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

Preparation and preliminary X-ray diffraction analysis of crystals of bacterial flagellar sigma factor σ^{28} in complex with the σ^{28}-binding region of its antisigma factor, FlgM

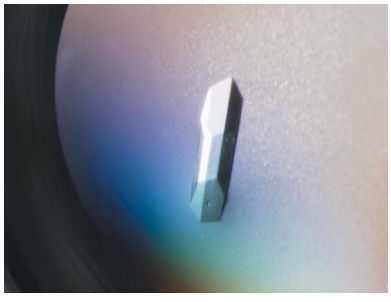
Kengo Okada,[a] Hisako Ichihara,[a]‡ Hiroyuki Takahashi,[a] Nobuyuki Fujita,[b]§ Akira Ishihama[b]¶ and Toshio Hakoshima[a,c]*

[a]Structural Biology Laboratory, Nara Institute of Science and Technology, Keihanna Science City, Nara 630-0192, Japan, [b]Department of Molecular Genetics, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan, and [c]CREST, Japan Science and Technology Agency, Keihanna Science City, Nara 630-0192, Japan

‡ Present address: Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan.
§ Present address: National Institute of Technology and Evaluation, Genome Analysis Center, Nishihara, Shibuya, Tokyo 151-0066, Japan.
¶ Present addresses: Hosei University, Faculty of Engineering and Research Center for Micro-Nano Technology, Koganei, Tokyo 184-8584, Japan and Nippon Institute for Biological Science, Division of Molecular Biology, Ome, Tokyo 198-0024, Japan.

Correspondence e-mail: hakosima@bs.naist.jp

Received 22 December 2006
Accepted 31 January 2007



The sigma 28 kDa (σ^{28}) factor is a transcription factor specific for the expression of bacterial flagellar and chemotaxis genes. Its antisigma factor, FlgM, binds σ^{28} factor and inhibits its activity as a transcription factor. In this study, crystals of the complex between *Escherichia coli* σ^{28} and the C-terminal σ^{28}-binding region of FlgM were obtained. The crystals belong to space group $P3_121$ or $P3_221$, with unit-cell parameters $a = b = 106.7$ (2), $c = 51.74$ (3) Å, containing one complex in the crystallographic asymmetric unit. An X-ray intensity data set was collected to a resolution of 2.7 Å.

1. Introduction

Transcription is facilitated by RNA polymerase, which has a sequence-specific affinity for the regions 5′ to coding sequences, which are known as promoters. Bacterial RNA polymerase is a hetero-multimeric complex of five essential protein subunits. Four subunits, $\alpha_2\beta\beta'$, form a stable complex known as core RNA polymerase, although this complex exhibits no specific binding to the promoter. The promoter recognition requires another subunit, represented by sigma factors, which binds core RNA polymerase to form the holoenzyme. Bacteria possess two families of sigma factors, σ^{70} and σ^{54}. In the σ^{70} family, multiple sigma factors are present and share high sequence homology at the amino-acid level. These sigma factors display different promoter specificities, providing one mechanism for global regulation of transcription. A set of genes can be coordinately expressed if their promoter sequences are recognized by a single sigma factor.

Regulation of transcription in the flagella biosynthetic pathway organizes the flagellar genes into a transcriptional hierarchy of three classes: class 1, class 2 and class 3 (Kutsukake *et al.*, 1990). σ^{28} is a flagellar-specific alternative sigma factor that belongs to the σ^{70} family and is necessary for transcription from class 3 promoters (Liu & Matsumura, 1995). Interestingly, FlgM directly binds σ^{28} to prevent its interactions with core RNA polymerase and with class 3 promoter DNAs (Kutsukake & Iino, 1994). Thus, FlgM is the negative regulatory protein for class 3 gene expression and is referred to as an anti-σ^{28} factor. FlgM is responsible for sensing completion of the hook–basal body (HBB) structure in order for class 3 promoters to be transcribed. The interaction between σ^{28} and FlgM has been investigated using NMR spectroscopy, which has shown that FlgM was mostly unstructured by itself in solution. Interestingly, the C-terminal half of FlgM is responsible for σ^{28} binding and the adoption of a structured form in the presence of σ^{28} (Daughdrill *et al.*, 1997). This finding suggests the biological importance of unfolded protein (Plaxco & Gross, 1997).

