Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The F4 fimbrial chaperone FaeE is stable as a monomer that does not require self-capping of its pilin-interactive surfaces

Inge Van Molle,[a,b]* Kristof Moonens,[a,b] Lieven Buts,[a,b] Abel Garcia-Pino,[a,b] Santosh Panjikar,[c] Lode Wyns,[a,b] Henri De Greve[a,b] and Julie Bouckaert[a,b]

[a]Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium, [b]Structural Biology Brussels, Department of Molecular and Cellular Interactions, VIB, Pleinlaan 2, 1050 Brussels, Belgium, and [c]EMBL Hamburg c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

Correspondence e-mail: ivmolle@vub.ac.be

Received 11 December 2008
Accepted 12 February 2009

PDB References: FaeE, 3f65, r3f65sf; 3f6i, r3f6isf; 3f6l, r3f6lsf.

Many Gram-negative bacteria use the chaperone–usher pathway to express adhesive surface structures, such as fimbriae, in order to mediate attachment to host cells. Periplasmic chaperones are required to shuttle fimbrial subunits or pilins through the periplasmic space in an assembly-competent form. The chaperones cap the hydrophobic surface of the pilins through a donor-strand complementation mechanism. FaeE is the periplasmic chaperone required for the assembly of the F4 fimbriae of enterotoxigenic *Escherichia coli*. The FaeE crystal structure shows a dimer formed by interaction between the pilin-binding interfaces of the two monomers. Dimerization and tetramerization have been observed previously in crystal structures of fimbrial chaperones and have been suggested to serve as a self-capping mechanism that protects the pilin-interactive surfaces in solution in the absence of the pilins. However, thermodynamic and biochemical data show that FaeE occurs as a stable monomer in solution. Other lines of evidence indicate that self-capping of the pilin-interactive interfaces is not a mechanism that is conservedly applied by all periplasmic chaperones, but is rather a case-specific solution to cap aggregation-prone surfaces.

1. Introduction

The attachment of pathogenic bacteria to host cells is a key event in their infection process and is typically mediated by adhesins. These are often located on the bacterial surface in polymeric proteinaceous appendages called fimbriae or pili (Soto & Hultgren, 1999). It is a major challenge for Gram-negative bacteria to generate these organelles on their surface, a process that requires protein synthesis, folding, translocation over two membranes and across the periplasmic space and self-assembly into the final polymeric structure. Many Gram-negative bacteria use the chaperone–usher pathway to tackle this problem (Thanassi *et al.*, 1998), employing two proteins that will not be part of the final structure: a periplasmic chaperone and an outer membrane gatekeeper known as the usher. Once the fimbrial subunits have been translocated over the inner membrane by the general secretory pathway, they are bound by a cognate chaperone and guided to the usher, the pilus-assembly platform in the outer membrane. Here, they are assembled into the final adhesive structure and transported across the outer membrane.

The crystal structures of the periplasmic chaperones PapD (uropathogenic *Escherichia coli* P pili; Holmgren & Brändén, 1989; Hung, Pinkner *et al.*, 1999), FimC (uropathogenic *E. coli* type 1 pili; Pellecchia *et al.*, 1998), SfaE (uropathogenic *E. coli* S pili; Knight *et al.*, 2002), Caf1M (*Yersinia pestis* F1 antigen; Zavialov *et al.*, 2003) and SafB (*Salmonella enterica* Saf fimbriae (Remaut *et al.*, 2006) show two immunoglobulin (Ig) domains oriented at a right angle in a boomerang shape. Genetic studies have revealed several conserved residues that occur throughout the family of PapD-like chaperones and are important for their structure and function (Hung, Knight *et al.*, 1999). Those conserved residues can mainly be found in the cleft between the two Ig domains.

Based on the length of the loop between the *F*1 and *G*1 strands, periplasmic chaperones have been classified as FGL (long loop between Ig-fold β-strands *F* and *G*) or FGS (short *FG* loop) chaperones (Hung *et al.*, 1996). Historically, FGS chaperones were believed to be involved in the biosynthesis of fimbriae with a complex subunit organization, presenting one adhesin at the tip, while FGL chaperones assisted in the assembly of polyadhesive structures consisting of mainly one subunit. Following this classification, Zavialov and coworkers assigned the F4 and the related F5 fimbriae as an intermediate family of adhesive fibres: FGS chaperone-assembled polyadhesive fibres (Zavialov *et al.*, 2007). The assembly of these fimbriae is assisted by FGS chaperones, but although the fimbrial operon contains ten genes the fimbrial structure is dominantly composed of one major subunit that also determines the adhesive properties of the fimbriae. More recent phylogenetic analysis based on usher ancestry revealed that only the FGL organelles form a monophyletic group, while the FGS structures can be divided into different clades (Nuccio & Baumler, 2007). Following this classification, F4 and F5 fimbriae belong to the family of κ fimbriae.

The crystal structures of chaperone–pilin complexes show that pilins have an incomplete Ig fold that lacks the seventh β-strand. Upon translocation across the inner membrane, the C-terminal carboxylate of the fimbrial subunits is bound by the conserved Arg8 and Lys112 residues in the cleft of the chaperone. Binding by the chaperone facilitates the release of the pilins from the inner membrane and provides them with a folding template. The pilins fold directly on the surface of the chaperone through a β-zippering mechanism along the *G*1 strand of the chaperone. By β-strand pairing of the *G*1 strand of the chaperone with the *F* strand of the pilin and the insertion of a conserved pattern of hydrophobic residues into the groove on the surface of the pilin, the chaperone complements the fold of the pilin (Jones *et al.*, 1997). In the case of the FGL chaperones, in addition to the *G*1 strand of the chaperone, the *A*1 strand is also involved in the donor-strand complementation mechanism. The *A*1 strand interacts with the *A* strand of the pilin (Zavialov *et al.*, 2003). The chaperone thus protects the pilins from aggregation, misfolding and proteolytic degradation. Periplasmic chaperones are essential for pilus assembly since they catalyze the folding of the pilins (Vetsch *et al.*, 2004) and keep them in an assembly-competent conformation (Sauer *et al.*, 2002; Zavialov *et al.*, 2003).

In pilin-free chaperones, the hydrophobic residues of the pilin-binding motif are surface-exposed. As several crystal structures of periplasmic chaperones have revealed oligomers with capped pilin-binding interfaces, a self-capping mechanism by which the chaperones also protect their pilin-binding surfaces in solution has been suggested (Hung, Pinkner *et al.*, 1999; Knight *et al.*, 2002; Zavialov & Knight, 2007). The *Y. pestis* Caf1M chaperone indeed forms tetramers in solution that are created by tight packing of the subunit binding surfaces into a hetero-sandwich. This self-capping mechanism protects the chaperone from proteolytic degradation and aggregation (Zavialov & Knight, 2007). Also, expression of a Q108C mutant of PapD without the P-pilus subunits resulted in the presence of up to 30% disulfide-linked PapDQ108C dimers in the periplasm, showing dimer formation *in vivo* (Hung, Pinkner *et al.*, 1999).

