Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

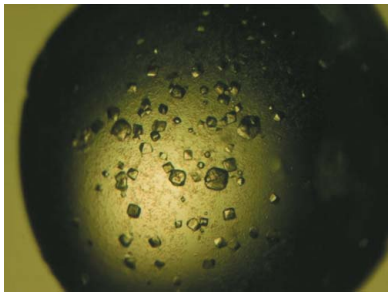
Crystallization and initial crystallographic analysis of AafA: the major adhesive subunit of the enteroaggregative *Escherichia coli* AAF/II pilus

Yi Yang, James A. Garnett and Stephen Matthews*

Centre for Structural Biology, Division of Molecular Biosciences, Department of Life Sciences, Imperial College London, South Kensington, London SW7 2AZ, England

Correspondence e-mail: s.j.matthews@imperial.ac.ukchkim@krict.re.kr

Received 26 November 2010
Accepted 10 January 2011



© 2011 International Union of Crystallography
All rights reserved

AafA is the major adhesive pilin subunit of the aggregative adherence fimbriae (AAF) from enteroaggregative *Escherichia coli*, which play an important role by attaching to the host cells during the initial phase of bacterial colonization and invasion. AafA has been crystallized at pH 3.4 and diffraction data have been collected to 2.1 Å resolution. Molecular replacement was unsuccessful and selenomethionine-substituted protein and heavy-atom derivatives are being prepared for phasing.

1. Introduction

Enteroaggregative *Escherichia coli* (EAEC) causes acute and persistent diarrhoea worldwide and leads to watery, mucoid, secretory diarrhoea with low-grade fever (Weintraub, 2007; Okhuysen & Dupont, 2010). Previous studies have shown that EAEC infection is the most common cause of diarrhoea resulting from bacterial infection (Nataro *et al.*, 2006).

EAEC is characterized by its 'stacked-brick' aggregative adherence pattern to HEp-2 cells or other bacteria (Nataro *et al.*, 1987). Like other pathogenic *E. coli*, EAEC possesses fimbriae, named aggregative adherence fimbriae (AAF), as adherence factors. To date, four different AAFs have been identified, all of which are encoded on a 55–65 MDa plasmid (pAA; Nataro *et al.*, 1993; Czeczulin *et al.*, 1997; Bernier *et al.*, 2002; Boisen *et al.*, 2008). All four of these AAF variants show a similar gene organization to those of the chaperone–usher Afa/Dr family of adhesins from diffusely adherent *E. coli* (DAEC), comprising genes expressing a chaperone, an usher, a major pilin which makes up the majority of the fibre and a minor capping pilin subunit (Servin, 2005). The organelle encoded by one of the AAF variants, namely AAF/II, possesses two structural subunits: a major subunit called AafA and a minor subunit called AafB (Elias *et al.*, 1999). During EAEC colonization, AafA is responsible for the initial attachment to the host cells by recognizing extracellular matrix (ECM) proteins such as fibronectin, laminin and type IV collagen (Farfan *et al.*, 2008). Recent studies show that AafA, rather than AafB, is able to disrupt epithelial cell tight junctions during colonization, which could potentially be followed by direct secretion of fluid and electrolytes into host cells (Okhuysen & Dupont, 2010).

AafA is a 14.5 kDa protein which possesses an unstructured N-terminal extension. During the construction of the pilus this flexible N-terminal extension binds in a hydrophobic groove of an adjacent subunit, completing the Ig fold. This process is termed donor-strand exchange (Remaut *et al.*, 2006). In this study, a donor-strand complemented version of AafA (AafA-dsc, 16.9 kDa; hereafter referred to as AafA) was used; it was obtained by cloning the N-terminal extension onto the C-terminus connected through a short linker (Fig. 1; Anderson *et al.* 2004). Here, we report the preliminary crystallographic analysis of AafA.