Bacteria utilize different classes of antisigma factors (Hughes & Mathee, 1998). Studies in recent years have investigated the crystal structures of bacterial core RNA polymerase and holoenzyme from *Thermus aquaticus* (Zhang *et al.*, 1999; Murakami, Masuda, Campbell *et al.*, 2002; Murakami, Masuda & Darst, 2002) and bacterial sigma factors, including antisigma factors such as SpoIIB and RseA (Campbell *et al.*, 2002, 2003). These structural investigations have revealed different modes of interactions of sigma factors with core enzymes and antisigma factors and also variable interactions between

each sigma factor and its antisigma factor. Although the amino-acid sequences of σ^{28} factors are conserved with more than 35% identity, those of anti-σ^{28} factors display less than 20% identity. This may provide the basis for versatility of interactions between σ^{28} factors and anti-σ^{28} factors in different species of bacteria. In this study, we investigated crystals of *Escherichia coli* σ^{28} factor in complex with the σ^{28}-binding region of its anti-σ^{28} factor, FlgM.

2. Methods

2.1. Cell culture

The expression plasmid for full-length σ^{28}, pETSF (Kundu *et al.*, 1997), was transformed into *E. coli* strain BL21Star(DE3) (Invitrogen). The cells were grown at 310 K in Luria–Bertani (LB) medium containing 50 µg ml^{-1} ampicillin. When cultures reached an OD_{660} of 0.8, isopropyl β-D-thiogalactopyranoside (IPTG) was added to a concentration of 0.1 m*M* to induce expression of the gene. Cultures were grown for an additional 3 h after IPTG induction at 310 K. The cells were then harvested by centrifugation at 6000 rev min^{-1} (Beckman J2-M1 JA10 rotor) for 20 min at 277 K. *E. coli* FlgM consists of 97 amino-acid residues. The cDNA sequence corresponding to residues 47–97 of *E. coli* FlgM was amplified by PCR and cloned into a pET vector using *Nde*I and *Xho*I restriction sites, a His_6 tag being fixed at the carboxyl-terminus by means of a two-residue (Leu-Glu) tether (Dubendorff & Studier, 1991). The absence of errors in the sequence was confirmed by DNA sequencing. The expression plasmid for the FlgM fragment was transformed into *E. coli* strain BL21(DE3) (Invitrogen). When cultures reached an OD_{660} of 0.6 at 310 K, IPTG was added to a concentration of 1 m*M*. Cultures were grown for an additional 5 h after IPTG induction at 298 K. The cells were then harvested in the manner used for σ^{28}.

For multiwavelength anomalous dispersion (MAD) experiments, expression plasmids for both σ^{28} and the σ^{28}-binding region of FlgM were introduced into the *E. coli* methionine auxotroph B834(DE3)pLysS (Novagen). The selenomethionine-substituted (SeMet) proteins of σ^{28} and the σ^{28}-binding region of FlgM were induced in a manner similar to the native proteins, with the exception that LeMaster medium (Hendrickson *et al.*, 1990) was used to guarantee complete substitution of the methionines with selenomethionine. However, the expression levels of the genes for both SeMet proteins were very low. Next, the protocol for selenomethionine incorporation by metabolic inhibition was attempted (Van Duyne *et al.*, 1993; Doublié, 1997). This technique is based on the inhibition of methionine biosynthesis and does not need the use of a methionine-auxotroph strain, requiring neither transformation nor transduction. Cells were grown to the mid-log phase before addition of the amino acids needed for the inhibition of methionine biosynthesis. Induction by IPTG to a concentration of 0.1 m*M* was performed 30 min after addition of the amino acids. Cultures were grown for an additional 5 h after IPTG induction. The cells were then harvested in the manner used for native proteins.

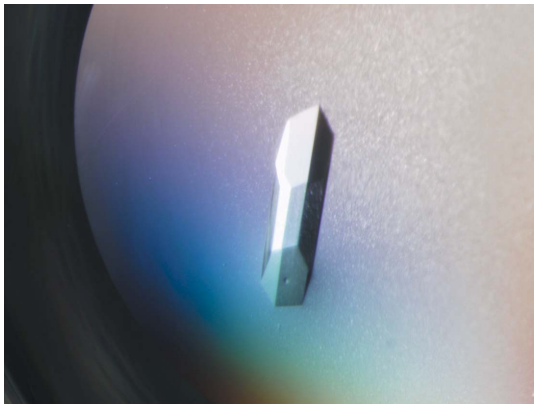


Figure 1
A native crystal of full-length σ^{28} in complex with the σ^{28}-binding region of FlgM. The approximate dimensions of this crystal are 80 × 120 × 400 µm.

