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

Crystallization and preliminary X-ray diffraction analysis of CfaE, the adhesive subunit of the CFA/I fimbriae from human enterotoxigenic *Escherichia coli*

Yong-Fu Li,[a] Steven Poole,[b] Fatima Rasulova,[b] Lothar Esser,[a] Stephen J. Savarino[b,c] and Di Xia[a]*

[a]Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD 20892-4256, USA, [b]Enteric Diseases Department, Infectious Diseases Directorate, Naval Medical Research Center, Silver Spring, MD 20910-7500, USA, and [c]Department of Pediatrics, Uniformed Services University of the Health Sciences, Bethesda, MD 20814-4799, USA

Correspondence e-mail: dixia@helix.nih.gov

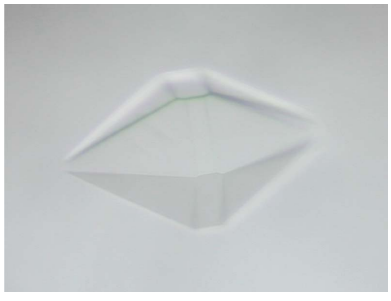
Received 23 November 2005
Accepted 24 December 2005

Enterotoxigenic *Escherichia coli* (ETEC) represents a formidable food and waterborne diarrheal disease threat of global importance. The first step in ETEC pathogenesis is bacterial attachment to small-intestine epithelial cells *via* adhesive fimbriae, many of which are genetically related to the prototype colonization factor antigen I (CFA/I). The minor fimbrial subunit CfaE is required for initiation of CFA/I fimbrial assembly and mediates bacterial attachment to host cell-surface receptors. A donor-strand complemented variant of CfaE (dscCfaE) was expressed with a hexahistidine tag, purified to homogeneity and crystallized using the hanging-drop vapor-diffusion method. X-ray diffraction data sets were collected to 2.4 Å resolution for both native and derivatized crystals and showed the symmetry of space group $P6_222$, with unit-cell parameters $a = b = 142.9$, $c = 231.9$ Å. Initial phases were derived from the SIRAS approach and electron density showed two molecules in the crystallographic asymmetric unit. Sequence assignments were aided by anomalous signals from the selenium of an SeMet-derivatized crystal and from S atoms of a native crystal.

1. Introduction

Enterotoxigenic *Escherichia coli* (ETEC) is a major cause of diarrhea in travelers to and in young children of resource-limited countries, affecting hundreds of millions of persons worldwide annually (Wenneras & Erling, 2004). ETEC ingested with contaminated food or water colonize the small intestine and secrete heat-labile and/or heat-stable enterotoxins, thereby causing diarrhea (Black, 1990; Huilan *et al.*, 1991). As the first step in pathogenesis, ETEC express adhesive fimbriae, the filamentous surface protein structures referred to as colonization factors (CFs), which specifically recognize and bind to receptors on intestinal epithelial cells. To date, over 22 different CFs have been identified, some of which have been clearly shown to be critical in pathogenesis (Gaastra & Svennerholm, 1996). Among them is the first reported human-specific ETEC CF, colonization factor antigen I (CFA/I; Evans *et al.*, 1975), the archetype of a family of eight related but serologically distinct fimbriae that share similar genetic and biochemical features. These include coli surface antigen 1 (CS1), CS2, CS4, CS14, CS17, CS19 and the putative CF O71 (PCFO71) (Anantha *et al.*, 2004). CFA/I biogenesis requires four protein components: the major (CfaB) and minor (CfaE) structural subunits and two assembly proteins, a periplasmic chaperone (CfaA) that shuttles subunits to the outer membrane and an outer membrane usher protein (CfaC) that orchestrates ordered tip-to-base assembly of fimbriae. Bioassembly of CFA/I and related fimbriae has been dubbed the 'alternate chaperone pathway' (Sakellaris & Scott, 1998) to distinguish it from the classic chaperone–usher pathway that guides assembly of class 1 pilus structures such as P and type I pili (Soto & Hultgren, 1999). Recent studies have clearly demonstrated that the minor CfaE subunit and its homologs are tip-localized and serve as the fimbrial adhesin in *in vitro* models of intestinal adherence (Anantha *et al.*, 2004; Sakellaris *et al.*, 1999; Poole *et al.*, submitted).