2. Materials and methods

2.1. Protein purification and crystallization

To create a donor-strand complemented *aafA* construct (Anderson *et al.*, 2004; Remaut *et al.*, 2006), PCR was used to clone amino-acid

residues 11–135 at the N-terminus, followed by a ten-residue linker and then residues 1–10 of AafA at the C-terminus (Fig. 1). This was then ligated into the pQE-30 vector (Qiagen) *via* *Bam*HI and *Hin*dIII restriction sites and expressed in *E. coli* strain M15 cells with pREP4 plasmids. The cells were induced with 1 m*M* IPTG when the OD_{600} reached 0.6, which was followed by overnight incubation at 310 K before harvesting by centrifugation. The cells were lysed using a French press under denaturing conditions before being purified with Ni–NTA (Qiagen). The eluate was first dialysed against 50 m*M* sodium acetate pH 5, 50 m*M* NaCl, 1 *M* urea, which was followed by a

Figure 1
Schematic representation of donor-strand complemented AafA. The N-terminal 12 residues consists of a His_6 tag (green) followed by residues 11–135 of mature AafA (dark blue), a DNKQ tetrapeptide linker (orange), six residues of periplasmic signal peptide (light blue) as an additional linker sequence and finally residues 1–10 of mature AafA (dark blue) at the C-terminus. The residue numbering of AafA is shown above the figure.

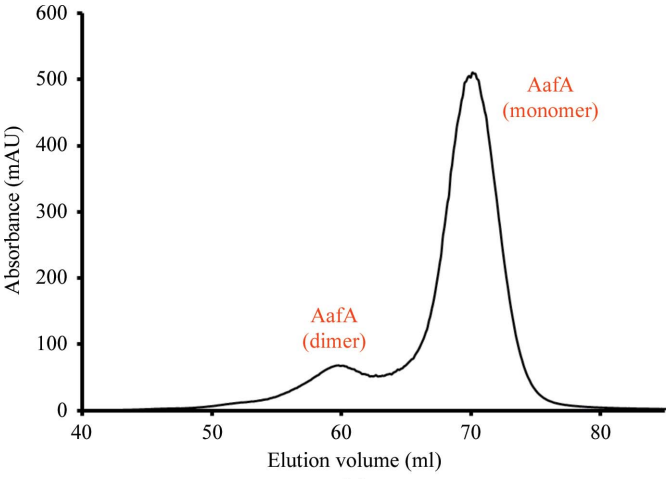


Figure 2
Purification of AafA. (*a*) Superdex 75 (GE Healthcare) gel-filtration profile of AafA. (*b*) SDS–PAGE of monomeric AafA after gel filtration.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

Space group	*I*222 or $I2_12_12_1$
Unit-cell parameters (Å)	$a = 42.3$, $b = 76.3$, $c = 81.1$
Resolution range (Å)	23.3–2.1 (2.21–2.10)
Wavelength (Å)	1.5418
Total reflections	36590 (4761)
Mosaicity (°)	0.78
Unique reflections	8372 (1159)
Completeness (%)	97.0 (94.5)
Multiplicity	4.4 (4.1)
R_{merge}† (%)	9.8 (31.4)
$\langle I/\sigma(I)\rangle$	6.8 (2.4)
Molecules per asymmetric unit‡	1
Solvent content (%)	40.8

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of the observations $I_i(hkl)$ of reflection hkl. ‡ Most probable value.

second dialysis against the same buffer with no urea. AafA was further purified by gel filtration using a Superdex 75 gel-filtration column (GE Healthcare; Fig. 2); monomeric AafA was pooled, dialysed against 50 m*M* sodium acetate pH 5, 50 m*M* NaCl and then concentrated to 5.5 mg ml^{-1}. Conditions for crystallization were initially screened by the sitting-drop method of vapour diffusion at 293 K with sparse-matrix crystallization kits (Hampton Research, USA; Emerald BioStructures, USA; Molecular Dimensions Ltd, USA) in 96-well MRC plates with 400 nl protein solution and 400 nl reservoir solution using a Mosquito nanolitre high-throughput robot (TTP Labtech). Small crystals (~50 μm^3) were obtained from 3.2 *M* NaCl, 100 m*M* citric acid pH 3.4 within three months (Fig. 3). Although optimization of these conditions was attempted by altering the protein concentration and crystallization components, screening additives (Hampton Research, USA), increasing the drop size and using different crystallization plates (24-well Limbro plates from Hampton Research, USA and a 48-well MRC Maxi plate from Molecular Dimensions, USA), there was no increase in the production time, the size or the quality of the crystals.

