2PtCl_4 derivative data gave an overall phasing power of 0.895 ($R_{cullis} = 0.857$) for centric and 1.168 ($R_{cullis} = 0.848$) for acentric data and a discouraging overall figure of merit of 0.205. Nevertheless, *SOLOMON* (Abrahams & Leslie, 1996) solvent flipping of the *SHARP* map assuming 50% solvent content gave an electron-density map with clear density for two copies of SfaE in the asymmetric unit. A homology model of SfaE based on the 2.0 Å model of PapD (Holmgren, 1993) could easily be inserted into the density for one of the two SfaE protomers, placing the S^{γ} atom of Met166 next to one of the two Pt sites. Difference Fourier maps for the SeMet crystals were calculated using the solvent-flattened *SHARP* phases and peaks in these maps were compared with the Se positions predicted by our model, allowing the identification of four Se positions corresponding to SeMet residues (Met57, Met75, Met93 and Met166) in the placed model. In addition, a fifth Se position tentatively corresponding to Met166 in the second SfaE copy was identified. *SHARP* refinement of all five selenium and the two platinum positions followed by solvent flipping with *SOLOMON* assuming 50% solvent content resulted in an excellent electron-density map which permitted tracing of the entire molecule using *O* (Jones *et al.*, 1991). Initial refinement of a model consisting of two identical copies of SfaE in the asymmetric unit carried out against the 2.8 Å native data using a standard simulated-annealing protocol with *X-PLOR* (Brünger, 1993) gave $R = 49.5\%$ and $R_{free} = 46.4\%$. Subsequent rounds of remodelling and refinement against the 2.2 Å native data using both *X-PLOR* and *REFMAC* (Collaborative Computational Project, Number 4, 1994; Murshudov *et al.*, 1997) resulted in a model with $R = 19.0\%$ and $R_{free} = 24.4\%$. Refinement statistics are given in Table 2.

Table 1
Data-collection statistics.

Data set	d_{min} (Å)	N_{meas}	N_{unique}	Completeness (%)	Redundancy	R_{sym} (%)	$\langle I/\sigma(I)\rangle$	R_{iso}
Native 1	2.80	26397	9632	83.7	2.8	4.7	15.5	na
Native 2	2.20	151034	22920	99.9	6.6	6.8	8.9	na
Deriv. 1†	4.20	7270	2377	71.0	3.1	12.9	5.4	18.3
Deriv. 2‡	4.50	3905	1806	67.4	2.2	11.7	6.2	27.1
Deriv. 3§	3.00	25952	8938	97.9	2.9	8.8	11.2	28.4
SeMet 1	3.20	20275	7401	97.6	2.7	6.9	10.3	16.7
SeMet 2	2.60	31765	12499	89.8	2.5	5.2	13.5	13.8

† $Pt(NH_4)(NO_2)_4$. ‡ $K_2Pt(SCN)_4$. § K_2PtCl_4.

3. Results and discussion

The crystal structure of the SfaE chaperone involved in biogenesis of S pili was solved by the multiple isomorphous replacement method and refined against data to 2.2 Å Bragg spacing (Tables 1 and 2). The SfaE crystals contain two closely associated SfaE protomers in the asymmetric unit. Each SfaE protomer has the characteristic overall L-shape of periplasmic chaperones, with a pair of Ig-like domains approximately perpendicular to each other (Fig. 1). The primary structure of

Table 2
Refinement statistics.

Resolution (Å)	15.0–2.2
No. of reflections (work)	20174
No. of reflections (free)	2222
R_{work} (%)	19.0
R_{free} (%)	24.4
No. of atoms	
Protein	2883
Water	103
B factor	
Protein (chain *A*) (Å^2)	28.9
Protein (chain *B*) (Å^2)	30.3
Water (Å^2)	31.3
R.m.s. deviations	
Bond lengths (Å)	0.014
Bond-angle distances (Å)	0.037
Ramachandran plot	
Residues in core region (%)	93.6
Residues in additional allowed regions (%)	6.4

Table 3
Comparison of the backbone of SfaE with FimC and PapD.