Antibiotic treatment options have been eroded by the inexorable rise in antimicrobial resistance to first-line drugs, stimulating interest in effectors of virulence as vaccine components as well as expanded targets for the development of new drugs or small-molecule inhibitors to prevent or treat related bacterial diseases. Towards these



goals, a detailed structural knowledge of the fimbria, especially the minor adhesive subunit, will further our understanding of the interaction between host receptors and the bacterial fimbria as well as the fimbrial assembly process. In this communication, we describe the crystallization and preliminary X-ray crystallographic analysis of a form of the CfaE minor adhesive subunit of CFA/I created by *in cis* donor-strand complementation (Poole *et al.*, submitted).

2. Materials and methods

2.1. Cloning, expression and purification of recombinant dscCfaE

A detailed description of the cloning and characterization of dcs_{19}CfaE[His]$_6$ (dscCfaE) will be given elsewhere (Poole *et al.*, submitted). Briefly, a pET-24(a)+T7 expression plasmid was engineered to express CfaE with a C-terminal extension consisting of a short hairpin linker (DNKQ) followed by the N-terminal 19 amino-acid residues of CfaB (CFA/I major subunit) and a hexahistidine affinity tag. The 19-amino-acid CfaB segment was incorporated to provide *in cis* a β-strand predicted to be missing from the native CfaE protein. The resultant plasmid, pET24-dsc19cfaE(his)6, was transformed into *E. coli* BL21(DE3) and used to produce dscCfaE. This strain was grown at 305 K in Alternative Protein Source (APS) superbroth (Difco, Detroit, MI, USA) (with kanamycin) to late log phase and induced for 3 h by addition of 1 m*M* IPTG. Cell paste was derived by centrifugation and resuspended in Binding Buffer A containing Benzonase (Novagen) and disrupted by microfluidization (Model 1109 Microfluidizer, Microfluidics Corp., Newton, MA, USA). After centrifugation, the supernatant was loaded onto a HisTrap HP column (GE Healthcare Bio-Sciences Corp., Piscataway, NJ, USA), and equilibrated with binding buffer. The dscCfaE protein was eluted with a gradient to 300 m*M* imidazole. Fractions containing dscCfaE were determined by SDS–PAGE and Western blotting, pooled and diluted tenfold with 25 m*M* MES buffer, pH 6, before loading onto a HiTrap SP column (GE Healthcare Bio-Sciences). Protein elution from the column was achieved using a gradient to 0.5 *M* NaCl, and fractions containing dscCfaE were pooled and dialyzed against PBS, pH 6.7. After concentration, the protein sample was loaded onto a Superdex 200 column. Protein fractions from the dscCfaE peak were concentrated using a Centricon filter unit to a concentration of 10 mg ml^{-1} (Fig. 1*a*).

Selenomethionine-substituted protein (SeMet-dscCfaE) was expressed in a methionine-auxotrophic *E. coli* strain (B834-DE3, Novagen) and cells were grown in a minimal medium supplemented with 19 amino acids plus selenomethionine. The derivatized protein was purified using the same procedure as used for the native protein.

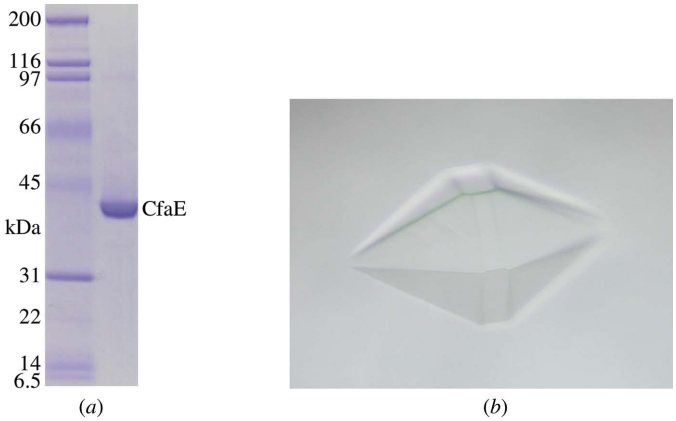


Figure 1
(*a*) SDS–PAGE of purified dscCfaE. (*b*) A dscCfaE crystal.