Here, we describe the crystal structure of the F4 fimbrial chaperone FaeE and apply differential scanning calorimetry, chemical cross-linking and dynamic light scattering to explore the self-capping mechanism of FaeE in solution. Additionally, we try to answer the question whether self-capping is a general mechanism applied by periplasmic chaperones or is a case-specific solution to protect aggregation-prone surfaces.

2. Material and methods

2.1. Protein expression, purification and crystallization

The *faeE* or *faeE-faeG* genes (plasmid pHD163 or pHD147, respectively) from the F4ad$^+$ *E. coli* strain C1360-79 were expressed in *E. coli* C43 (DE3) under the T7 promoter. FaeE and FaeE–FaeG were purified by ion-exchange chromatography in 20 m*M* TES pH 7.0 and elution with an NaCl gradient in the same buffer and were finally used for crystallization in 20 m*M* TES pH 7.0, 150 m*M* NaCl, all as described previously (Van Molle *et al.*, 2005). FaeE crystals were grown in 0.1 *M* Tris pH 7.5, 50%(*v*/*v*) 2-methyl-2,4-pentanediol (MPD), 0.2 *M* $NH_4H_2PO_4$. FaeE crystallized in three crystal forms, all of which belonged to space group *C*2. Crystallization drops set up with FaeE at 20 mg ml^{-1} were found to contain crystals of both form 1 (unit-cell parameters $a = 195.7$, $b = 78.5$, $c = 184.6$ Å, $\beta = 102.2^\circ$) and form 2 (unit-cell parameters $a = 136.4$, $b = 75.7$, $c = 69.4$ Å, $\beta = 92.8^\circ$). $FaeE_{SeMet}$ crystals (set up at 25 mg ml^{-1}) belonged to crystal form 2. Crystals of form 3 (unit-cell parameters $a = 109.7$, $b = 78.6$, $c = 87.8$ Å, $\beta = 96.4^\circ$) were found in drops originally set up with the FaeE–FaeG complex at 15 mg ml^{-1}. In these drops only FaeE crystallized, while FaeG polymerized and precipitated (Van Molle *et al.*, 2005).

2.2. X-ray data collection and structure determination

X-ray data were collected on the ESRF beamline ID14-1 (Grenoble, France) for crystal form 1 and the DESY/EMBL beamline BW7A (Hamburg, Germany) for crystal forms 2 and 3 (Van Molle *et al.*, 2005). An initial partial FaeE model was obtained by phase determination using a three-wavelength MAD data set from the $FaeE_{SeMet}$ crystal form 2 (Van Molle *et*

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Crystal form 1	Crystal form 2	Crystal form 3
Data collection			
Space group	*C*2	*C*2	*C*2
Unit-cell parameters			
a (Å)	195.7	136.4	109.7
b (Å)	78.5	75.7	78.6
c (Å)	184.6	69.4	87.8
β (°)	102.2	92.8	96.4
Resolution range (Å)	49.8–2.3 (2.4–2.3)	39.0–2.7 (2.8–2.7)	39.3–2.8 (2.9–2.8)
Redundancy	5.2 (4.0)	7.5 (7.4)	6.1 (6.2)
$\langle I/\sigma(I)\rangle$	14.56 (3.7)	19.3 (5.3)	14.91 (7.0)
Completeness (%)	99.5 (95.6)	98.0 (97.1)	99.8 (100.0)
$R_{\rm merge}$† (%)	8.3 (39.0)	7.1 (32.9)	12.2 (30.2)
Refinement			
No. of unique reflections	122580	17123	18291
$R_{\rm work}$	0.2238	0.235	0.255
$R_{\rm free}$ (5% test set)	0.2596	0.296	0.303
No. of protein atoms	11097	2799	3049
No. of solvent atoms	328	13	32
Average *B* factor (Å^2)	64.49	65.674	49.259
R.m.s.d. stereochemistry			
Bonds (Å)	0.004	0.005	0.003
Angles (°)	0.653	1.016	0.805
Ramachandran plot			
Residues in allowed regions (%)	98.5	98.1	97.3
Residues in disallowed regions (%)	1.5	1.9	2.7
PDB code	3f65	3f6i	3f6l

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity for symmetry-related reflections.

al., 2005) using the *Auto-Rickshaw* software pipeline (Panjikar *et al.*, 2005). Substructure determination, phase calculation, solvent flattening and model building were performed automatically using *SHELXD* (Sheldrick, 2008), *MLPHARE* (Collaborative Computational Project, Number 4, 1994), *DM* (Cowtan & Zhang, 1999) and *ARP*/*wARP* (Perrakis *et al.*, 2001), respectively, within *Auto-Rickshaw* (Panjikar *et al.*, 2005). This initial model was further completed by manual model building in *Coot* (Emsley & Cowtan, 2004) and was refined against the absorption-edge data to 2.7 Å resolution (Table 1) using *REFMAC*5 (Murshudov *et al.*, 1997) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) and *phenix.refine* (Adams *et al.*, 2002) using TLS restraints. Chain *A* of this model was used to search for molecular-replacement solutions of the FaeE structures for the data sets of crystal forms 1 and 3 using *Phaser* from the *CCP*4 suite. Molecular-replacement solutions containing eight or two monomers in the asymmetric unit were searched for using a 3.5 Å resolution cutoff for crystal forms 1 and 3, respectively. The models obtained were improved by iterative manual model building in *Coot* and refinement using *REFMAC*5 and *phenix.refine* using TLS restraints to 2.3 and 2.8 Å resolution for crystal forms 1 and 3, respectively.

2.3. Analysis of packing interfaces in the crystal structures of FaeE, PapD, SfaE and Caf1M

The FaeE structures were validated using *MolProbity* (Table 1; Davis *et al.*, 2007) and the coordinates for the structures of FaeE in crystal forms 1, 2 and 3 were submitted to the Protein Data Bank (PDB codes 3f65, 3f6i and 3f6l, respectively). The root-mean-square deviations (r.m.s.d.s) between the FaeE structures in crystal forms 1, 2 and 3 were calculated using the *LSQKAB* program from the *CCP*4 suite (Kabsch, 1976). The interfaces present in the crystal structures of FaeE (PDB codes 3f65, 3f6i and 3f6l), PapD (1qpp), SfaE (1l4i) and Caf1M (2os7) were analyzed using the *PISA* program from the EMBL–EBI website (Krissinel & Henrick, 2007).

2.4. Differential scanning calorimetry (DSC)

All calorimetric measurements were performed on a MicroCal VP-DSC differential scanning microcalorimeter with a 0.515 ml cell and data were analyzed using the program *MicroCal Origin DSC* 7.0. All samples were filtered using 0.45 µm Minisart filters (Sartorius) and degassed for 10 min using a vacuum pump. DSC scans followed heating of the samples from 293 to 353 K. Scan rates of 90, 75, 60, 30 and 15 K h^{-1} were applied to assess the scan-rate dependence of the transition temperature ($T_{\rm m}$). To evaluate the pH-dependence of $T_{\rm m}$, DSC scans were recorded in 20 m*M* sodium acetate pH 5.2, 20 m*M* phosphate pH 7.5 and 20 m*M* *N*-cyclohexyl-3-aminopropanesulfonic acid (CAPS) pH 10.5 buffer solutions.