Root-mean-square deviation (r.m.s.d.) (Å) for C^α positions within 3.8 Å after superpositioning. The number in parentheses is the number of superimposed C^α atoms. Suffixes indicate individual domains used for superpositioning, with N for N-terminal domain and C for C-terminal domain.

	SfaE B	SfaE B_N	SfaE B_C	$FimC_N$	$FimC_C$	$PapD_N$	$PapD_C$
SfaE A	1.108 (182)	—	—	—	—	—	—
SfaE A_N	—	0.586 (103)	—	0.981 (104)	—	1.155 (103)	—
SfaE A_C	—	—	0.431 (79)	—	1.408 (83)	—	1.436 (71)
SfaE B_N	—	—	—	1.071 (104)	—	1.283 (103)	—
SfaE B_C	—	—	—	—	0.771 (81)	—	1.399 (69)

SfaE is quite similar to that of FimC (67% identity), but significantly different from PapD (32% identity). Nevertheless, the backbone structure of SfaE is very similar to that of FimC and PapD (Table 3). The N-terminal domain, which has a topology similar to Ig variable domains, is a β-barrel consisting of seven β-strands (labelled *A*1–*G*1) organized in two sheets with four and three strands each (Figs. 1*a* and 1*c*). The disulfide bond linking strands *B* and *F* in Ig variable domains is missing in SfaE as well as in other periplasmic chaperones. Instead, there is a pair of strand switches distinguishing the chaperone N-terminal domains from Ig variable domains. In SfaE, the *A*1 β-strand, which is initially hydrogen bonded to the *B*1 strand, moves towards the *F*1 strand starting from position 5 and becomes hydrogen bonded to it from position 7. Similarly, the *D*1 strand, which is initially hydrogen bonded to the *E*1 strand, makes a sharp turn at position 52 and becomes hydrogen bonded to the *C*1 strand at position 53. As in both FimC and PapD, residue Pro52 is a *cis*-proline, a feature no doubt required for the sharp change in direction made by the *D*1 β-strand at this point. The switching of strands *A*1 and *D*1 serves to link the two β-sheets making up the domain and probably plays a role in imparting structural stability to the domain in the absence of inter-sheet disulfide bonds. The C-terminal domains of periplasmic chaperones are more variable in structure, but all of them bear an overall similarity with Ig constant domains (Figs. 1*b* and 1*c*; Holmgren & Branden, 1989; Holmgren *et al.*, 1992). A major difference between the C-terminal domains of SfaE and PapD is in the connection between the *C*2 and *D*2 β-strands. In PapD, this connection is made *via* a short α-helix which packs onto the *C*2*D*2*F*2*G*2 sheet and which is involved in interactions with PapK in the PapD–PapK crystal structure. In SfaE, as well as in FimC, the connection is instead a β-turn with no intervening helix, effectively widening the subunit-binding cleft in these two chaperones compared with PapD.

There are three disordered regions in both of the SfaE chains. There is essentially no density for the *F*1–*G*1 loop, the *A*2–*B*2 loop or the *E*2–*F*2 loop. In the NMR structure of FimC, the *F*1–*G*1 loop is exceptionally mobile (Pellecchia *et al.*, 1998). In the 2.0 Å structure of PapD (Holmgren, 1993) and in crystals of R8A and Q108C PapD (Hung, Pinkner *et al.*, 1999), the *F*1–*G*1 loop either has very high temperature factors or is not seen at all. The loop becomes much more ordered on binding of pilus subunits or subunit-derived peptides in the subunit-binding cleft (Choudhury *et al.*, 1999; Kuehn *et al.*, 1993; Sauer *et al.*, 1999; Soto *et al.*, 1998). Flexibility of this loop appears to be a general feature of periplasmic chaperones that might be required for efficient binding to all of the different pilus subunits during pilus assembly. In addition, ordering of the *F*1–*G*1 loop on binding provides a negative entropy component that might be necessary to avoid too tight chaperone–subunit complexation and to allow uncapping of the subunit during pilus assembly.