2.2. Crystallization, X-ray diffraction data collection and reduction

The purified dscCfaE protein at 10 mg ml^{-1} was used for initial crystallization screening by the vapor-diffusion method at 288 K. Hexagonal bipyramidal-shaped crystals were grown from condition C9 of the Index HT kit (1.1 *M* sodium malonate, 0.1 *M* HEPES pH 7.0, 0.5% Jeffamine ED-2001; Hampton Research). Subsequent optimization of parameters such as precipitant concentrations, pH range and additives led to the final well solution of 1.3–1.4 *M* sodium malonate, 0.1 *M* HEPES pH 7.0, 0.1% Jeffamine ED-2001, 100 m*M* NaCl and 10% PEG MME 500. For crystal production, 2.0 µl protein sample at 5 mg ml^{-1} was mixed with 2.0 µl well solution; the droplets were incubated at 288 K. Within 7–10 d, dscCfaE crystals grew to dimensions of 0.2 × 0.1 × 0.1 mm (Fig. 1*b*). Crystals of SeMet-dscCfaE were seeded with native crystals and grew in a similar way.

Solutions containing different salts, glycerol and PEGs of various concentrations were tested for cryoprotection and crystal stability. dscCfaE crystals were stable and diffracted X-rays well in a solution containing 2.4 *M* sodium acetate, 0.1 *M* HEPES pH 7.0, 0.1% Jeffamine ED-2001, 100 m*M* NaCl and 10% PEG 200. Various heavy-metal compounds (such as Hg, Pb, Os, Au, Cs and lanthanides) solubilized in the stabilization solution were used to derivatize the crystals. The cross-linking agent glutaraldehyde was also used to provide mechanical stability to the crystals in various soaking experiments.

The X-ray diffraction data sets reported in this study were collected at the SER-CAT beamline at the Advanced Photon Source (APS), Argonne National Laboratory (ANL) with a MAR 300 CCD detector. The crystal-to-detector distance was set between 250 and 300 mm and the oscillation range was either 0.5 or 1°. The raw data frames were indexed and intensities integrated using the program *DENZO*; integrated intensities from each diffraction image of the

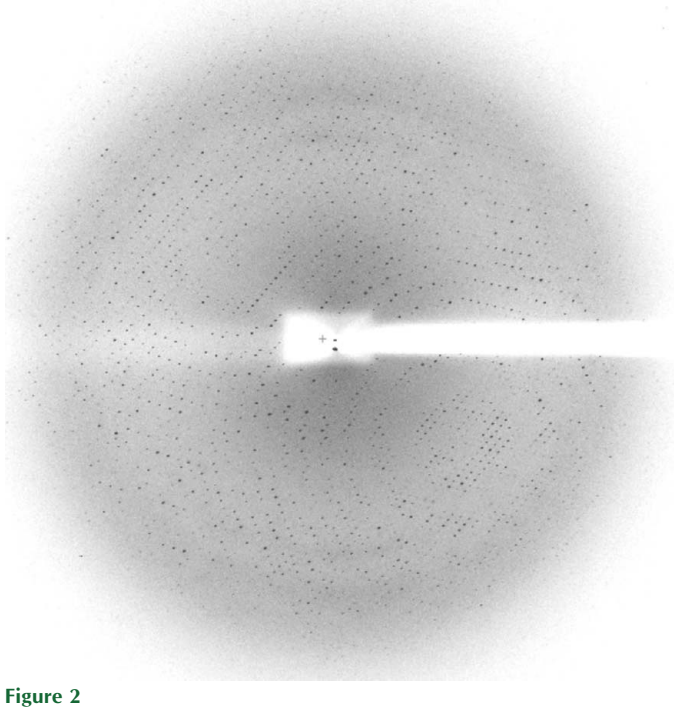


Figure 2
A typical X-ray diffraction pattern of a dscCfaE crystal using a synchrotron-radiation source.

Table 1
Statistics of the X-ray diffraction data sets.

Values in parentheses are for the outermost shell.

