Crystallographic studies of the extracytoplasmic function σ factor σ^J from *Mycobacterium tuberculosis*

Kapil Goutam, Arvind Kumar Gupta and Balasubramanian Gopal*

Received 7 March 2015

Accepted 19 May 2015

Edited by R. Sankaranarayanan, Centre for Cellular and Molecular Biology, Hyderabad, India

Keywords: σ factor; sigma factor; SnoaL_2 domain; potential drug target.

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India. *Correspondence e-mail: bgopal@mbu.iisc.ernet.in

Mycobacterium tuberculosis has multiple σ factors which enable the bacterium to reprogram its transcriptional machinery under diverse environmental conditions. σ^J , an extracytoplasmic function σ factor, is upregulated in late stationary phase cultures and during human macrophage infection. σ^J governs the cellular response to hydrogen peroxide-mediated oxidative stress. σ^J differs from other canonical σ factors owing to the presence of a SnoaL_2 domain at the C-terminus. σ^J crystals belonged to the tetragonal space group *I*422, with unit-cell parameters $a = b = 133.85$, $c = 75.08$ Å. Diffraction data were collected to 2.16 Å resolution on the BM14 beamline at the European Synchrotron Radiation Facility (ESRF).

1. Introduction

Transcription in prokaryotes is primarily regulated at the initiation step. σ factors play a prominent role in transcription initiation as they govern both promoter recognition and promoter melting. Changes in gene expression are thus dependent on the reversible association of the σ factor with the RNA polymerase enzyme (Burgess *et al.*, 1969). *Mycobacterium tuberculosis* has 13 σ factors. These include a principal σ factor, σ^A , a principal-like or group 2 σ factor, σ^B , a group 3 σ factor, σ^F , and ten group 4 or extracytoplasmic function (ECF) σ factors, σ^C , σ^D , σ^E , σ^G , σ^H , σ^I , σ^J , σ^K , σ^L and σ^M (Cole *et al.*, 1998; Rodrigue *et al.*, 2006). ECF σ factors couple environmental changes with the transcriptional profile. The large number of ECF σ factors in *M. tuberculosis* suggests an elaborate mechanism to ensure the survival of this pathogen under diverse microenvironments in the host (Sachdeva *et al.*, 2010).

ECF σ factors are typically two-domain proteins with a conserved domain 2 (σ_2) that binds the Pribnow box element and domain 4 (σ_4) that recognizes the −35 element of the promoter (Lonetto *et al.*, 1994; Gruber & Gross, 2003). The activity of many ECF σ factors is post-translationally regulated by either a soluble anti- σ factor (in the case of σ^E and σ^H) or a membrane-bound anti- σ factor (regulating σ^D , σ^K , σ^L and σ^M) (Song *et al.*, 2003; Hahn *et al.*, 2005; Saïd-Salim *et al.*, 2006; Dainese *et al.*, 2006; Donà *et al.*, 2008; Sklar *et al.*, 2010). The release of the free ECF σ from an inactive σ -anti- σ complex is influenced by multiple mechanisms including phosphorylation, redox stimuli or specific degradation of the anti- σ domain by Clp proteases (Song *et al.*, 2003; Park *et al.*, 2008; Sklar *et al.*, 2010; Jaiswal *et al.*, 2013). Among ECF σ factors, σ^C , σ^G , σ^I and σ^J are not associated with an anti- σ factor. Interestingly, these ECF σ factors have either N- or C-terminal polypeptide extensions that can influence the function of these transcription factors. While the N-terminal

© 2015 International Union of Crystallography

polypeptide is likely to play a role in the auto-regulation of σ^C , the role of the C-terminal domain in σ^G , σ^I and σ^J has not been examined thus far. This C-terminal stretch of around 120 residues is conserved across orthologues (Fig. 1). The C-terminal polypeptide extension contains a putative SnoaL_2 domain in σ^G and σ^J . In eukaryotes, proteins with SnoaL domains play a role in facilitating protein–protein interactions. This inference was based on the similarity of the SnoaL_2 domain to a region of nuclear transport factor 2 (NTF) that interacts with the Ras GTPase (Rodrigue *et al.*, 2006).

The expression of σ^J is upregulated during the late stationary phase of growth and during human macrophage infection (Hu & Coates, 2001; Cappelli *et al.*, 2006). Interestingly, σ^J knockout mutants of *M. tuberculosis* have a comparable growth and survival ability to the wild-type strain in stationary-phase cultures. *M. tuberculosis* σ^J deletion-mutant strains are more susceptible to killing by hydrogen peroxide compared with the wild type, suggesting a role for

this σ factor in hydrogen peroxide resistance (Hu *et al.*, 2004). In a study using an *Escherichia coli* two-plasmid system, σ^J was shown to recognize a promoter upstream of the ECF σ factor σ^I , suggesting a direct role of σ^J in the expression of σ^I (Homerova *et al.*, 2008).