As was previously found for PapD in crystals of both wild-type and R8A PapD (Hung, Pinkner *et al.*, 1999), the conserved surfaces in the subunit-binding cleft of each protomer pack against each other to form a dimer with an extended 1000 Å^2 interface, although the packing is very different in PapD and SfaE dimers. In the PapD dimer, which exhibits near-perfect twofold symmetry, the two protomers pack with the *F*1–*G*1 loop of each protomer in the subunit-binding cleft of its neighbour and with a short stretch of β-sheet hydrogen bonding between the two antiparallel *G*1 donor strands. The SfaE dimer is formed mainly by packing of the N-terminal domain of one protomer (chain *A*) in the subunit-binding cleft of the second protomer (chain *B*) and does not posses proper symmetry (Fig. 2*a*). The *A*1*B*1*E*1 sheet of chain *A* packs against the *C*2*D*2*F*2*G*2 sheet on one side of the subunit-binding cleft of chain *B*. On the other side of the cleft, the *G*1 donor strand of chain *A* crosses the *A*1 and *G*1 strands of chain *B* more or less at right angles.

Interestingly, Phe105*A*, which is part of the conserved *G*1 donor-strand motif of periplasmic chaperones (Choudhury *et al.*, 1999; Hung *et al.*, 1996; Sauer *et al.*, 1999), is intercalated between these two strands and partly penetrates the hydrophobic core of chain *B* (Fig. 2*a*). The phenylalanine side chain is deeply buried in a hydrophobic pocket formed by Val2*B*, Leu4*B*, Val22*B*, Ile33*B*, Val87*B*, Ala89*B* and Ile107*B* (Fig. 2*b*). The *A*1 and *G*1 strands of chain *B* have moved apart by

approximately 1 Å compared with chain *A* in order to accommodate the phenylalanine side chain. The main chain on each side of Phe105*A* hydrogen bonds with the *A*1 and *G*1 strands of chain *B*. Ile107*A*, which is also part of the donor-strand motif in SfaE, is buried between the two protomers of the dimer but makes much less extensive contacts with the second SfaE chain compared with Phe105*A*.

In the FimC–FimH complex, residues Ile103, Leu105 and Ile107 of FimC are deeply embedded within the acceptor cleft of FimH and form an integral part of the hydrophobic core of the FimH pilin domain. In the SfaE dimer structure, two of the donor-strand residues, Phe105*A* and Ile107*A*, appear to interact with the second SfaE protomer in a structural mimic of the donor-strand complementation between chaperone and subunit observed in the FimC–FimH and PapD–PapK crystal structures. In these complexes, the *G*1 strand is longer compared with the free chaperones and is hydrogen bonded both to the first (*A*) and last (*F*) strand of the pilin domain. In FimC bound to FimH, the *G*1 strand starts at position 101 and ends at position 112. Comparison of the *G*1 strand in SfaE chain *A* with the *G*1 strand of FimC bound in the acceptor cleft of FimH shows very similar conformations starting from position 105. In contrast, the two first residues of the *G*1 strand that are visible in SfaE chain *A* (Leu103 and Gln104) have very different conformations compared with those in FimC. Hydrogen bonding between the *F*1 and *G*1 strand in SfaE is interrupted at position 91 in the F1 strand, which is a proline in SfaE as well as in all other known sequences of periplasmic chaperones. This defines the starting point of the flexible region of the *F*1–*G*1 connection in SfaE and probably in all other periplasmic chaperones. Phe105 in the *G*1 strand is opposite Pro91 and defines the end of the flexible region. Thus, the conformation of the first two hydrophobic residues in the *G*1 donor-strand motif (positions 103 and 105 in SfaE) is variable and can be adjusted to fit into the acceptor cleft of various pilus subunits. The last residue of the motif (position 107) has a more fixed conformation, perhaps reflecting greater similarity in the regions of the pilin domains with which this residue interacts in chaperone–pilin complexes.

