Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of a core fragment of the flagellar hook protein FlgE

Fadel A. Samatey,[a,b,c]* Hideyuki Matsunami,[a,b,c] Katsumi Imada,[a,b,c] Shigehiro Nagashima[a,c] and Keiichi Namba[a,b,c]

[a]Dynamic NanoMachine Project, ICORP, JST, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan, [b]Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan, and [c]Protonic NanoMachine Project, ERATO, JST, 3-4 Hikaridai, Seika, Kyoto 619-0237, Japan

Correspondence e-mail: alexis@npn.jst.go.jp

Received 23 June 2004
Accepted 13 September 2004

A core fragment of the bacterial flagellar hook protein FlgE was overexpressed, purified and crystallized. The crystal diffracted to 1.6 Å resolution using synchrotron X-radiation. The crystal belongs to the orthorhombic crystal system, with space group $P2_12_12$ and unit-cell parameters $a = 128.4$, $b = 48.8$, $c = 96.7$ Å. SeMet protein was also overexpressed, purified, crystallized and a set of 2.3 Å MAD data was collected.

1. Introduction

For chemotaxis, bacteria have developed a highly sophisticated machinery called the flagellum, which enables them to swim toward favourable or away from unfavourable conditions in their living environment (Berg & Anderson, 1973). The flagellum can be divided into three parts: the basal body, which is a rotary motor embedded in the cell membrane, the hook, which is a short flexible tubular segment acting as a universal joint, and the filament, a long tubular structure that functions as a helical propeller. The basal body is a complex structure made up of about 20 proteins, but the hook and the filament are both made of single proteins, the hook protein (FlgE) and flagellin (FliC), respectively (Namba & Vonderviszt, 1997). The hook and the filament grow into the cell exterior in this order and these two tubular structures are connected to each other by two junction proteins, hook-associated proteins 1 (HAP1) and 3 (HAP3) (Homma *et al.*, 1990). The filament is capped with a HAP2 pentamer complex, which is essential for the filament growth by polymerization of flagellin at the distal end (Homma *et al.*, 1984; Ikeda *et al.*, 1987).

The hook of *Salmonella* has a relatively well defined length of 55 nm (Hirano *et al.*, 1994) and is made up of about 130 molecules of FlgE, the molecular weight of which is 42.2 KDa (Homma *et al.*, 1990). In order to understand the mechanism of the universal joint function of the hook, structural information on FlgE at high resolution is essential. However, FlgE is difficult to crystallize because of its tendency to polymerize or aggregate under most crystallization conditions. To overcome this problem, we constructed a core fragment of FlgE by removing the N-terminal 70 and the C-terminal 44 residues, which are unfolded in the monomeric form of FlgE in solution but are well structured in the hook for its mechanically stable structure formation (Vonderviszt *et al.*, 1992). The same strategy worked well with flagellin (Samatey *et al.*, 2000). The molecular weight of this fragment is 31.2 kDa and it was therefore named FlgE31. Here, we describe the overexpression, purification and crystallization of the native and selenomethionyl (SeMet) derivative protein of FlgE31 and complete high-resolution X-ray diffraction data collection from crystals of these two proteins.

2. Materials and methods

2.1. Expression and purification

The DNA sequence encoding FlgE31 was amplified by polymerase chain reaction (PCR) from *S. typhimurium* strain SJW1103 chromosomal DNA (Yamaguchi *et al.*, 1984) with the 5′ primer CCCATATGGGGCTGGATGTCGCGATTAGC, generating an *Nde*I site and a start codon for additional methionine, and the 3′ primer GAGGATCCTATTCACCAGTTATTTACTC, generating a stop codon and *Bam*HI site. The PCR fragment was digested with *Nde*I and *Bam*HI and ligated into a T7-expression vector pET3c (Novagen) to create pHMK611. The plasmid was transformed into *Escherichia coli* BL21(DE3)pLysS cells.

E. coli cells carrying the plasmid were grown in Luria–Bertani broth supplemented with 50 mg l^{-1} ampicillin and 30 mg l^{-1} chloramphenicol at 310 K to a cell density of OD = 0.7 and protein expression was then induced with 0.5 m*M* isopropyl thiogalactopyranoside (IPTG) at 310 K for 2 h. The cells were harvested by centrifugation at 4500*g* for 15 min. The cell pellet was suspended in 20 m*M* Tris–HCl pH 7.5 and disrupted by sonication. The solution was centrifuged at 10 000*g* for 15 min to remove the cell debris. The supernatant was centrifuged at 200 000*g* for 30 min and the supernatant was loaded