Data set	Native	Au	SeMet	Native^{LW}†
Wavelength (Å)	1.00000	1.00052	0.97931‡	1.74000
Beamlines	22-BM, APS	22-ID, APS	22-ID, APS	22-BM, APS
Exposure time (s)	11	6	6	8
Resolution (Å)	50–2.32 (2.40–2.32)	50–2.70 (2.80–2.70)	50–2.80 (2.90–2.80)	50–2.80 (2.90–2.80)
Unit-cell parameters (Å)	$a = b = 143.920$, $c = 231.735$	$a = b = 143.435$, $c = 230.783$	$a = b = 143.763$, $c = 231.046$	$a = b = 144.322$, $c = 231.513$
No. of observations	417146	438974	475133	1361260
No. of unique reflections	61940	39333	35627	35888
Mosaicity (°)	0.336	0.313	0.333	0.326
R_{merge}§	0.095 (0.491)	0.090 (0.325)	0.115 (0.466)	0.105 (0.344)
Completeness (%)	95.2 (92.1)	99.9 (100.0)	96.4 (96.4)	99.9 (99.9)
Average $I/\sigma(I)$	15.5 (2.2)	21.3 (7.8)	21.6 (9.0)	45.3 (13.2)
Redundancy	7.1 (3.8)	11.2 (6.3)	13.8 (13.6)	38.0 (24.7)

† A native data set collected at 1.74 Å wavelength. ‡ The peak wavelength for the SeMet crystal was based on an extended X-ray absorption fine-structure (EXAFS) scan. § R_{merge} is defined as $\sum|I_{h,i} - \langle I_h\rangle|/\sum I_{h,i}$, where $I_{h,i}$ is the intensity of the ith observation of a reflection with Miller index h and $\langle I_h\rangle$ is the mean intensity for all measured I_hs and Friedel pairs.

Table 2
Phasing statistics.

The space group for CfaE crystals is $P6_222$.
(*a*) SIR phasing with Au.

Site	*X*	*Y*	*Z*	*O*	*B*
1	0.4655	0.4385	0.0074	0.3311	52.28
2	0.3645	0.0084	0.0979	0.1801	8.96
3	0.5536	0.0079	0.0679	0.1460	40.62

(*b*) Phasing statistics in resolution shells.

Resolution	12.78	8.50	6.78	5.80	5.15	4.67	4.31	4.02	Total
No. of centric reflections	276	294	286	287	288	259	274	259	2223
PP† for centric reflections	0.79	0.97	1.16	0.93	0.73	0.67	0.59	0.54	0.80
No. of acentric reflections	474	839	1115	1330	1508	1670	1817	1944	10697
PP for acentric reflections	1.03	1.37	1.67	1.36	1.10	0.99	0.86	0.77	1.07
Mean figure of merit	0.42	0.52	0.57	0.52	0.46	0.43	0.39	0.36	0.45

† Phasing power is defined as $\langle f_H\rangle$/residual, where $\langle f_H\rangle$ is the root-mean-square (r.m.s.) of the structure-factor contribution in amplitudes from heavy atoms in the derivative and residual is the r.m.s. of the lack-of-closure error.

same crystal were merged and scaled with the program *SCALEPACK*. Both programs are part of the *HKL*2000 package (Otwinowski & Minor, 1997).

3. Results and discussion

3.1. Protein purification, crystallization, cryoprotection and preliminary X-ray diffraction analysis

The donor-strand complemented minor adhesive subunit CfaE of the human-specific ETEC CFA/I fimbriae was cloned, overexpressed and purified to homogeneity by nickel-affinity, ion-exchange and size-exclusion chromatography (Fig. 1*a*). Both N-terminal sequence and mass-spectrometry analysis on the purified protein confirmed the cleavage of a 22-amino-acid signal peptide. The mature protein contains 369 residues with a molecular weight of 40.9 kDa. Purified dscCfaE was concentrated to approximately 10 mg ml^{-1}, from which hexagonal bipyramidal-shaped dscCfaE crystals were obtained (Fig. 1*b*). These crystals diffracted X-rays at room temperature when mounted in quartz capillaries, but were radiation-sensitive. Flash-freezing of crystals to 100 K led to an elevated mosaic spread, which prompted a search for a cryoprotection condition that could minimize the increase in mosaicity of dscCfaE crystals and is compatible with the heavy-metal derivatization process.

Crystals of dscCfaE typically diffracted X-rays to 2.2 Å resolution (Fig. 2) using a synchrotron-radiation source. The diffraction pattern is consistent with a hexagonal crystal lattice. Systematic absences for the 00l reflections showed diffraction conditions for the $l = 3n$ reflections only, indicating the symmetry of either space group $P6_222$ or $P6_422$. The native crystal has unit-cell parameters $a = b = 143.919$, $c = 231.734$ Å. Calculation of the Matthews coefficient (V_M) suggested that there are either two (4.22 Å^3 Da^{-1}) or three (2.82 Å^3 Da^{-1}) dscCfaE molecules per crystallographic asymmetric unit (Matthews, 1968).