2.2. X-ray data collection, processing and analysis

Crystals were cryoprotected by transferring them to 5 µl of 3.2 *M* NaCl, 100 m*M* citric acid pH 3.4, 7.5% glycerol, allowing 5 min for equilibration and then repeating the transfer with additional 7.5% increments of glycerol up to a final cryoprotectant concentration of 3.2 *M* NaCl, 100 m*M* citric acid pH 3.4, 30% glycerol. Crystals were immediately mounted in a cryoloop and flash-cooled in liquid

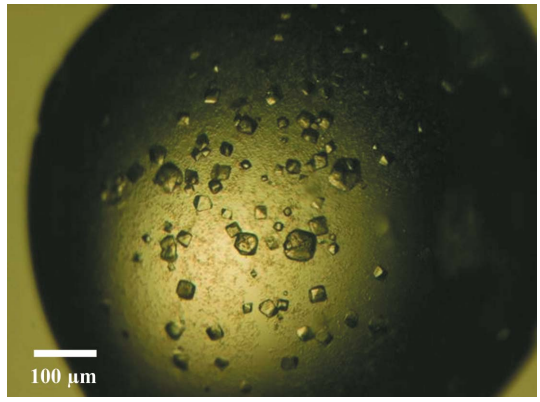


Figure 3
Representative native crystals of AafA.

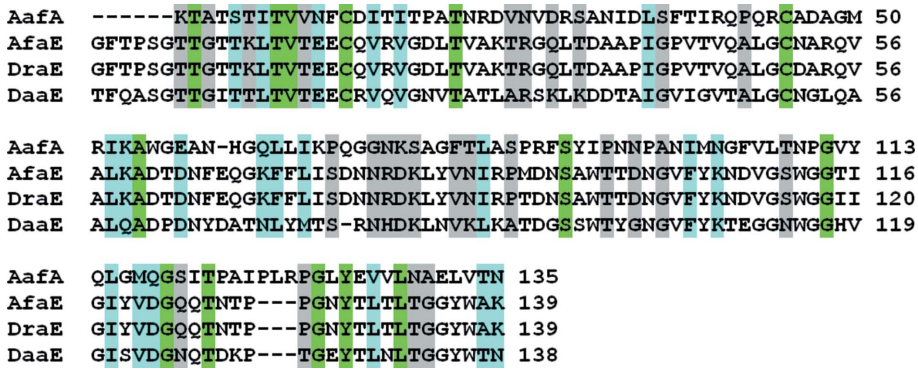


Figure 4
Sequence alignment of mature AafA (UniProtKB O30595), AfaE-III (UniProtKB Q57254), DraE (UniProtKB Q7BG35) and DaaE (UniProtKB P13719). The match strength is highlighted (highest to lowest: green, cyan and grey).

nitrogen. Diffraction data from a single native crystal were collected at a wavelength of 1.5418 Å on our in-house Rigaku MicroMax-007 HF-M high-flux generator coupled with a Rigaku Saturn 944+ CCD detector. Data were processed with *MOSFLM* (Leslie, 2006) and scaled with *SCALA* (Evans, 2006). Data-collection statistics are shown in Table 1. The contents of the unit cell were analyzed using the Matthews coefficient (Matthews, 1968) and molecular replacement was attempted with *AMoRe* (Navaza, 1994), *MOLREP* (Vagin & Teplyakov, 2010) and *Phaser* (McCoy *et al.*, 2005) with the crystal structures of DaaE (PDB entry 2bcm; Korotkova *et al.*, 2006), DraE (PDB entry 1ut1; Anderson *et al.*, 2004) and AfaE-III (PDB entry 1ut2; Anderson *et al.*, 2004) and the NMR structure of AfaE-III (PDB entry 1ut2; Anderson *et al.*, 2004) as search models.