2.2. Purification

All steps of protein purification were carried out at 277 K. Initially, full-length σ^{28} and the σ^{28}-binding region of FlgM were purified and concentrated separately. However, the purified σ^{28}-binding region of FlgM aggregated immediately in solution. Therefore, the protocol was changed to co-purify σ^{28} with the σ^{28}-binding region of FlgM as a protein complex because σ^{28} has a high binding affinity for FlgM (K_d = ~2 × 10^{-10} *M*; Chadsey *et al.*, 1998). Cell pellets for each protein were mixed and suspended in 20 volumes of ice-cold lysis buffer (50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl). The mixed pellets were homogenized with a BeadBeater (BioSpec Products). The crude lysate was centrifuged at 19 000 rev min^{-1} (Beckman J2-M1 JA20 rotor) for 50 min at 277 K. The supernatant fraction was then collected and filtered before being loaded onto an Ni-chelating column (Amersham Biosciences) equilibrated with a buffer containing 100 m*M* Tris–HCl pH 8.0 and 300 m*M* NaCl. The column was washed with five bed volumes of the equilibration buffer to remove nonspecific binding proteins and the sample was then eluted with a step-wise imidazole gradient (50–500 m*M*). This eluate containing the complex was concentrated to a volume of 2 ml and applied onto a Superdex 75 pg gel-filtration column (Amersham Biosciences) equilibrated with a buffer containing 10 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl and 0.1 m*M* EDTA. The presence of both full-length σ^{28} and the σ^{28}-binding region of FlgM in the eluate fraction containing the complex was confirmed by polyacrylamide gel electrophoresis in the presence of 0.1%(*w*/*v*) sodium dodecyl sulfate (SDS–PAGE) and this fraction was concentrated to 35 mg ml^{-1}.

N-terminal sequence analysis (M492, Applied Biosystems), matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy (MALDI–TOF MS; PerSpective) and dynamic light-scattering measurements (DynaPro-801, Protein Solutions) were performed to validate the purity and stability of the resulting sample.

2.3. Crystallization

Crystals were obtained at 293 K using the sitting-drop vapour-diffusion technique in six-depression glass plates (Waken, Japan). The protein solution contained 100 m*M* Tris–HCl pH 8.5, 100 m*M* NaCl, 1 m*M* σ^{28} and 1 m*M* of the σ^{28}-binding fragment of FlgM. A reservoir solution of 18%(*w*/*v*) polyethylene glycol 6000 (PEG 6000), 100 m*M* Tris–HCl pH 8.5 and 200 m*M* Li_2SO_4 was used. The protein solution was mixed in a 1:1 ratio with the reservoir solution. Clusters of very small needle crystals (maximum dimensions 5 × 5 × 50 µm) appeared in a week and some of the crystals grew in a month to typical dimensions of 80 × 120 × 400 µm.

2.4. Data collection and processing

All data were collected from frozen crystals at 100 K. Before freezing the crystals in a stream of nitrogen gas, the crystal for data collection was transferred into a cryoprotective solution containing 20%(*v*/*v*) PEG 200 and the reservoir solution [18%(*w*/*v*) PEG 6000, 100 m*M* Tris–HCl pH 8.5 and 200 m*M* Li_2SO_4] and then mounted on

Table 1
Crystal data and data-collection statistics of the native crystal.

Values in parentheses are for the last shell.