3.2. Preparation of heavy-metal-derivatized crystals

We designed an acetate-based stabilization solution and transferred dscCfaE crystals to a few droplets of this solution sequentially to remove malonate, since attempts to soak various heavy-metal compounds into dscCfaE crystals in a malonate-based stabilization solution did not yield useful derivatized crystals. This was presumably owing to the chelating power of malonate. The prolonged soaking procedure, although effective in replacing malonate with acetate, frequently caused damage to the crystals, which could be alleviated by treating crystals with cross-linking agents such as glutaraldehyde prior to soaking. Numerous heavy-metal compounds were tested; these included mercurial, lead, osmium, gold, cesium, platinum, selenium, tungsten and some lanthanide compounds. We also tested several halides such as bromide and iodide salts.

3.3. Phase determination

Although as many as 20 diffraction data sets were collected from heavy-metal-soaked crystals, only a few were useful for phasing (Tables 1 and 2). A native crystal diffracted X-rays to better than 2.2 Å resolution and the resulting data set was processed to 2.4 Å resolution with an R_{merge} of 9.5%. One data set was obtained from a gold-derivatized crystal and used for SIRAS phasing with the program *SOLVE* (Terwilliger, 2003). Three metal-binding sites were identified with reasonable occupancies, giving an overall figure of merit of 0.45 in the resolution range 15–4.0 Å for space group $P6_222$ (Table 1). The program *RESOLVE* (Terwilliger, 2003) was subsequently used for phase improvement by density modification and for automated chain tracing. The resulting electron-density map contains clearly recognizable features for protein. The SeMet and the native^{LW}

(collected at a wavelength of 1.74 Å) data sets were used to compute anomalous Fourier maps with phase information from the improved SIRAS. The positions of all S atoms in methionine and cysteine residues were unambiguously identified and were helpful in sequence assignment in model building, which is currently in progress. Based on the SIRAS electron-density map, there are two dscCfaE molecules in the crystallographic asymmetric unit.

The authors wish to thank the beamline staff of the SER-CAT at APS, ANL for their assistance in data collection. We acknowledge the help from Dr S. Hess of NIDDK for the mass-spectrometric analysis of dscCfaE. This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, by a grant from the Trans NIH/FDA Intramural Biodefense Program (to DX), by the US Army Military Infectious Diseases Research Program Work Unit Number A0307 (to SJS) and by the Henry M. Jackson Foundation for the Advancement of Military Medicine (to SJS). We thank Junio Colobong for technical assistance in preparing the SeMet derivative of dscCfaE.

References

Anantha, R. P., McVeigh, A. L., Lee, L. H., Agnew, M. K., Cassels, F. J., Scott, D. A., Whittam, T. S. & Savarino, S. J. (2004). *Infect. Immun.* **72**, 7190–7201.
Black, R. E. (1990). *Rev. Infect. Dis.*, **12**, *Suppl.* 1, S73–S79.
Evans, D. G., Silver, R. P., Evans, D. J. Jr, Chase, D. G. & Gorbach, S. L. (1975). *Infect. Immun.* **12**, 656–667.
Gaastra, W. & Svennerholm, A. M. (1996). *Trends Microbiol.* **4**, 444–452.
Huilan, S., Zhen, L. G., Mathan, M. M., Mathew, M. M., Olarte, J., Espejo, R., Khin Maung, U., Ghafoor, M. A., Khan, M. A., Sami, Z. & Sutton, R. G. (1991). *Bull. World Health Organ.* **69**, 549–555.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–494.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sakellaris, H., Munson, G. P. & Scott, J. R. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 12828–12832.
Sakellaris, H. & Scott, J. R. (1998). *Mol. Microbiol.* **30**, 681–687.
Soto, G. E. & Hultgren, S. J. (1999). *J. Bacteriol.* **181**, 1059–1071.
Terwilliger, T. C. (2003). *Methods Enzymol.* **374**, 22–37.
Wenneras, C. & Erling, V. (2004). *J. Health Popul. Nutr.* **22**, 370–382.