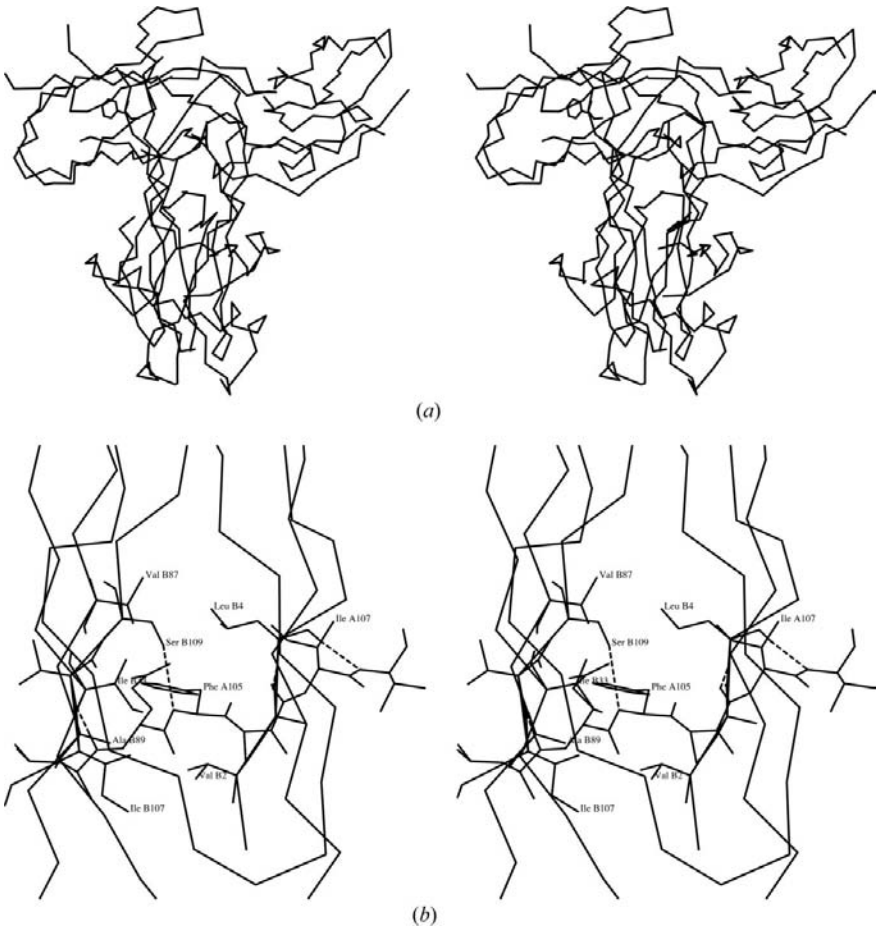

Figure 2
(*a*) Stereo diagram showing packing of the two SfaE molecules in the asymmetric unit. SfaE chain *A* is on top, with the tip of the N-terminal domain inserted into the subunit-binding cleft of chain *B*. Also shown is the side chain of Phe105*A* in the *G*1 strand of chain *A* which is intercalated between the *F*1 and *G*1 strands of chain *B*. (*b*) Close-up of the region around Phe105*A*. The phenylalanine side chain is inserted in a hydrophobic pocket formed by Val2*B*, Leu4*B*, Val22*B*, Ile33*B*, Val87*B*, Ala89*B* and Ile107*B*. In addition, four hydrogen bonds are formed between the *G*1 strand in chain *A* and the *A*1 and *G*1 strands of chain *B*: Val108*A* N–Leu4*B* O, Ala106*A* O–Leu4*B* N, Phe105*A* N–Ser109*B* O^{γ} and Leu103*A* O–Val108*B* N.