DSC thermograms of the denaturation of FaeE were analyzed assuming a two-state reversible transition,

$$\mathrm{N} \overset{K}{\longleftrightarrow} \mathrm{D}, \qquad K = \frac{\alpha}{(1-\alpha)} = \exp\left(\frac{-\Delta G_T^\circ}{RT}\right),$$

where K is the equilibrium denaturation constant, R is the universal gas constant and α is the fraction of protein in the denatured state.

The model function for the DSC thermograms ($C_p - C_{p,N}^\circ$) can be obtained from the first partial derivative of the enthalpy of the denaturation function,

$$C_p - C_{p,N}^\circ = \alpha\Delta C_p^\circ + \frac{\alpha(1-\alpha)}{RT^2}[\Delta H_{\rm v} + \Delta C_p^\circ(T - T_{1/2})]^2,$$

where $\Delta H_{\rm v}$ is the van't Hoff enthalpy of denaturation at the transition temperature ($T_{1/2}$), α is the fraction of denatured protein and ΔC_p° is the corresponding heat-capacity change, which is assumed to be independent of the temperature. The thermodynamic parameters that characterize the melting curves ($\Delta H_{\rm v}$, ΔC_p° and $T_{1/2}$) were obtained by fitting the model function for the DSC endotherms to the experimental temperature profiles using the *MicroCal Origin DSC* 7.0 software package.

2.5. Chemical cross-linking

FaeE at 0.75 mg ml^{-1} concentration was incubated with different concentrations of glutaraldehyde in 20 m*M* phosphate buffer pH 7.5 for 30 min at room temperature. The total reaction volume was 100 µl. The reaction was stopped by adding 10 µl 1 *M* Tris–HCl pH 8.0. Cross-linking products were analyzed on SDS–PAGE.

2.6. Dynamic light scattering (DLS)

Prior to DLS measurements, FaeE samples (20 mg ml^{-1}) were cleared from dust and air by centrifugation for 1 h at 277 K and 14 000 rev min^{-1} in a microcentrifuge. DLS measurements were conducted on a RiNA Laser-Spectroscatter 210 (Netzwerk RNA-Technologien GmbH, Germany). To assess the particle-size distribution in the sample, ten measurements of 10 s each were performed.

2.7. Prediction of β-aggregation-prone regions

The Caf1M (UniProtKB entry P26926), FaeE (P25401), PapD (P15319), FimC (P31697), SfaE (Q9EXJ6) and SafB (Q8ZRK3) amino-acid sequences were analyzed using the statistical mechanics algorithm *TANGO* (Fernandez-Escamilla *et al.*, 2004), with input parameters set at physiological conditions (pH 7, ionic strength 0.02 *M*). Protein stability was set at −42 kJ mol^{-1}.

3. Results and discussion

3.1. Structure of the chaperone FaeE

Although the structure of FaeE was solved in the three crystal forms previously reported (Van Molle *et al.*, 2005), we will mainly discuss the FaeE structure from crystal form 1, which was determined at the highest resolution. Details of data-collection and refinement statistics are listed in Table 1. Crystal form 1 contains eight molecules in the asymmetric unit. The asymmetric units of crystal forms 2 and 3 each contained two molecules. The FaeE structure from crystal form 1 has an r.m.s.d. with the FaeE structures of crystal forms 2 and 3 of 1.04 and 0.88 Å, respectively (199 and 204 C^{α} atoms superposed, respectively). FaeE shows the typical boomerang-shaped structure containing two immunoglobulin domains that has been observed for all periplasmic chaperones. In the first domain, strands *A*1′, *A*1″, *B*1, *C*1, *D*1′, *D*1″, *E*1, *F*1 and *G*1 are organized in two sheets, one containing strands *A*1′, *B*1, *D*1′ and *E*1 and the other containing strands *C*1, *D*1″, *F*1 and *G*1. The second domain also comprises two sheets: one containing strands *A*2, *B*2 and *D*2 and one containing strands *C*2, *E*2 and *F*2. A α-helix connects strands *C*2 and *D*2 (Fig. 1). PapD and FaeE are longer than the other FGS chaperones. In the structure of PapD, this results in an additional β-strand in the second domain, *H*2 (Holmgren *et al.*, 1992). In the FaeE structures, there is no clear electron density for residues 205–224 corresponding to strand *G*2 and possibly *H*2.

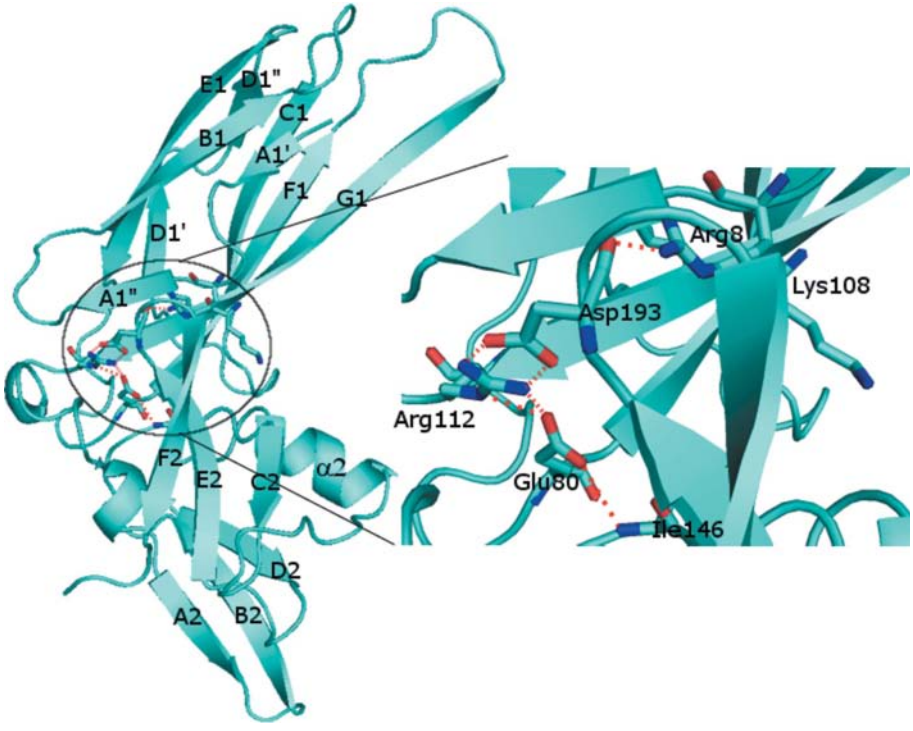

Figure 1
Ribbon diagram of the FaeE monomer (left), as present in all three crystal forms, highlighting its pilin-anchoring cleft (right). The conserved salt bridges between residues Glu80, Arg112 and Asp193 (Glu83, Arg116 and Asp196 in PapD nomenclature) and the hydrogen bonds formed between Glu80 and Ile146 and between Arg8 and Asp193 are indicated. Arg8 and Lys108 are the conserved subunit-anchoring residues. The chaperone *G*1 strand extends beyond its antiparallel β-strand interactions with the *F*1 strand because of crystal-packing interactions in the chaperone dimer (Fig. 2).

