Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic analysis of FlhD/FlhC complex from *Escherichia coli*

Shu-ying Wang,[a] Philip Matsumura[a,b]* and Edwin M. Westbrook[b]

[a]Department of Microbiology and Immunology, University of Illinois, Chicago, IL 60612-7344, USA, and [b]Molecular Biology Consortium, Chicago, IL 60612, USA

Correspondence e-mail: philip.matsumura@uic.edu

Received 21 October 2000
Accepted 16 February 2001

The heterotetrameric (C_2D_2) FlhD/FlhC complex was first discovered as a transcriptional activator of the flagellar genes in *Escherichia coli*. Recent studies now show that FlhD/FlhC also regulates several non-flagellar target genes in *E. coli*. The FlhD/FlhC complex also plays several important roles in other microorganisms. The molecular interactions between FlhD and FlhC, as well as the mechanisms by which the complex may vary its DNA-binding specificity, are not clear. Determination of the FlhD/FlhC crystal structure will provide insight into these protein–protein and protein–DNA interactions. The initial steps in this investigation are reported here: the overexpression, purification and crystallization of the FlhD/FlhC complex, the characterization of this crystal form and the recording and processing of an initial diffraction data set. The obtained crystal form of the FlhD/FlhC complex is hexagonal (space group $P6_1$, unit-cell parameters $a = b = 150.5$, $c = 115.9$ Å). The crystal density is very low ($V_M = 5.5$), with 81.7% of its volume occupied by solvent. A single C_2D_2 tetramer is present in the crystallographic asymmetric unit. A complete native data set has been collected to 4.5 Å resolution.

1. Introduction

The transcription of flagellar genes in *E. coli* is controlled in a regulatory hierarchy (Macnab, 1992). At the top of the hierarchy is the *flhDC* master operon, which is composed of two genes, *flhD* and *flhC*. Two molecules each of FlhD (116 residues; MW 13 317 Da) and FlhC (192 residues; MW 21 544 Da), the proteins expressed by *flhD* and *flhC*, respectively, form a tetrameric (C_2D_2) complex (Liu & Matsumura, 1994) that is required for the expression of all other flagellar genes (Bartlett *et al.*, 1988). It has recently been observed that the FlhD/FlhC complex also regulates several non-flagellar genes in *E. coli* (Prüss *et al.*, 2001, in the press). These studies show that the FlhD/FlhC complex is more appropriately viewed not just as a flagellar transcriptional activator but rather as a global regulator involved in many stationary-phase processes. These include the TCA cycle, anaerobic respiration and the synthesis of transporters and rod-shape determination proteins. FlhD, but not FlhC, regulates the cell-division rate in the stationary phase through the *cadBA* operon (Prüss *et al.*, 1997). Regulatory actions of the FlhD/FlhC complex have also been observed in bacteria other than *E. coli*. For example, the *flhD* operon controls the expression and secretion of virulence-associated phospholipase in both *Serratia liquefaciens* (Givskov *et al.*, 1995) and *Yersinia enterocolitica* (Young *et al.*, 1999). In the major insect pathogen *Xenorhabdus nematophilus*, the FlhD/FlhC complex is an important global regulator that affects the expression of virulence factors in the stationary phase and is required for full virulence (Givaudan & Lanois, 2000). In *Salmonella enterica* serovar *typhimurium*, FlhD/FlhC modulates *hilA* expression and this regulation pathway is integrated into the regulatory network of type III secretion systems (Lucas *et al.*, 2000). In *Proteus mirabilis*, the *flhDC* operon is central to differentiation into elongated hyperflagellated swarm cells (Claret & Hughes, 2000*b*). All of these studies suggest that FlhD/FlhC plays important and complex roles in various microorganisms. Interestingly, all roles of the complex so far observed have related in some way to the adaptation of each organism to environmental conditions. To date, FlhD/FlhC is the first and the only heteromeric transcriptional factor found in prokaryotes.