The unusual domain architecture of *M. tuberculosis* σ^J and its role in hydrogen peroxide-mediated stress resistance and high expression levels during the late stationary phase makes this σ factor an interesting target for structure–function analysis. Here, we describe biochemical characterization and crystallization experiments on this protein

2. Materials and methods

2.1. Macromolecule production

The gene encoding σ^J was PCR-amplified from *M. tuberculosis* H37Rv genomic DNA. The amplified product was

Figure 1
Domain architecture of σ^J . (a) σ^J contains two DNA-binding domains (shown here as Sigma70_r2 and Sigma70_r4) preceding the SnoaL_2 domain. (b) Multiple sequence alignment of σ^J homologues. The SnoaL_2 domain is conserved across orthologues. The tentative domain boundaries are marked using black, blue and green outlined boxes. The abbreviations used are Mtb, *M. tuberculosis*; Mbo, *M. bovis*; Mav, *M. avium*; Min, *M. intracellulare*; Mma, *M. marinum*; Msm, *M. smegmatis*.

Table 1

Macromolecule-production information.

Source organism	<i>M. tuberculosis</i> H37Rv
DNA source	<i>M. tuberculosis</i> genomic DNA (obtained from BEI Resources)
Forward primer†	ATATATGCTAGCATGGAGGTTCCGAATTCGAGGC
Reverse primer‡	ATATATCTCGAGATTCCGGTGATGCCTGCC
Cloning vector	Modified pET-15b§
Expression vector	Modified pET-15b
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced	MGSSHHHHHSSGLVPRGSHMASMEVSEFEALRQ-HLMSVAYRLTGTVADAEDIVQEAWLWDSPT-VIADPRAWLTTVVSRLGLDKLRSAHRRREYTG-TWLPEPVVTGLDATDPLAAVVAEDARFAAM-VVLERLRPDQRFVFLHDFAVPFAEVAEVLG-TSEAAARQLASARKAVTAQPALISGDPDPAH-NEVVGRLLMAAMAAGDLDTVVSLLHPDVTFTGD-SNGKAPTAVRAVRGSDKVVRFILGLVQRYGPG-LFGANQLALVNGELGAYTAGLPVDGYRAMAP-RITAITVRDGVKVCALWDIANPDKFTGSPLKER-RAQPTGRGRHHRNLE

† Contains an *NheI* site. ‡ Contains an *XhoI* site. § The vector has been modified to contain an additional restriction site (*NheI*) and a stop codon (after *XhoI*).

cloned between the *NheI* and *XhoI* sites of a modified pET-15b vector (Novagen Inc.; Table 1). The clone was confirmed using single primer-based sequencing (Amnion Biosciences Pvt. Ltd). The expression clone was transformed into *Escherichia coli* BL21(DE3) strain (Novagen Inc.). A single colony was inoculated in Luria–Bertani (LB) medium containing ampicillin ($100 \mu\text{g ml}^{-1}$). Cultures were grown at 310 K prior to induction with 0.2 mM IPTG. The cultures were grown for a further 12 h at 291 K. Cells were then spun at 7000g for 15 min. The pellet was resuspended in buffer A (50 mM Tris–HCl pH 8.0, 300 mM NaCl) containing 2 mM PMSF and EDTA-free protease-inhibitor tablets (Roche). The cells were lysed by sonication and the cell debris was removed by centrifugation at 30 000g for 45 min at 277 K. The supernatant was incubated with Ni–NTA resin (Sigma–Aldrich) for 1 h. The protein was then eluted using buffer A with a gradient

Table 2

Crystallization.

Method	Microbatch under oil
Plate type	72-well microbatch plates
Temperature (K)	291
Protein concentration (mg ml^{-1})	10
Buffer composition of protein solution	50 mM Tris–HCl pH 8.0, 300 mM NaCl
Composition of reservoir solution	0.2 M potassium sodium tartrate, 0.1 M sodium citrate tribasic pH 5.6, 2 M ammonium sulfate
Volume and ratio of drop	4 μl drop, 2 μl protein solution: 2 μl crystallization condition

of imidazole (10–200 mM). The protein was further purified by size-exclusion chromatography using a Sephacryl S200 HiPrep 16/60 column (GE Healthcare). The fractions containing the purified protein were concentrated to $\sim 10 \text{ mg ml}^{-1}$ for crystallization trials. The purity and molecular mass of the protein were verified using SDS–PAGE and mass spectrometry.