Mainly owing to the intercalation of Phe105*A* between the *A*1 and *G*1 strands of chain *B*, the two SfaE molecules in the asymmetric unit were sufficiently different for refinement with non-crystallographic symmetry restraints to perform significantly worse than unrestrained refinement as judged by the $R_{\rm free}$ value (data not shown). Nevertheless, except for the differences arising from packing interactions in the asymmetric unit dimer, the two copies of SfaE are essentially identical (Table 3). One exception is the difference in conformation of the last strand in the C-terminal domain between the two polypeptide chains in the dimer (Fig. 3). In chain *A*, the *G*2 strand is hydrogen bonded to the *F*2 strand until position 203, after which the *G*2 strand moves slightly away from the *F*2 strand. The *G*2 strand in chain *B* instead undergoes a strand switch starting at position 202 and becomes hydrogen bonded to the *A*2 strand for the last three residues of the chain, *i.e.* positions 204–206. The difference in conformation of the *G*2 strand would appear to arise from crystal packing, as the side chain of Gln74 from a symmetry-related molecule is inserted between the *F*2 and *G*2 strands of chain *B* with hydrogen bonds from the side chain to the main-chain carbonyl O atom of Gly203*B* and to the peptide N atom of Ile186*B*. However, an almost identical switch occurs in both the crystal structure of FimC bound to FimH (Choudhury *et al.*, 1999) and in the NMR

structure of free FimC (Pellecchia *et al.*, 1998). In PapD, the additional *H*2 strand at the C-terminal end of the protein is inserted between the *A*2 and *G*2 strands. The three final residues of the *G*2 strand in both SfaE chain *B* and FimC occupy essentially the same position as the three first residues (211–213) of PapD's *H*2 strand. Perhaps owing to the absence of an *H*2 strand in SfaE, the C-terminal domain appears to have a degenerate fold, permitting two alternative positions for its C-terminal strand.

The two Ig-like domains of SfaE are linked *via* an irregular six-residue linker (residues 118–123). The domain interface consists of three rows of residues in the *A*1′′, *F*1 and *G*1 strands of the N-terminal domain and of residues in the *C*2 and *D*2 strands as well as in the *F*2–*G*2 turn of the C-terminal domain. As in all PapD-like chaperones, the domain interface buries a salt-bridge system formed by the side chains of three invariant residues (Glu83, Arg116 and Asp196 in PapD; Holmgren *et al.*, 1992; Hung, Knight *et al.*, 1999; Hung *et al.*, 1996). The corresponding residues in SfaE are Glu80, Arg116 and Asp193. Two of these residues, Glu80 and Arg116, constitute the bottom row of the N-terminal domain interface.

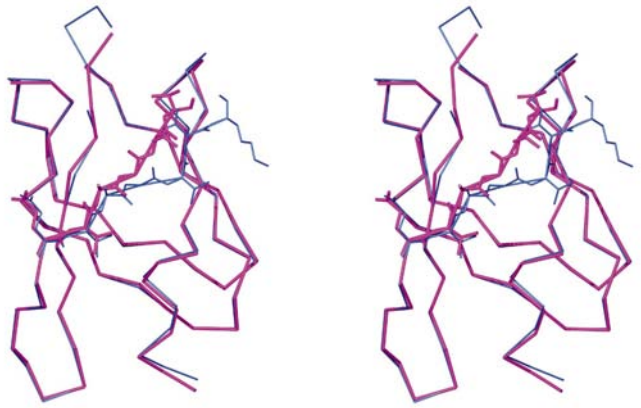

Figure 3
Stereo C^{α} trace of the C-terminal domain of chain *B* (blue) superimposed on the C-terminal domain of chain *A* (magenta) showing the different conformation of the *G*2 strand in the two domains. Side chains are shown for residues in the *G*2 strand.

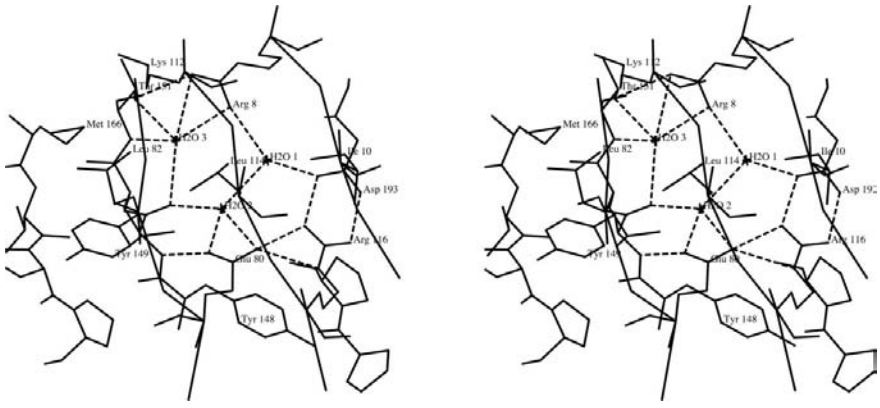

Figure 4
Close-up of the inter-domain interface of SfaE.