onto a Hi-Trap Q-Sepharose column (Amersham Pharmacia) equilibrated with 20 m*M* Tris–HCl pH 7.5. Elution was performed with the same buffer solution with a linear gradient of NaCl concentration from 0 to 0.5 *M*. The main peak fraction was dialyzed against 20 m*M* Tris–HCl pH 7.5, applied onto a Poros H20Q anion-exchange column (Applied Biosystems) and eluted with a linear gradient of NaCl concentration from 0 to 0.5 *M*. The fraction containing the protein was collected and the protein was concentrated to a final concentration of 20 mg ml^{-1} using a Centriprep centrifugal filter device (Amicon). During the concentration process, the buffer solution was replaced by Milli-Q water. The protein solution was kept at 277 K as a stock. The sample purity was checked by SDS–PAGE and the molecular mass of the purified protein was examined by matrix-assisted laser desorption ionization–time-of-flight (MALDI–TOF) mass spectrometry using a Voyager-DE/PRO spectrometer (Applied Biosystems). These experiments showed that the sample was pure, free from any protein degradation or contamination.

2.2. Crystallization and data collection

The initial screening of crystallization conditions was performed by the hanging-drop vapour-diffusion method at 289 K with two screening kits, Crystal Screen and Crystal Screen 2 from Hampton Research. The best result obtained from these trials was a crystal-like precipitate in 18% polyethyleneglycol (PEG) 8000, 0.1 *M* sodium cacodylate, 0.2 *M* zinc acetate (ZnAc) pH 6.5 (Fig. 1*a*). Reduction of the precipitant or protein concentration or different temperatures did not produce any significant improvements in crystallization. However, an improvement was observed when the molecular weight of PEG was reduced from PEG 8000 to PEG 2000. The crystals obtained were still small and stacked together (Fig. 1*b*). We were able to grow large thin plate-like crystals using the seeding technique, but these crystals were still stacked (Fig. 1*c*) and were not suitable for X-ray diffraction experiments. A dramatic improvement was observed when the ZnAc in the solution was replaced by copper acetate (CuAc). The stacking tendency was suppressed and the crystals became much thicker. Many crystals appeared in a few days in a solution containing 25 m*M* CuAc, 16% PEG 2000, 50 m*M* sodium cacodylate pH 4.5 (Fig. 1*d*). These crystals were crushed and the solution containing the crushed crystals was diluted and used for micro-seeding. Large thick crystals suitable for X-ray diffraction were obtained in a solution containing 12% PEG 2000, 3 m*M* CuAc, 50 m*M* sodium cacodylate pH 4.5 and at a protein concentration of 12 mg ml^{-1} (Fig. 1*e*). Crystals of the SeMet FlgE31 were obtained under the same conditions, but grew to a smaller size. To freeze the crystals for cryocrystallographic data collection, crystals were transferred into a solution containing 12% PEG 2000, 3 m*M* CuAc, 50 m*M* sodium cacodylate pH 4.5 and 20% glycerol for about 20 s, mounted in a cryo-loop and frozen in liquid propane.

X-ray diffraction data collection was carried out at 100 K. The crystal quality was assessed using our in-house X-ray facility, which consists of an FR-D rotating-anode X-ray generator equipped with Yale-type mirror optics and an R-AXIS IV image-plate detector (Rigaku). Even with this X-ray source of relatively high brilliance, a few hours of exposure time was needed to observe high-resolution reflections from the crystals. X-ray diffraction data from the native crystal were collected at beamline ID29 of the European Synchrotron Radiation Facility (ESRF). The X-ray wavelength used was 0.92 Å. Data from the SeMet derivative for multi-wavelength anomalous diffraction (MAD) phasing (Hendrickson, 1991) were collected at SPring-8 beamline BL41XU. Three data sets were collected from one crystal (Table 1).