The structure of the FGS chaperones and particularly the relative orientation of the two domains is stabilized by salt bridges (Hung, Knight *et al.*, 1999). A salt-bridge network is also present in FaeE and is formed between Glu80, Arg112 and Asp193 (Fig. 1). The side chain of Glu80 forms an additional hydrogen bond to the main-chain N atom of Ile146.

As in the other FGS chaperones, the conserved residues Arg8 and Lys108 (Arg8 and Lys112 in the PapD nomenclature) are located at the bottom of the pilin-binding anchoring cleft. The Arg8 side chain points away from the cleft and reorients upon interaction with the pilin, allowing anchoring of the pilin by salt-bridge formation with the N-terminus of the pilin (Kuehn *et al.*, 1993). In the free FaeE chaperone, Arg8 makes a hydrogen bond to the main-chain O atom of Asp193 (Fig. 1).

The eight molecules in the asymmetric unit of crystal form 1 show a high diversity in loop visibility. Monomers *A*, *B*, *C* and *G* are complete up to residue 206, but the electron density is missing for a major part of the C-terminal

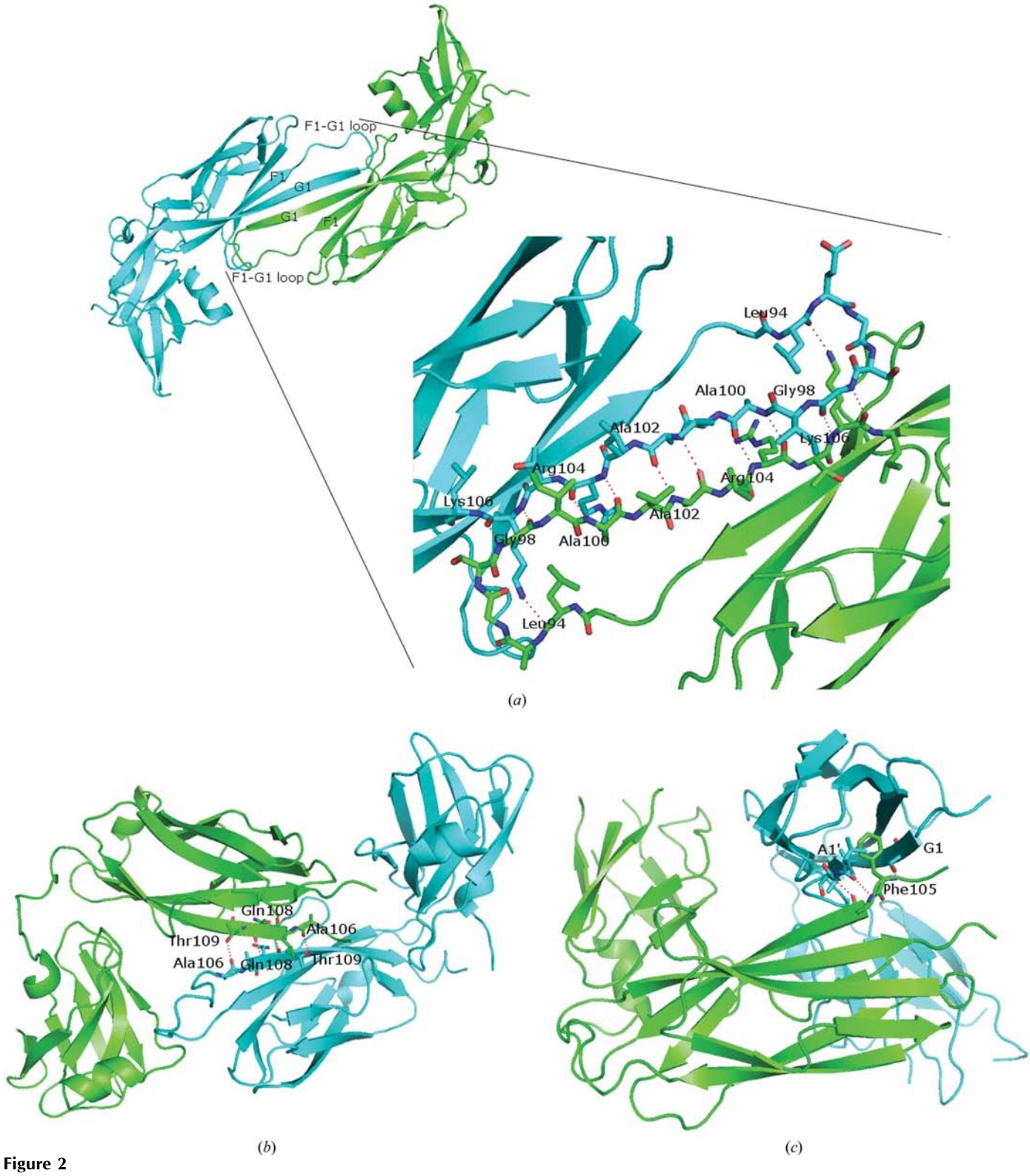

Figure 2
(*a*) Detailed view of the interface between the FaeE monomers in FaeE dimer 1 (left). The dimer is formed by extensive hydrogen bonding between the *G*1 strands of both monomers. An additional hydrogen bond is formed between Lys106 in the *G*1 strand of one monomer and Leu94 in the *F*1–*G*1 loop of the other monomer. The *F*1 and *G*1 strands and the *F*1–*G*1 loops of both monomers are indicated in the inset on the left; the hydrogen-bonding residues of the *G*1 strands and *F*1–*G*1 loops of both monomers are indicated in the enlarged view. (*b*) The PapD dimer (PDB code 1qpp) is formed around a noncrystallographic twofold axis, orienting the two monomers at an almost right angle. Hydrogen bonds are found between the Gln108 residues and the Thr109 and Ala106 residues of both monomers. (*c*) The SfaE dimer (PDB code 1l4i) is maintained by hydrophobic packing of the N-terminal domain of one monomer against the subunit-binding interface of the other monomer. The Phe105 side chain of one monomer (green) intercalates between the *A*1′ and *G*1 strands of the other monomer (cyan). Main-chain hydrogen bonds between Ala3 (cyan) and Ala103 (green) and between Ser109 (cyan) and Phe105 (green) are also indicated.

domain of monomers *D*, *F* and *H*. Monomer *E* has an intermediate completeness. Specifically, the loops between strands *A*2 and *B*2 and between *D*2 and *E*2 (Fig. 1) show no or poor electron density in those monomers that make fewer packing contacts. In monomers *A*, *B*, *C* and *G* these loops are stabilized by packing with symmetry-related molecules.

3.2. Dimer interfaces in the crystal structures of FaeE

The three crystal forms reveal a dimeric FaeE structure formed through antiparallel β-strand hydrogen bonding between the *G*1 strands of two FaeE monomers (FaeE dimer 1, Fig. 2*a*). Additionally, hydrogen bonds are formed between Lys106 N^{ζ} in one monomer and the main-chain O atom of Leu94 in the *F*1–*G*1 loop of the other monomer. The *G*1 strand and the *F*1–*G*1 loop are also involved in pilin binding by the chaperone. The *AB* dimer of crystal form 1 was chosen for the study of FaeE dimerization because it had been determined with the highest completeness. It has r.m.s.d. deviations of 2.07 and 1.09 Å with the dimers present in crystal forms 2 and 3, respectively (383 and 408 C^{α} atoms superposed, respectively). The initial complexation of FaeE with FaeG in the solution used to obtain crystals of form 3 does not influence the dimerization of FaeE in the crystal.