The molecular interactions between FlhD and FlhC are not currently understood, nor are the mechanisms by which this complex is apparently able to modify its DNA-binding specificity. The crystal structure of FlhD by itself has been determined (Campos *et al.*, 1998, 2001). It is a homodimer, predominantly helical, with a tightly folded core. The structures of the 34 C-terminal residues of both

FlhD chains are poorly defined in this crystal structure. Crystal packing disrupted most of this region, while those residues of the C-termini that were observed in the crystal structure exhibited high temperature factors. A putative helix–turn–helix (HTH) DNA-binding motif has been identified in this region of the FlhD structure, but only the first helix of one FlhD monomer was observed in the crystal structure. FlhD alone cannot bind to the promoter of class II flagellar genes (Liu & Matsumura, 1994), yet mutagenesis analysis suggests that it is FlhD that binds to DNA in the FlhD/FlhC complex (Campos & Matsumura, 2001). Therefore, it was initially proposed that the FlhD C-termini are weakly folded and mobile, that association with FlhC may stabilize these C-termini and that the stable FlhD structure includes a well defined DNA-binding HTH motif. A recent suggestion has been made that FlhC may be the DNA-binding protein, with FlhD acting as the helper protein to enhance DNA-binding affinity (Claret & Hughes, 2000*a*). Therefore, determination of the FlhD/FlhC complex crystal structure will permit us to test both hypotheses and will be valuable for our further general understanding of macromolecular recognition processes. In this study, we report the overexpression, purification, crystallization and preliminary X-ray crystallographic analysis of the *E. coli* FlhD/FlhC complex.

2. Materials and methods

2.1. Protein expression and purification

The *E. coli flhDC* gene was digested from pXL27 (Liu & Matsumura, 1994) and subcloned into the pET24a vector (Novagen) in *Nde*I/*Hin*dIII sites. B834(DE3) *E. coli* cells transformed with a pSW28 (pET24a-*flhDC*) plasmid were grown at 310 K in LB medium containing 50 µg ml^{-1} kanamycin. Expression of *flhDC* was induced by adding 1 m*M* IPTG at an OD_{600} of 0.6. After 3 h induction, cells were harvested, resuspended in buffer *A* (20 m*M* Tris–HCl pH 7.9) and lysed by sonication. The cell lysate was clarified by centrifugation at 31 000*g* for 30 min. The supernatant was loaded onto a 5 ml HiTrap heparin column (Pharmacia) which had been pre-equilibrated with buffer *A*. The unbound protein was washed out with buffer *A* containing 0.2 *M* NaCl and then a linear NaCl gradient (0.2–1.0 *M* in buffer *A*) was applied. FlhD/FlhC complex eluted at a salt concentration of about 0.5 *M* NaCl. The eluted fractions were pooled and dialyzed against 0.2 *M* NaCl in 100 m*M* Tris–HCl pH 7.9. The protein complex was concentrated to 30 mg ml^{-1} using Centriprep and Centricon 30K (Amicon) in preparation for crystallization. The protein is at least 95% pure as estimated by analysis of an SDS–polyacrylamide gel (Fig. 1).

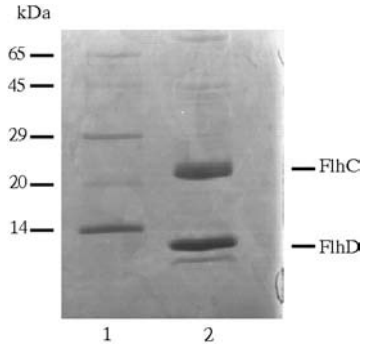

Figure 1
The purity of the FlhD/FlhC complex was analyzed on 15% SDS–polyacrylamide gel. Lane 1, molecular markers; lane 2, purified FlhD/FlhC complex. The molecular weights are 13.6 and 22 kDa for FlhD and FlhC, respectively.

2.2. Crystallization

Crystals were grown by the hanging-drop vapor-diffusion method. A large number of conditions were screened with Crystal Screen I and II, Natrix formulation and additive screen kits (Hampton Research). Each drop was prepared by mixing 2 µl protein complex (30 mg ml^{-1} in 100 m*M* Tris–HCl pH 7.9 containing 0.2 *M* NaCl) and 2 µl reservoir solution. Each reservoir volume was 0.7 ml. Hexagonal rod-shaped crystals grew at room temperature from reservoir solutions containing 1–2% polyethyleneimine, 0.1 *M* sodium citrate pH 5.6 and 0.5 *M* NaCl. We optimized the crystal-growth conditions and the best crystals were obtained in drops containing 2 µl protein-complex solution, 1 µl reservoir solution and 1 µl 30% ethanol. Crystals grew to dimensions of approximately 0.15 × 0.15 × 1.0 mm within a week (Fig. 2).