X-ray source	BL41XU, SPring-8
Wavelength (Å)	1.0000
Detector	MAR CCD
Oscillation angle (°)	1.0
Exposure time (s)	5.0
Distance (mm)	180
No. of frames	135
Temperature (K)	100
Resolution range (Å)	20–2.9 (3.06–2.90)
Space group	$P3_121$ or $P3_221$
Unit-cell parameters (Å)†	$a = b = 106.7$ [2], $c = 51.74$ [3]
Observed reflections	61026
Unique reflections	7540
R_{sym}‡	0.077 (0.316)
$\langle I/\sigma(I)\rangle$	7.8 (2.3)
Completeness	0.982 (0.982)
Multiplicity	8.1
Mosaicity (°)	0.6

† The values in square brackets are standard deviations. ‡ $R_{sym} = \sum_{h,i} |I_{h,i} - \langle I_h\rangle| / \sum_{h,i} I_{h,i}$, where $I_{h,i}$ is the intensity of the ith measurement of reflection h and $\langle I_h\rangle$ is the mean value of $I_{h,i}$ for all i measurements.

a rayon loop. Intensity data were collected at SPring-8. All data were processed and scaled with *MOSFLM* and *SCALA* (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

Optimization of the protein-purification protocol described above was crucial for successful crystallization because although full-length σ^{28} and the σ^{28}-binding region of FlgM were highly purified separately, the latter aggregated immediately in the σ^{28}-free state. An additional gel-filtration column-chromatography step was also needed to remove other aggregated proteins. As a result of the modified purification protocol, a more stable protein complex was consistently obtained and single crystals grew within a month with typical dimensions of 80 × 120 × 400 µm (Fig. 1). SDS–PAGE and MALDI–TOF MS analysis confirmed that the crystals contained both full-length σ^{28} and the σ^{28}-binding region of FlgM. The gel showed that the crystals contain both proteins (Fig. 2). MALDI–TOF MS peaks of 27 506 Da (calculated molecular weight of 27 521 Da) and 6789 Da (calculated weight of 6777 Da) were observed for σ^{28} and the σ^{28}-binding region of FlgM, respectively. The crystals diffracted X-rays to a resolution of 2.7 Å (Fig. 3) and belong to space group $P3_121$ or $P3_221$, with unit-cell parameters $a = b = 106.7$ (2), $c = 51.74$ (3) Å. A complete data set was collected to a resolution of 2.9 Å using a single crystal. Table 1 summarizes the detailed statistics for data collection. A Matthews coefficient (Matthews, 1968) of 2.4 Å^3 Da^{-1} was calculated assuming one protein complex with an approximate molecular weight of 35 k Da in the asymmetric unit, which corresponds to 49% solvent content by volume.

In the early stages of this project, we tried to crystallize the complex between full-length FlgM and σ^{28} factor. However, these attempts were unsuccessful. The N-terminal half of full-length FlgM is flexible even in the complex with σ^{28} factor (Daughdrill *et al.*, 1997), which probably made it difficult to crystallize the complex. Recently, the crystal structure of the complex between *Aquifex aeolicus* σ and FlgM has been reported (Serenson *et al.*, 2004). This structure contains a full-length FlgM and most of the FlgM polypeptide was found to be structured, which is contradictory to the NMR and biochemical data (Daughdrill *et al.*, 1997). Our structure could shed light on this contradiction.

To facilitate structure determination of the protein complex between full-length σ^{28} and the σ^{28}-binding region of FlgM using the MAD method, SeMet proteins for both full-length σ^{28} and the σ^{28}-binding region of FlgM were prepared and crystallized. Although the preparation using the *E. coli* methionine-auxotroph strain failed, selenomethionine incorporation by metabolic inhibition was successful and sufficient sample for crystallization was obtained. Validation of the success of the substitution of methionine for selenomethionine was obtained using MALDI–TOF MS (PerSpective). Peaks of 27 737 Da (calculated molecular weight of 27 755 Da) and 6917 Da (calculated weight of 6917 Da) were observed for the SeMet proteins σ^{28} and the σ^{28}-binding region of FlgM, respectively. Crystallization conditions similar to the native protein complex produced very small single crystals, with maximum dimensions of 30 × 30 × 250 µm. Data collection of SeMet protein crystals and an associated structure determination using MAD data is in progress. The crystal

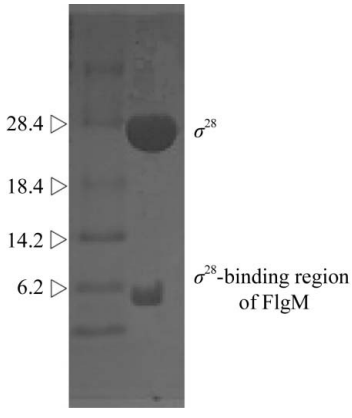


Figure 2
SDS–PAGE of molecular-weight markers (left lane) and the complex crystals (right lane). The molecular weights of the markers are indicated (kDa). The bands for full-length σ^{28} and the σ^{28}-binding region of FlgM are also indicated.