A second type of dimer (FaeE dimer 2), created by crystal packing, can be observed in the asymmetric unit of crystal form 1 between monomers *A* and *D* and between monomers *E* and *H* (Fig. 3). This dimer is also present in crystal forms 2 and 3 and is formed through crystal packing with symmetry-related molecules. FaeE dimer 2 is formed around a noncrystallographic twofold-symmetry axis, allowing the *C*1–*D*1′ loop of one monomer to contact the *C*1 strand of the other monomer. Hydrogen bonds are formed between Thr45 and Asp38, Thr45 and Trp36, Arg46 and Gln34 and between Pro47 and Gln34. In addition, Asp44 forms two coordinated salt bridges to Arg104.

The contact area between the two monomers in FaeE dimer 2 is larger than the interface area in dimer 1 (698.5 *versus* 615.2 $Å^2$). Despite this fact, the *AB* dimer is the more relevant one to study, because this dimer has also been observed for other fimbrial chaperones such as PapD and SfaE (Hung, Pinkner *et al.*, 1999; Knight *et al.*, 2002). In all observed cases the subunit binding G1 strands and F1-G1 loops are implicated in this dimer.

Although crystal packing shields the subunit-binding sequences by dimer formation in every case, superposition of FaeE dimer 1 and the SfaE and PapD dimers shows that the dimer interfaces are very different (Fig. 2). Ten hydrogen bonds are formed between main-chain residues of the *G*1 strands of the FaeE monomers in FaeE dimer 1 (Fig. 2*a*). Compared with the well aligned *G*1 strands of the FaeE monomers, the *G*1 strands of the PapD monomers are slightly twisted (Fig. 2*b*). Therefore, only a limited number of hydrogen bonds are formed between the PapD *G*1 strands (Hung, Pinkner *et al.*, 1999). In the SfaE dimer, the N-terminal domain of one monomer (monomer *A*; green in Fig. 2*c*) is packed against the subunit-binding surface of the other (monomer *B*; cyan in Fig. 2*c*; Knight *et al.*, 2002). The *G*1 strand of the first monomer crosses the *A*1 strand of the second monomer at nearly right angles. Phe105, which is part of the conserved pilin-binding patch of the *G*1 strand of monomer *A*, is intercalated between the *A*1′ and *G*1 strands of monomer *B*.

Dimerization of PapD occurs through interaction of the *F*1–*G*1 and *C*2–*D*2 loops of both monomers and significant conformational changes in those loops can be seen when comparing monomeric and dimeric PapD (Hung, Pinkner *et al.*, 1999). There is also a difference in the *F*1–*G*1 loop

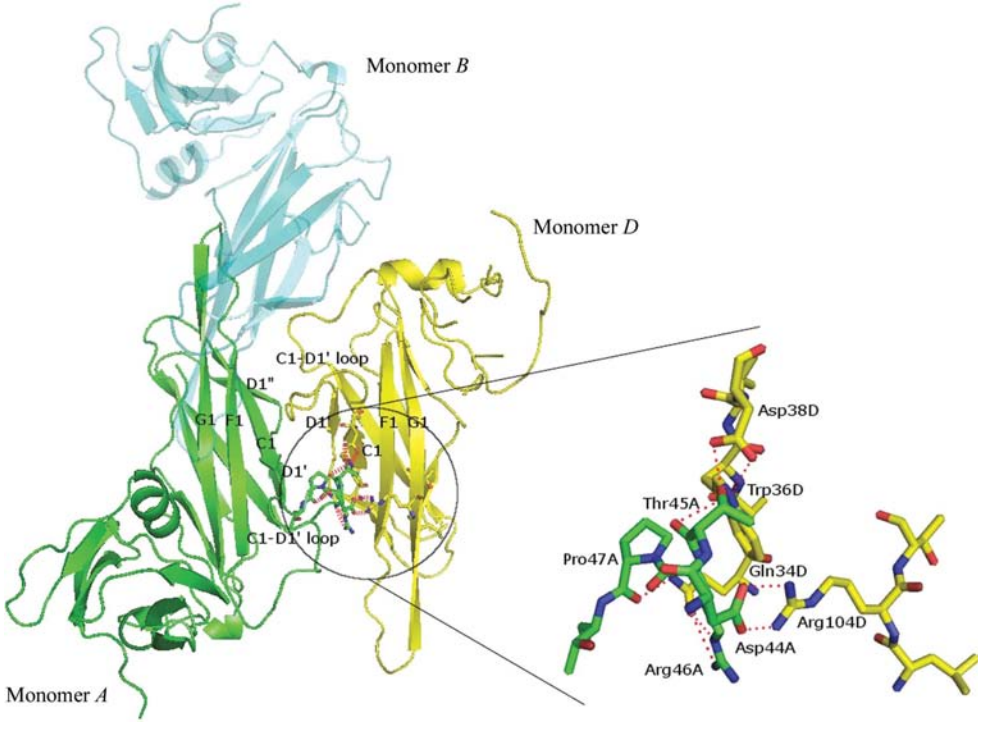

Figure 3
Superposition of both dimer interfaces formed by monomer *A* in crystal form 1. FaeE dimer 1 is formed between monomers *A* (green) and *B* (cyan), while FaeE dimer 2 is formed between monomers *A* and *D* (yellow). The inset shows a detailed view of the hydrogen bonds between the *C*1–*D*1′ loop of chain *A* and the *C*1 strand of chain *D*, as well as the salt bridge formed between Asp44*A* and Arg104*D*. The intramolecular hydrogen bond between Asp38*D* $O^{\delta 2}$ and Trp36*D* $N^{\eta 2}$ is also indicated. Here again the *G*1 strand of both monomers is elongated beyond the contacts it makes with the *F*1 strand because of the contacts it makes within the FaeE dimer 1 interface, as shown for the FaeE dimer 1 formed between monomers *A* and *B*. This figure also demonstrates that the C-terminal domain of monomer *D* is unresolved.

conformation when comparing dimerized and subunit-bound PapD. The *F*1–*G*1 loop of PapD is in a more elongated conformation in the PapD–PapK complex and is buried in the groove of PapK. This conformation cannot be adopted in the PapD dimer because of steric hindrance (Hung, Pinkner *et al.*, 1999). In FaeE dimer 1, the *F*1–*G*1 and *C*2–*D*2 loops of both monomers make no contacts (Fig. 2). As mentioned above, the *G*1 strand extends beyond the contacts it makes with the *F*1 strand because of contacts in the dimer interface. It can thus be presumed that the *G*1 strand is shorter in monomeric FaeE. In the SfaE dimer structure, the *F*1–*G*1 loop is invisible and no structural information is available about monomeric or subunit-bound SfaE (Knight *et al.*, 2002).