3. Data collection and analysis

Diffraction data were collected at beamline 17ID of the Advanced Photon Source, Argonne National Laboratory, Argonne, IL, USA (IMCA-CAT) and at beamline 5.0.2 of the Advanced Light Source, Berkeley, CA, USA (Berkeley Center for Structural Biology). To minimize radiation damage, the crystals were cryoprotected by rapid immersion in liquid nitrogen (Rogers, 1997). To permit freezing, crystals were transferred briefly to a cryoprotectant solution containing 30% glycerol in addition to the crystallization buffer solutes. Crystals were kept at about 100 K in a cold nitrogen stream during data collection.

At APS beamline 17ID the MAR165 detector was 300 mm from the crystal sample and the wavelength was λ = 0.9919 Å. 100 data frames, each of 1° rotational width and 30 s exposure, were recorded. Data in the resolution range 14.0–6.0 Å were processed with the programs *d*TREK* (Pflugrath, 1999) and *XGEN* (Howard *et al.*, 1987). Data extended to only 6.0 Å *d* spacing (*B* factor from *XGEN*, 148 Å^2). The data exhibited sixfold rotational symmetry and no mirror symmetry normal to the sixfold. The only reflections on the *l* axis above background were (0, 0, 12) and (0, 0, 18). Therefore, the space group is either $P6_1$ or $P6_5$. The unit-cell parameters are $a = b = 152.8$, $c = 115.8$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. A total of 18 676 observations were processed and merged to 3 593 unique reflections with $R_{\rm merge} = 10.4\%$. The merged data set is 99.4% complete in the resolution range 14–6 Å.

At ALS beamline 5.0.2 the ADSC Q4 detector was 171 mm from the crystal sample and the wavelength was λ = 1.5498 Å. 120 data frames, each of 1° rotational width and 120 s exposure, were recorded. Data in the resolution range 20.0–4.5 Å were processed with the program *HKL*2000 (Otwinowski & Minor, 1997) as space group *P*6, but examination of the (0, 0, *l*) line confirmed that only $l = 6n$ reflections are non-zero, consistent with the assignment of space group $P6_1$ (or $P6_5$). The unit-cell parameters refined as $a = b = 150.5$, $c = 115.9$ Å, slightly different to those from the previous data set. A total of 63 150 observations were processed and merged to 8 958 unique reflections with $R_{\rm merge} = 9.2\%$.

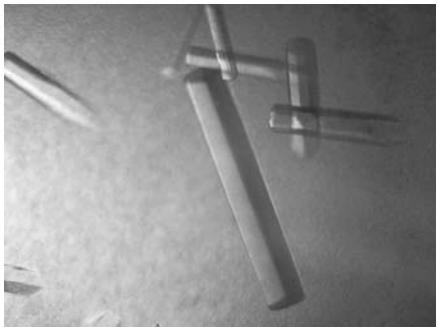

Figure 2
Hexagonal rod-shaped crystals of the FlhD/FlhC complex from *E. coli* were grown at room temperature. The approximate dimensions of the largest crystal in the center are 0.15 × 0.15 × 1.0 mm.

The merged data set is 97.5% complete in the 20–4.5 Å resolution range (89.0% in the outer shell). The Debye–Waller temperature factor, calculated by a Wilson plot (Wilson, 1949), is B = 231 Å^2, consistent with the large value of the previously calculated B factor.

It is well known that B factors calculated from low-resolution data do not necessarily predict the dependence of $\langle I \rangle$ on resolution. In practical terms, extending the data set from 6 to 4.5 Å (a 2.5-fold increase in the number of observed unique data) required quadrupling the exposure time (30 to 120 s) of each data image. The Bragg intensities decrease rapidly with diffraction angle and there is no apparent inflection in this drop-off at higher angles. Although we expect to be able to extend to 4.0 Å minimum d spacing in the next data-collection round, the high B-factor value clearly indicates structural disorder within the crystal lattice that will prevent us from observing significantly higher diffraction angles from this crystal form.