3. Results and discussion

Crystals of AafA were obtained after three months and grew to $\sim$50 µm^3 (Fig. 3). Data were collected to 2.1 Å resolution and were indexed in space group *I*222. Examination of the systematic absences could not confirm the true space group, which may be either *I*222 or $I2_12_12_1$. Analysis of the crystal content indicated that there is a single molecule in the asymmetric unit, with a Matthews coefficient of 2.08 Å^3 Da^{-1} (Matthews, 1968) and a corresponding solvent content of 40.8%. Data-collection and processing statistics are listed in Table 1.

Molecular replacement was attempted with *AMoRe* (Navaza, 1994), *MOLREP* (Vagin & Teplyakov, 2010) and *Phaser* (McCoy *et al.*, 2005) using the crystal structures of DaaE (PDB entry 2bcm; Korotkova *et al.*, 2006), DraE (PDB entry 1ut1; Anderson *et al.*, 2004) and AfaE-III (PDB entry 1ut2; Anderson *et al.*, 2004) and the NMR structure of AfaE-III (PDB entry 1ut2; Anderson *et al.*, 2004) as search models. Unfortunately, no solutions were found; however, the sequence identity between AafA and these homologues is less than 20% (Fig. 4). We are currently preparing selenomethionine-substituted protein and heavy-atom derivatives with a view to solving the phase problem using anomalous dispersion techniques.

This work was supported by a grant from the Wellcome Trust.

References

Anderson, K. L. *et al.* (2004). *Mol. Cell*, **15**, 647–657.
Bernier, C., Gounon, P. & Le Bouguénec, C. (2002). *Infect. Immun.* **70**, 4302–4311.
Boisen, N., Struve, C., Scheutz, F., Krogfelt, K. A. & Nataro, J. P. (2008). *Infect. Immun.* **76**, 3281–3292.
Czeczulin, J. R., Balepur, S., Hicks, S., Phillips, A., Hall, R., Kothary, M. H., Navarro-Garcia, F. & Nataro, J. P. (1997). *Infect. Immun.* **65**, 4135–4145.
Elias, W. P., Czeczulin, J. R., Henderson, I. R., Trabulsi, L. R. & Nataro, J. P. (1999). *J. Bacteriol.* **181**, 1779–1785.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Farfan, M. J., Inman, K. G. & Nataro, J. P. (2008). *Infect. Immun.* **76**, 4378–4384.
Korotkova, N., Le Trong, I., Samudrala, R., Korotkov, K., Van Loy, C. P., Bui, A. L., Moseley, S. L. & Stenkamp, R. E. (2006). *J. Biol. Chem.* **281**, 22367–22377.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Nataro, J. P., Kaper, J. B., Robins-Browne, R., Prado, V., Vial, P. & Levine, M. M. (1987). *Pediatr. Infect. Dis. J.* **6**, 829–831.
Nataro, J. P., Mai, V., Johnson, J., Blackwelder, W. C., Heimer, R., Tirrell, S., Edberg, S. C., Braden, C. R., Morris, J. G. & Hirshon, J. M. (2006). *Clin. Infect. Dis.* **43**, 402–407.
Nataro, J. P., Yikang, D., Giron, J. A., Savarino, S. J., Kothary, M. H. & Hall, R. (1993). *Infect. Immun.* **61**, 1126–1131.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Okhuysen, P. C. & Dupont, H. L. (2010). *J. Infect. Dis.* **202**, 503–505.
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.
Servin, A. L. (2005). *Clin. Microbiol. Rev.* **18**, 264–292.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Weintraub, A. (2007). *J. Med. Microbiol.* **56**, 4–8.