3.3. The oligomerization state of FaeE in solution

Subunit binding by periplasmic chaperones occurs through the complementation of the hydrophobic core of the subunit by the *G*1 strand of the chaperone (plus the *A*1 strand for the FGL chaperones), which consists of a pattern of alternating hydrophobic residues (Choudhury *et al.*, 1999; Sauer *et al.*, 1999; Zavialov *et al.*, 2003). Upon subunit binding, the often bulky side chains of these residues are inserted into hydrophobic pockets on the surface of the pilin subunit. When the chaperone is not in complex with a pilin, these hydrophobic side chains are exposed to the solvent. Chaperone oligomerization could therefore serve as a self-capping mechanism that serves to protect the chaperone from aggregation or degradation when not in complex with a pilin (Hung, Pinkner *et al.*, 1999; Knight *et al.*, 2002; Zavialov & Knight, 2007). Such a self-capping mechanism would be confirmed by chaperone oligomerization in solution. It has been shown that the Caf1M F1-antigen chaperone from *Y. pestis* tetramerizes (Zavialov & Knight, 2007) in solution and that PapD dimerizes (Hung, Pinkner *et al.*, 1999). In order to investigate FaeE dimerization in solution, we conducted differential scanning calorimetry (DSC), glutaraldehyde cross-linking and dynamic light-scattering (DLS) measurements.

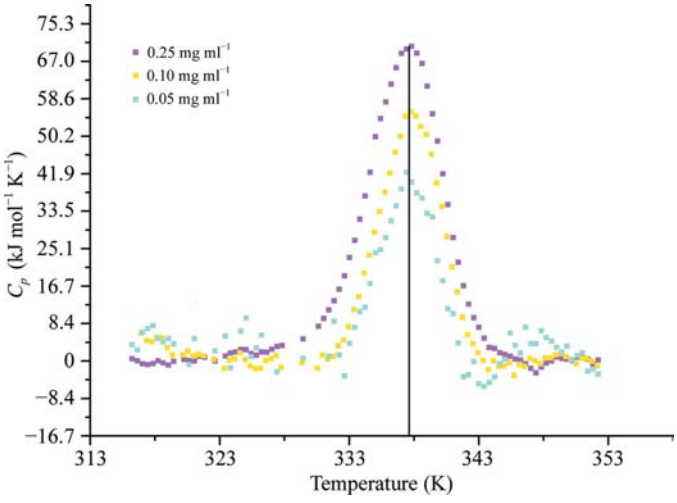

Figure 4
Thermal denaturation of FaeE in 20 m*M* acetate buffer pH 5.2. All scans were recorded at 90 K h^{-1}.

Table 2
Thermodynamic parameters of the thermal denaturation of the chaperone FaeE at different pH values and concentrations.

The error on $T_{1/2}$ is 0.5 K and the error on ΔH_v is 13 kJ mol^{-1}. For pH 5.2 and pH 10.5, ΔC_p° was determined for the highest protein concentration and kept constant in the fitting of the DSC profiles for the other protein concentrations. For pH 7.5, ΔC_p° was determined at 0.33 mg ml^{-1} and at 0.13 mg ml^{-1} and kept constant in the fitting of the DSC profile at 0.06 mg ml^{-1}. For pH 5.2 + 5% PEG 6000, ΔC_p° was determined at 0.2 mg ml^{-1} and kept constant in the fitting of the DSC profiles for the other protein concentrations.

pH	Concentration (mg ml^{-1})	ΔC_p° (J mol^{-1} K^{-1})	$T_{1/2}$ (K)	ΔH_c (kJ mol^{-1})	ΔH_v (kJ mol^{-1})
5.2	0.05	1176	337.8	331	419
	0.1	1176	337.8	599	532
	0.25	1176	337.7	540	557
5.2 + 5% PEG 6000	0.04	6865	339.0	578	645
	0.10	6865	338.9	645	632
	0.20	6865	338.7	712	682
	0.32	6865	338.4	762	795
7.5	0.066	2093	338.6	783	674
	0.13	2093	338.5	682	657
	0.33	2093	338.1	615	687
10.5	0.05	1423	337.2	439	511
	0.1	1423	337.2	573	561
	0.5	1423	336.9	607	573

Equilibrium thermodynamics could be used to analyze the DSC data for FaeE at all pH values when samples were heated to 353 K, the point at which the unfolding of FaeE is completed (Fig. 4). The folding–unfolding transition is either almost completely (96% at pH 5.2) or partially reversible (60% at pH 7.5 and 77% at pH 10.5). Moreover, the transition temperature is independent of the scan rate, indicating that all irreversible processes occur after the unfolding transition. This allowed equilibrium thermodynamics to be applied even in those cases where the transition is not completely reversible. All thermodynamic parameters (ΔC_P°, $T_{1/2}$ and ΔH_v) can be calculated from the fitting of the model function to the experimental DSC data (Table 2).

DSC measurements are a valuable tool to investigate the oligomerization state of a protein in solution. In one approach, the dependence of the DSC profile on the protein concentration can be analyzed. The DSC transition of an oligomeric protein depends on the protein concentration because the energy required for the disruption of the intermolecular interactions depends on the protein concentration. DSC measurements of FaeE at different concentrations show that the DSC profile is independent of the protein concentration (Fig. 4 and Table 2) and thus suggest that FaeE is a monomer. To assess the influence of pH on the DSC transitions, the measurements were repeated at pH 7.5 and 10.5. These results confirmed the independence of the $T_{1/2}$ of the protein concentration (Table 2) and thus the monomeric nature of FaeE. Using a second approach, the ratio of the calorimetric and van't Hoff enthalpies determined from the DSC profiles (ΔH_c and ΔH_v, respectively) gives an estimate of the oligomerization state of a given protein. ΔH_c expresses the enthalpy of denaturation per mole and ΔH_v the enthalpy of denaturation per cooperative unit. Consequently, the ratio between ΔH_c and ΔH_v can be seen as the number of co-

operative units per mole. This ratio is always about 1.0 for FaeE (Table 2). This also points out that both the FaeE N-terminal and C-terminal domains unfold simultaneously as one cooperative unit. In conclusion, both DSC approaches identify FaeE as a monomer in solution.