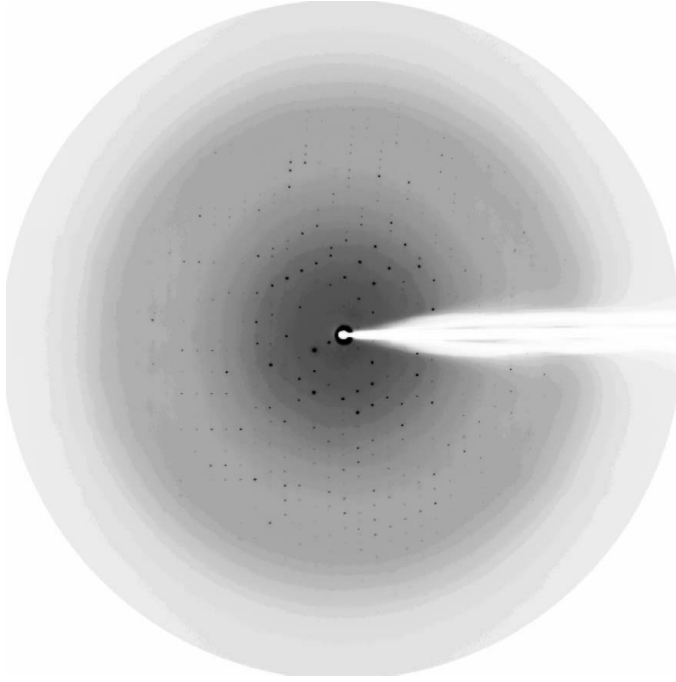


Figure 3
X-ray diffraction image of the native crystal of full-length σ^{28} in complex with the σ^{28}-binding region of FlgM.

structure of *E. coli* flagellar sigma factor σ^{28} in complex with the σ^{28}-binding fragment of its anti-σ^{28} factor FlgM will not only provide valuable results on the role of FlgM in the flagellar regulatory system, but insights into the notion that proteins are dynamic systems.

This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan to AI and TH. This work was also supported by a Grant-in-Aid for Encouragement of Young Scientists from the Ministry of Education, Science and Culture of Japan to KO. We are grateful to Ms J. Tsukamoto for her assistance with the mass spectroscopy. We also would like to thank all the staff for their assistance with the MAD data collection at SPring-8.

References

Campbell, E. A., Masuda, S., Sun, J. L., Muzzin, O., Olson, C. A., Wang, S. & Darst, S. A. (2002). *Cell*, **108**, 795–807.
Campbell, E. A., Tupy, J. L., Gruber, T. M., Wang, S., Sharp, M. M., Gross, C. A. & Darst, S. A. (2003). *Mol. Cell*, **11**, 1067–1078.
Chadsey, M. S., Karlinsey, J. E. & Hughes, K. T. (1998). *Genes Dev.* **12**, 3123–3136.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Daughdrill, G. W., Chadsey, M. S., Karlinsey, J. E., Hughes, K. T. & Dahlquist, F. W. (1997). *Nature Struct. Biol.* **4**, 285–291.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Dubendorff, J. W. & Studier, F. W. (1991). *J. Mol. Biol.* **219**, 45–59.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hughes, K. T. & Mathee, K. (1998). *Annu. Rev. Microbiol.* **52**, 231–286.
Kundu, T. K., Kusano, S. & Ishihama, A. (1997). *J. Bacteriol.* **179**, 4264–4269.
Kutsukake, K. & Iino, T. (1994). *J. Bacteriol.* **176**, 3598–3605.
Kutsukake, K., Ohya, Y. & Iino, T. (1990). *J. Bacteriol.* **172**, 741–747.
Liu, X. & Matsumura, P. (1995). *Gene*, **164**, 81–84.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. (2002). *Science*, **296**, 1285–1290.
Murakami, K. S., Masuda, S. & Darst, S. A. (2002). *Science*, **296**, 1280–1284.
Plaxco, K. W. & Gross, M. (1997). *Nature (London)*, **386**, 657–658.
Serenson, M. K., Ray, S. S. & Darst, S. A. (2004). *Mol. Cell*, **14**, 127–138.
Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.
Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. & Darst, S. A. (1999). *Cell*, **98**, 811–824.