The self-rotation function calculated with *AMoRe* (Navaza, 1994) exhibits just one non-crystallographic symmetry: a dyad, perpendicular to the crystallographic sixfold axis and 36.9° from the crystallographic *a* axis. The self-Patterson map does not contain any large peaks that might indicate the presence of parallel dyad rotational axes in the crystal packing. This finding is thus consistent only with a single heterotetramer of approximate molecular mass 72 kDa per asymmetric unit, containing an internal twofold symmetry. The implied solvent content of this crystal form is 81.7% and the Matthews coefficient is 5.48 Å^3 Da^{-1}.

We have now grown crystals of selenomethionyl-substituted FlhD/FlhC, but these crystals are not as good as the native crystals. We are therefore searching for heavy-atom derivatives (platinum, mercury, lanthanides *etc*.) formed by random soaks. It is our intention to solve the structure of this complex at least to 4.0 Å d spacing by the multiple-energy anomalous dispersion (MAD) method. There are two goals in solving this admittedly low-resolution structure: to determine the folding topology of FlhC in the complex and to determine whether, in the complex, the FlhD C-termini display well ordered helix–turn–helix motifs for DNA binding. We are also crystallizing FlhD/FlhC complexes from related bacteria in the hope of finding a more densely packed crystal form that would diffract to higher angle. However, it is important to obtain the baseline tetrameric protein structure before beginning our next task: co-crystallizing the tetramer with its cognate DNA.

Use of the Advanced Photon Source and the Advanced Light Source was supported by the US Department of Energy Office of Basic Energy Science. We wish to thank Dr Andrew Howard and his colleagues in the Industrial Macromolecular Crystallography Association Collaborative Access Team (IMCA-CAT) at the Advanced Photon Source. These facilities are supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with Illinois Institute of Technology (IIT) executed through the IIT center for Synchrotron Radiation Research and Instrumentation. We wish to thank Dr Gerry McDermott and his colleagues in the Berkeley Center for Structural Biology (BCSB) at ALS beamline 5.0.2. The BCSB is supported by the US Department of Energy Office of Biological and Environmental Research. This work has been supported by grant 5R01GM59484-02 from the National Institutes of Health.

References

Bartlett, D. H., Frantz, B. B. & Matsumura, P. (1988). *J. Bacteriol.* **170**, 1575–1581.
Campos, A. & Matsumura, P. (2001). *Mol. Microbiol.* **39**(3), 581–594.
Campos, A., Matsumura, P. & Volz, K. (1998). *J. Struct. Biol.* **123**, 269–271.
Campos, A., Zhang, R. G., Alkire, R. W., Matsumura, P. & Westbrook, E. M. (2001). *Mol. Microbiol.* **39**(3), 567–580.
Claret, L. & Hughes, C. (2000*a*). *J. Mol. Biol.* **303**, 467–478.
Claret, L. & Hughes, C. (2000*b*). *J. Bacteriol.* **182**, 833–836.
Givaudan, A. & Lanois, A. (2000). *J. Bacteriol.* **182**, 107–115.
Givskov, M., Eberl, L., Christiansen, G., Benedik, M. J. & Molin, S. (1995). *Mol. Microbiol.* **15**, 445–454.
Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H. & Salemme, F. R. (1987). *J. Appl. Cryst.* **20**, 383–387.
Liu, X. & Matsumura, P. (1994). *J. Bacteriol.* **176**, 7345–7351.
Lucas, R. L., Lostroh, C. P., DiRusso, C. C., Spector, M. P., Wanner, B. L. & Lee, C. A. (2000). *J. Bacteriol.* **182**, 1872–1882.
Macnab, R. M. (1992). *Annu. Rev. Genet.* **26**, 131–158.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Prüss, B. M., Liu, X., Hendrickson, W. & Matsumura, P. (2001). In the press.
Prüss, B. M., Markovic, D. & Matsumura, P. (1997). *J. Bacteriol.* **179**, 3818–3821.
Rogers, D. W. (1997). *Methods Enzymol.* **276**, 183–203.
Wilson, A. J. C. (1949). *Acta Cryst.* **3**, 397–398.
Young, G. M., Schmiel, D. H. & Miller, V. L. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 6456–6461.