The middle row consists of three invariantly hydrophobic residues (Ile10, Leu82 and Leu114 in SfaE). The top row consists of the two invariant residues Arg8 and Lys112 which define the base of the subunit-binding cleft and which are both critical for subunit binding and chaperone function.

Alignment of SfaE with PapD and FimC from the FimC–FimH structure shows that the buried salt-bridge residues have very similar conformations and interact in the same way in all three chaperones. In SfaE, the salt-bridge system is linked to Arg8 in the subunit-binding cleft *via* hydrogen bonds to two buried inter-domain water molecules (Fig. 4). Interestingly, these two water molecules are structurally conserved in the 2.0 Å structure of PapD (Holmgren, 1993) as well as in a 2.5 Å structure of free FimC (S. D. Knight, unpublished data). The inter-domain hydrogen-bond network is further extended in SfaE by hydrogen bonds involving the side chain of Thr151, main-chain atoms in the region from residue 148–151 in the C-terminal domain and a third inter-domain water molecule. The atoms participating in this buried hydrogen-bond network form a cusp-shaped surface onto which two of the side chains (Ile10 and Leu114) in the middle row of the N-terminal domain interface pack (Fig. 4). The third residue in this row, Leu82, packs against Tyr149 and Met166 in the C-terminal domain. The relative orientation between domains in the *A* chain of SfaE compared with the *B* chain is different by about 10°. The residues and the water molecules forming the inter-domain hydrogen-bond network behave as part of the C-terminal domain. The difference in orientation is mostly owing to rotation centred around Leu114, which packs with its side chain against the central region of the cusp-shaped hydrogen-bond network.

Given that the N-terminal domain interface contains a row of three large hydrophobic residues whereas the C-terminal domain interface is much more hydrophilic, the finding that the C-terminal but not the N-terminal domain of FimC could be expressed as a stable protein on its own (Hermanns *et al.*, 2000) is perhaps not so surprising. The very close interaction between chaperone and pilus subunits seen in the FimC–FimH and PapD–PapK structures, with the *G*1 donor strand of the chaperone becoming essentially part of the subunit fold, suggests that the two proteins will have to at least partially unfold in order to dissociate. This in turn suggests the intriguing possibility that subunit release during pilus assembly might involve partial unfolding of the N-terminal domain owing to destabilization of the domain interface. Site-directed mutagenesis studies using the PapD chaperone have shown that the Glu-Arg-Asp triad in the domain interface is important for chaperone stability (Hung, Knight *et al.*, 1999). Interestingly, mutation of the aspartic acid to asparagine had a relatively limited effect on chaperone stability, but blocked pilus assembly at a stage after interaction of chaperone–

subunit complex with the outer membrane usher; it was suggested that the aspartate side chain might be involved in displacing the pilus subunit C-terminus from Arg8 to promote release of subunit. This would require a conformational change in both Arg8 and the aspartate side chain, which would in turn cause disruption of the domain-interface hydrogen-bond network and possibly result in partial unfolding of the N-terminal domain to further promote subunit release. Such a model for subunit release is consistent with recent NMR studies (Bann *et al.*, 2002) suggesting that folding of the N-terminal domain of PapD is dependent on the folding of the C-terminal domain.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.

Bann, J. G., Pinkner, J., Hultgren, S. J. & Frieden, C. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 709–714.

Brünger, A. T. (1993). *X-PLOR Manual, Version* 3.1. *A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT, USA.

Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J. & Knight, S. D. (1999). *Science*, **285**, 1061–1066.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Hacker, J. (1992). *Can. J. Microbiol.* **38**, 720–727.

Hermanns, U., Sebbel, P., Eggli, V. & Glockshuber, R. (2000). *Biochemistry*, **39**, 11564–11570.