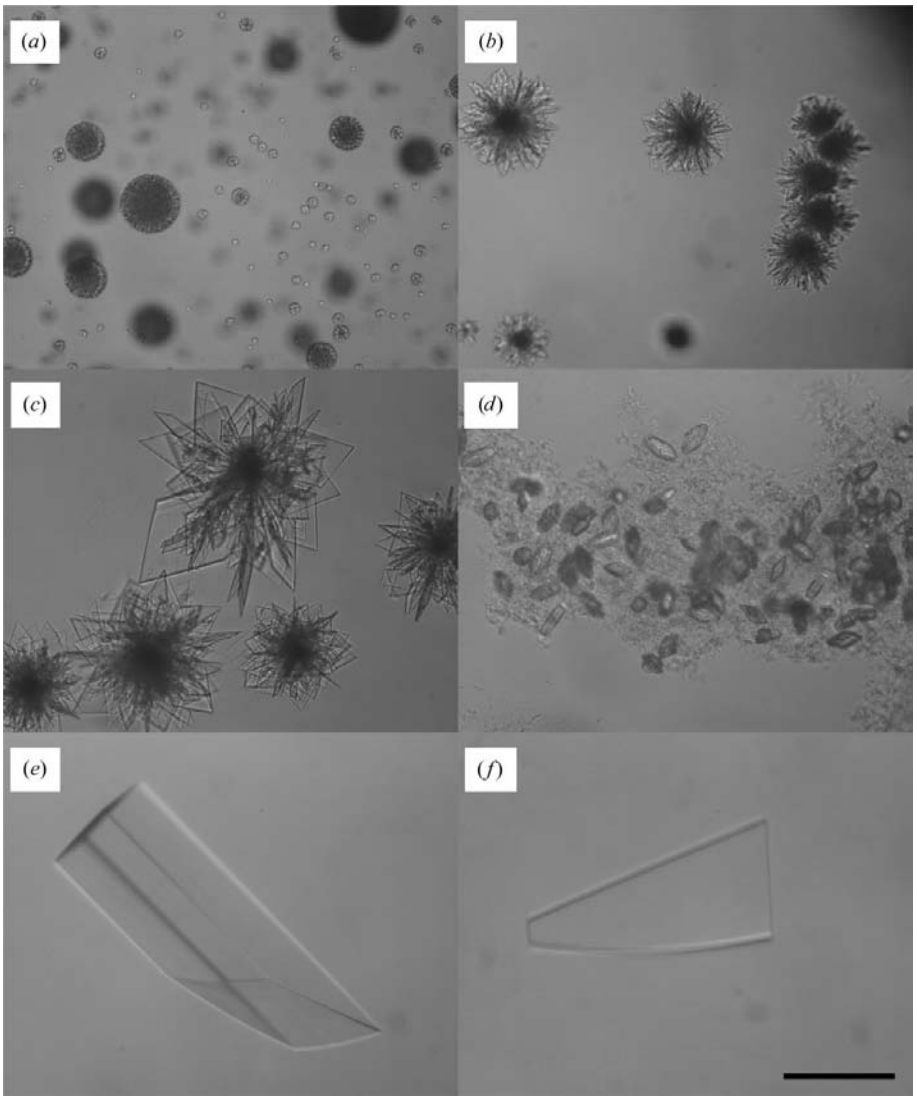

Figure 1
Crystals of FlgE31 from *Salmonella*. (*a*) Crystal-like precipitates formed in a solution containing zinc acetate and PEG 8000; (*b*) stacked crystals formed by replacing PEG 8000 with PEG 2000; (*c*) stacked thin-plate crystals formed by seeding; (*d*) thicker crystal obtained by replacing zinc acetate with copper acetate; (*e*) large single crystal obtained by microseeding; (*f*) single crystal obtained by seeding of crystals from (*b*) in a solution containing zinc acetate, glycerol and 2-propanol. Scale bar, 0.2 mm.

Table 1
Summary of data statistics.

Values in parentheses are for the last resolution shell.

	Native crystal	SeMet		
X-ray source	ID29-ESRF	BL41XU SPring-8		
Wavelength (Å)	0.92	0.9791	0.9794	0.9707
Space group	$P2_12_12$	$P2_12_12$		
Unit-cell parameters				
a	128.7	128.5	128.5	128.5
b	49.0	48.8	48.8	48.8
c	96.9	96.7	96.7	96.7
Resolution (Å)	1.8 (1.89–1.8)	2.3 (2.4–2.3)	2.3 (2.4–2.3)	2.3 (2.4–2.3)
Observed reflections	349273	239366	225105	230998
Unique reflections	56214	39914	39161	39467
Data completeness (%)	97.6 (91.7)	99.3 (98.0)	99.0 (97.3)	98.8 (97.4)
$R_{\rm merge}$† (%)	11.9 (23.3)	10.1 (27.0)	10.2 (30.4)	10.5 (33.0)
$R_{\rm ano}$‡ (%)		7.1 (13.2)	5.6 (15.1)	6.1 (14.6)

† $R_{\rm merge} = \sum|I_{\rm obs} - \langle I\rangle|/\sum\langle I\rangle$, where $I_{\rm obs}$ is an observed intensity and $\langle I\rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections. ‡ $R_{\rm ano} = \sum|I_+ - I_-|/\sum|I_+ + I_-|$, where I_+ and I_- are intensity measurements of anomalous pairs.