Chemical cross-linking and dynamic light-scattering (DLS) experiments add further evidence for the stable nature of monomeric FaeE. SDS–PAGE analysis of FaeE samples incubated with glutaraldehyde showed the formation of FaeE oligomers with increasing glutaraldehyde concentration, without the accumulation of a particular oligomeric species (Fig. 5). This is in contrast to the Caf1M chaperone, for which chemical cross-linking clearly showed accumulation of the tetrameric species without the formation of higher oligomers (Zavialov *et al.*, 2007), and the PapD chaperone, which demonstrated the accumulation of PapD dimers (Hung, Pinkner *et al.*, 1999). Thus, the absence of the accumulation of a dimeric FaeE species in the cross-linking experiment

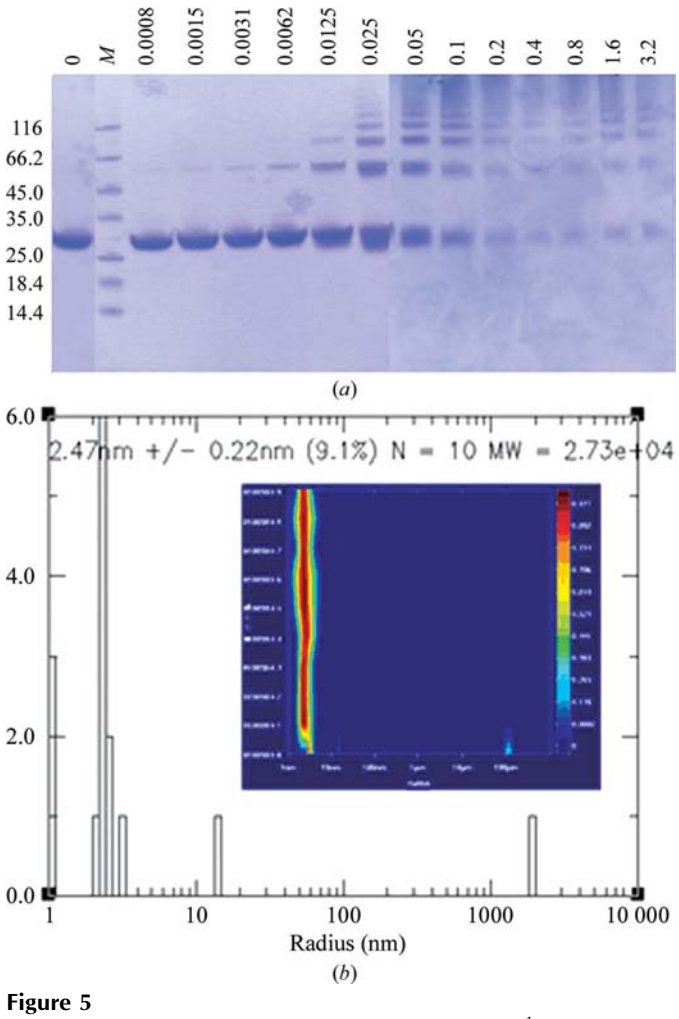

Figure 5
(*a*) Chemical cross-linking of FaeE at 0.75 mg ml^{-1} concentration incubated with varying concentrations of glutaraldehyde as indicated at the top [%(*w*/*v*)]. Lane *M*, protein molecular-weight markers; molecular weights are indicated in kDa on the left. (*b*) Histogram and distribution profile (inset) of the DLS measurement on FaeE at 20 mg ml^{-1}, clearly showing a monodisperse distribution of particles with a Stokes radius of 2.5 nm.

provides additional evidence for the monomeric nature of FaeE. Finally, the monomeric character of FaeE was confirmed by comparing the Stokes radius of FaeE determined by DLS measurements prior to crystallization (2.5 nm; Fig. 5) with the radius of the FaeE monomer in its crystal structures (3.3 nm for a monomer, 5 nm for a dimer).

Overall, our results demonstrate that FaeE can occur as a monomer in solution and does not require self-capping for stabilization and that FaeE dimerization is merely a result of crystallization effects. At the high concentrations used, crystallization of FaeE might only be favoured upon aggregation when the hydrophobic surfaces are capped, in this case by dimerization. The complexation significance score (CSS) calculated by the *PISA* algorithm (0.0) indeed indicates that the dimers formed in the crystal structure are a result of crystal packing only. The same is true for the dimers present in the crystal structures of PapD and SfaE (CSS of 0.0 and 0.1, respectively). In contrast, the CSS for the Caf1M tetramer is 1.0, indicating that the tetramer present in the crystal structure is representative of oligomerization in solution. As during crystallization, *in vivo* crowding effects in the periplasm might urge the chaperone to shield its hydrophobic surfaces. In order to investigate the effect of molecular crowding on the dimerization of FaeE, the DSC experiments were repeated in the presence of 5% PEG 6000. This technique was previously applied by Zavialov & Knight (2007) to show the effect of crowding on the tetramerization of the Caf1M chaperone. This concentration of PEG 6000 clearly stimulated Caf1M tetramerization. However, the DSC profile of FaeE in 5% PEG 6000 is independent of the protein concentration (Table 2), supporting the fact that FaeE dimerization is not required for the stability of the chaperone when it is not in complex with a pilin.

The oligomerization of the Caf1M chaperone seems to be a result of the high β-aggregation propensity of its *G*1 strand (Fig. 6*a*; Zavialov & Knight, 2007). The sequence-based statistical algorithm *TANGO* (Fernandez-Escamilla *et al.*, 2004) does not predict any β-aggregation propensity for the *G*1 strands of PapD, FimC and FaeE, supporting the notion that they could occur as stable monomers without the need for capping of their interactive surfaces. Direct evidence for the monomeric nature of the type 1 fimbrial chaperone FimC in solution comes from its NMR structure (Pellecchia *et al.*, 1998). The fast amide-proton exchange for the *F*1–*G*1 loop confirmed that it is solvent-exposed. The fact that 70% of the PapDQ108C was still present as monomers and that no dimers were found for wild-type PapD (Hung, Pinkner *et al.*, 1999) indicates that dimer formation is not strictly required for PapD stability. Additionally, the PapD chaperone was also found as a monomer in its earliest crystal structure (Holmgren & Brändén, 1989), as was the SafB chaperone (H. Remaut, personal communication). Surprisingly, the *TANGO* predictions for PapD and FaeE are very similar and both indicate the same stretch of residues as prone to β-aggregation (residues 149–156 for PapD and 145–153 for FaeE, in both cases belonging to the *B*2–*C*2 loop and the *C*2 strand; Figs. 6*b* and 6*c*). These residues are not involved in any crystal-packing

interfaces in the crystal structures. For SfaE, *TANGO* predicts a high β-aggregation propensity for the dimerization interface in the crystal structure, which could support SfaE dimerization in solution. However, no evidence is present for this.

4. Concluding remarks

The pilin-free form of the F4 fimbrial chaperone FaeE is predominantly present as a monomer in solution. In its crystal structure, FaeE dimerizes by means of antiparallel β-strand pairing of its pilin-interactive *G*1 strands (FaeE dimer 1). This form of oligomerization has also been observed in the crystal structures of PapD and SfaE and tetramers have been observed both in crystals and in solution for the Caf1M chaperone. However, we find that oligomerization or self-capping of fimbrial chaperones, which has been suggested to protect the chaperone from degradation, is not an absolute requirement for FaeE stability in solution, nor is it a general mechanism to shield the pilin-interactive motif in free periplasmic chaperones. Oligomerization of the Caf1M chaperone seems to be the result of the high β-aggregation propensity of its *G*1 strand. Understandably, the presence of an excess of free periplasmic chaperone, conditional for its crystallization, would be rather unusual during pilus biogenesis, hence the need for self-capping and thus protection of free chaperone would be minor.