Holmgren, A. (1993). Department of Molecular Biology, Swedish University of Agricultural Sciences, Uppsala.

Holmgren, A. & Branden, C. I. (1989). *Nature (London)*, **342**, 248–251.

Holmgren, A., Kuehn, M. J., Branden, C. I. & Hultgren, S. J. (1992). *EMBO J.* **11**, 1617–1622.

Hung, D. L., Knight, S. D. & Hultgren, S. J. (1999). *Mol. Microbiol.* **31**, 773–783.

Hung, D. L., Knight, S. D., Woods, R. M., Pinkner, J. S. & Hultgren, S. J. (1996). *EMBO J.* **15**, 3792–3805.

Hung, D. L., Pinkner, J. S., Knight, S. D. & Hultgren, S. J. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 8178–8183.

Jones, C. H., Dodson, K. & Hultgren, S. J. (1996). *Urinary Tract Infection: Molecular Pathogenesis to Clinical Management*, edited by J. W. Warren, pp. 175–219. Washington DC: ASM.

Jones, C. H., Pinkner, J. S., Nicholes, A. V., Slonim, L. N., Abraham, S. N. & Hultgren, S. J. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 8397–8401.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Khan, A. S., Muhldorfer, I., Demuth, V., Wallner, U., Korhonen, T. K. & Hacker, J. (2000). *Mol. Gen. Genet.* **263**, 96–105.

Knight, S. D., Berglund, J. & Choudhury, D. (2000). *Curr. Opin. Chem. Biol.* **4**, 653–660.

Korhonen, T. K., Parkkinen, J., Hacker, J., Finne, J., Pere, A., Rhen, M. & Holthofer, H. (1986). *Infect. Immun.* **54**, 322–327.

Korhonen, T. K., Vaisanen-Rhen, V., Rhen, M., Pere, A., Parkkinen, J. & Finne, J. (1984). *J. Bacteriol.* **159**, 762–766.

Kuehn, M. J., Ogg, D. J., Kihlberg, J., Slonim, L. N., Flemmer, K., Bergfors, T. & Hultgren, S. J. (1993). *Science*, **262**, 1234–1241.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Morschhauser, J., Vetter, V., Korhonen, T., Uhlin, B. E. & Hacker, J. (1993). *Zentralbl. Bakteriol.* **278**, 165–176.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pellecchia, M., Guntert, P., Glockshuber, R. & Wuthrich, K. (1998). *Nature Struct. Biol.* **5**, 885–890.

Prasadarao, N. V., Wass, C. A., Hacker, J., Jann, K. & Kim, K. S. (1993). *J. Biol. Chem.* **268**, 10356–10363.

Prasadarao, N. V., Wass, C. A. & Kim, K. S. (1997). *Infect. Immun.* **65**, 2852–2860.

Saren, A., Virkola, R., Hacker, J. & Korhonen, T. K. (1999). *Infect. Immun.* **67**, 2671–2676.

Sauer, F. G., Futterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J. & Waksman, G. (1999). *Science*, **285**, 1058–1061.

Sauer, F. G., Knight, S. D., Waksman, G. & Hultgren, S. J. (2000). *Semin. Cell Dev. Biol.* **11**, 27–34.

Sauer, F. G., Mulvey, M. A., Schilling, J. D., Martinez, J. J. & Hultgren, S. J. (2000). *Curr. Opin. Microbiol.* **3**, 65–72.

Slonim, L. N., Pinkner, J. S., Branden, C. I. & Hultgren, S. J. (1992). *EMBO J.* **11**, 4747–4756.

Soto, G. E., Dodson, K. W., Ogg, D., Liu, C., Heuser, J., Knight, S., Kihlberg, J., Jones, C. H. & Hultgren, S. J. (1998). *EMBO J.* **17**, 6155–6167.

Virkola, R., Parkkinen, J., Hacker, J. & Korhonen, T. K. (1993). *Infect. Immun.* **61**, 4480–4484.

Virkola, R., Westerlund, B., Holthofer, H., Parkkinen, J., Kekomaki, M. & Korhonen, T. K. (1988). *Infect. Immun.* **56**, 2615–2622.