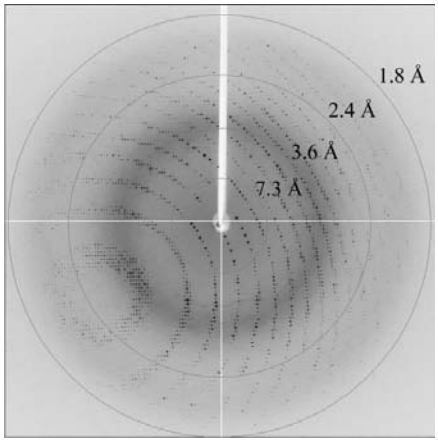

Figure 2
Diffraction pattern from the native crystal of FlgE31.

3. Results and discussion

For the crystal grown in the presence of CuAc, a diffraction pattern from the native crystal is shown in Fig. 2. The native data were processed with *MOSFLM* (Leslie, 1992) and scaled with *SCALA* from the *CCP*4 software suite (Collaborative Computational Project, Number 4, 1994). The MAD data were processed and scaled using the *HKL*2000 program package (Otwinowski & Minor, 1997). Data reduction and a search for systematic absences of reflections indicated that the FlgE31 crystal belongs to space group $P2_12_12$. Assuming two FlgE31 molecules per asymmetric unit, the Matthews coefficient $V_{\rm M}$ was calculated to be 2.4 Å^3 Da^{-1} (Matthews, 1968), which corresponds to a solvent content of 48%. Ten Se sites were found using the *SOLVE* program (Terwilliger & Berendzen, 1999) and the figure of merit of the phases was 0.48 between 20.0 and 2.3 Å resolution. The phases were extended to 1.8 Å with the density-modification program *DM* (Cowtan, 1994; Collaborative Computational Project, Number 4, 1994). The atomic model building of FlgE31 based on the density map at 1.8 Å resolution will soon be published.

Metal ions are often contained in kit solutions for protein crystallization or are sometimes specifically added in crystallization attempts. However, their effects on the crystallization process are not predictable. In the initial crystallization screening of FlgE31, the zinc ion present in a crystallization kit solution was found to be effective, but we could not improve the crystal quality using solutions containing zinc. Because the zinc ion has a similar size to the copper ion but different chemical properties, we opted to substitute ZnAc with CuAc in crystallization solutions. To fully understand the effect of the copper ion in the crystallization of FlgE31 and to compare it with that of the zinc ion, it is necessary to also grow high-quality crystals of FlgE31 in a ZnAc-containing solution and to solve the structures of both crystals to determine the difference in the ion–protein coordination. Long after the copper-type crystals of FlgE31 had been obtained, we were able to produce good-quality crystals of FlgE31 in 12% PEG 2000, 50 m*M* ZnAc, 12% glycerol, 10% 2-propanol and 50 m*M* sodium cacodylate pH 4.5 (Fig. 1*f*). Preliminary X-ray diffraction data from these crystals show that they belong to space group *C*2, with unit-cell parameters a = 275.3, b = 56.6, c = 77.6 Å, β = 96.4°. This work is still in progress.

We thank the staff members of beamline ID29 at ESRF in Grenoble and beamline BL41XU at SPring-8 in Harima for their help in data collection.

References

Berg, H. C. & Anderson, R. A. (1973). *Nature (London)*, **245**, 380–382.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACMB Newsl. Protein Crystallogr.* **31**, 34–38.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Hirano, T., Yamaguchi, S., Oosawa, K. & Aizawa, S.-I. (1994). *J. Bacteriol.* **176**, 5439–5449.
Homma, M., DeRosier, D. J. & Macnab, R. M. (1990). *J. Mol. Biol.* **213**, 819–832.
Homma, M., Kutsukake, K., Iino, T. & Yamaguchi, S. (1984). *J. Bacteriol.* **157**, 100–108.
Ikeda, T., Homma, M., Iino, T., Asakura, S. & Kamiya, R. (1987). *J. Bacteriol.* **169**, 1168–1173.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACMB Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Namba, K. & Vonderviszt, F. (1997). *Q. Rev. Biophys.* **30**, 1–65.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Samatey, F. A., Imada, K., Vonderviszt, F., Shirakihara, Y. & Namba, K. (2000). *J. Struct. Biol.* **132**, 106–111.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Vonderviszt, F., Ishima, R., Akasaka, K. & Aizawa, S.-I. (1992). *J. Mol. Biol.* **226**, 575–579.
Yamaguchi, S., Fujita, H., Sugata, K., Taira, T. & Iino, T. (1984). *J. Gen. Microbiol.* **130**, 255–265.