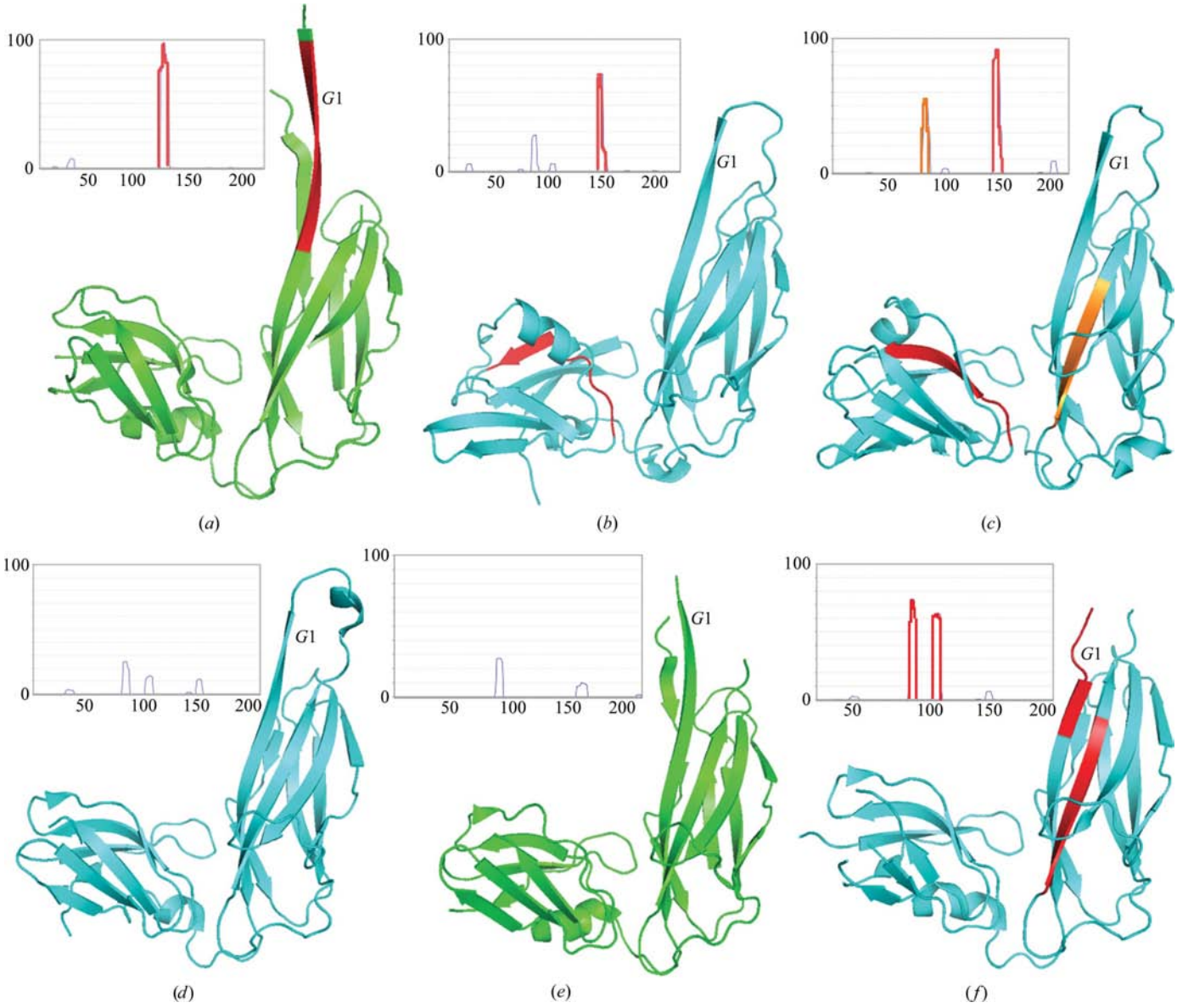

Figure 6
TANGO β-aggregation propensity prediction for Caf1M (*a*), FaeE (*b*), PapD (*c*), FimC (*d*), SafB (*e*) and SfaE(*f*). Regions with the highest β-aggregation propensity are coloured red. For PapD, a second region with ~50% β-aggregation propensity is coloured orange. A cutoff value of 30% β-aggregation propensity was used. The Caf1M, FaeE, PapD, FimC, SafB and SfaE structures were deduced from PDB entries 1p5v, 3f65, 1pdk, 1qun, 2co7 and 1l4i, respectively. FGL chaperones are coloured green and FGS chaperones are coloured cyan. The vertical axis in the insets indicates the percentage of β-aggregation propensity, while the horizontal axis gives the residue number.

The authors acknowledge the use of the EMBL beamline BW7A at the DESY synchrotron (Hamburg, Germay) and of ESRF beamline ID14-1 (Grenoble, France). The Fonds voor Wetenschappelijk Onderzoek-Vlaanderen is thanked for doctoral and postdoctoral grants for KM and LB, and for grant G.0389.07. The authors thank Dr Han Remaut for carefully reading the manuscript.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.

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.

Cowtan, K. D. & Zhang, K. Y. (1999). *Prog. Biophys. Mol. Biol.* **72**, 245–270.

Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B. III, Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007). *Nucleic Acids Res.* **35**, W375–W383.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. & Serrano, L. (2004). *Nature Biotechnol.* **22**, 1302–1306.

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

Holmgren, A., Kuehn, M. J., Brändén, 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., Danese, P. N., Pinkner, J. S., Silhavy, T. J. & Hultgren, S. J. (1997). *EMBO J.* **16**, 6394–6406.

Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.

Knight, S. D., Choudhury, D., Hultgren, S., Pinkner, J., Stojanoff, V. & Thompson, A. (2002). *Acta Cryst.* D**58**, 1016–1022.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

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

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

Nuccio, S. P. & Baumler, A. J. (2007). *Microbiol. Mol. Biol. Rev.* **71**, 551–575.

Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.

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

Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001). *Acta Cryst.* D**57**, 1445–1450.

Remaut, H., Rose, R. J., Hannan, T. J., Hultgren, S. J., Radford, S. E., Ashcroft, A. E. & Waksman, G. (2006). *Mol. Cell*, **22**, 831–842.

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

Sauer, F. G., Pinkner, J. S., Waksman, G. & Hultgren, S. J. (2002). *Cell*, **111**, 543–551.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Soto, G. E. & Hultgren, S. J. (1999). *J. Bacteriol.* **181**, 1059–1071.

Thanassi, D. G., Saulino, E. T. & Hultgren, S. J. (1998). *Curr. Opin. Microbiol.* **1**, 223–231.

Van Molle, I., Buts, L., Coppens, F., Qiang, L., Wyns, L., Loris, R., Bouckaert, J. & De Greve, H. (2005). *Acta Cryst.* F**61**, 427–431.

Vetsch, M., Puorger, C., Spirig, T., Grauschopf, U., Weber-Ban, E. U. & Glockshuber, R. (2004). *Nature (London)*, **431**, 329–333.

Zavialov, A. V., Berglund, J., Pudney, A. F., Fooks, L. J., Ibrahim, T. M., MacIntyre, S. & Knight, S. D. (2003). *Cell*, **113**, 587–596.

Zavialov, A. V. & Knight, S. D. (2007). *Mol. Microbiol.* **64**, 153–164.

Zavialov, A., Zav'yalova, G., Korpela, T. & Zav'yalov, V. (2007). *FEMS Microbiol. Rev.* **31**, 478–514.