2.2. Crystallization

Initial crystallization trials were performed using screens from Hampton Research. The sparse-matrix conditions were examined using the microbatch method at 291 K. A crystallization drop consisted of 2 μl protein solution and 2 μl crystallization condition. Crystals were obtained in a condition consisting of 0.2 M potassium sodium tartrate, 0.1 M sodium citrate tribasic pH 5.6, 2 M ammonium sulfate (Fig. 2). Crystals were soaked in mother liquor containing 20% glycerol followed by flash-cooling in liquid nitrogen (Table 2).

2.3. Data collection and processing

Diffraction data were collected on the BM14 beamline at the European Synchrotron Radiation Facility (ESRF) at 100 K. The diffraction data were collected using an oscillation of 0.5° per image. The diffraction data were processed using

Figure 2

Purification and crystallization of σ^J . (a) 12% silver-stained SDS–PAGE. Lane 1, Bio-Rad dual colour protein markers (molecular masses are labelled in kDa). Lane 2, an analysis of σ^J crystals suggests the crystallized species to be full length. (b) Tetragonal crystals of σ^J . (c) A representative diffraction image for crystals of σ^J .

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BM14, ESRF
Wavelength (Å)	0.9762
Temperature (K)	100
Detector	MAR225
Crystal-to-detector distance (mm)	232.09
Rotation range per image (°)	0.5
Total rotation range (°)	62.5
Exposure time per image (s)	3
Space group	<i>I</i> 422
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 133.85, <i>c</i> = 75.08
Mean <i>I</i> half-set correlation CC _{1/2}	0.99 (0.54)
Mosaicity (°)	0.73
Resolution range (Å)	66.93–2.16 (2.28–2.16)
Total No. of reflections	91937 (12661)
No. of unique reflections	18347 (2640)
Completeness (%)	99.4 (99.9)
Multiplicity	5.0 (4.8)
$\langle I/\sigma(I) \rangle$	13.1 (1.4)
<i>R</i> _{meas}	0.062 (1.152)
Overall <i>B</i> factor from Wilson plot (Å ²)	56.5

iMosflm (Battye *et al.*, 2011) and were scaled (Table 3) using *SCALA* (Evans, 2006).

3. Results and discussion

The *M. tuberculosis sigJ* gene (Rv3328c) encoding the ECF σ factor σ^J was cloned into a modified pET-15b expression vector (Novagen). The expression clone was confirmed using a single primer-based sequencing method (Sanger *et al.*, 1977). The recombinant σ^J protein could be purified to greater than 90% purity using a two-step purification process involving affinity and size-exclusion chromatography. The molecular mass of the recombinant protein was found to be 36 270 Da using liquid-chromatography electrospray ionization mass spectrometry (LC-ESI-MS). The oligomeric status of *M. tuberculosis* σ^J was examined using analytical size-exclusion chromatography. These experiments, which were performed on a Superdex 200 column (GE Healthcare),

Figure 3

M. tuberculosis σ^J is a monomer in solution. Freshly purified σ^J (100 μ l at 10 mg ml⁻¹ concentration) was loaded onto a Superdex S200 column (GE Healthcare) in this size-exclusion chromatography experiment.

suggested that this protein is a monomer in solution (Fig. 3). The protein could be concentrated to 10 mg ml⁻¹. Tetragonal-shaped crystals of σ^J could be obtained in approximately two weeks (Fig. 2). These crystals diffracted to a resolution of 2.16 Å on the BM14 beamline at ESRF, Grenoble. The diffraction data statistics are compiled in Table 3. The Matthews coefficient (2.52 Å³ Da⁻¹) suggests that this crystal form has one molecule of σ^J in the asymmetric unit, with a solvent content of ~51% (Matthews, 1968). Analysis of the diffraction data using *SFCHECK* (Winn *et al.*, 2011) suggested the absence of translational pseudo-symmetry or twinning. *M. tuberculosis* σ^J has two DNA-binding domains, σ_2^J and σ_4^J , that are likely to interact with the Pribnow box (−10 element) and the −35 element of the promoter. These domains are attached to a SnoaL_2 domain (Rodrigue *et al.*, 2006). *M. tuberculosis* σ^J shares very poor sequence similarity with other structurally characterized σ factors (~13%). This observation, as well as the unique domain organization of *M. tuberculosis* σ^J , precludes the possibility of a straightforward molecular-replacement strategy using the structures of other characterized σ factors. The role of the SnoaL_2 domain in *M. tuberculosis* σ^J activity or regulation remains to be determined. The crystal structure is likely to aid in this process as well as in the functional characterization of this transcription factor. Evaluation of heavy-atom derivatives of these crystals and efforts to obtain crystals of selenomethionine-labelled protein are in progress.

Acknowledgements

We acknowledge Dr Hassan Belrhali and Dr Babu A. Manjasetty for their help in data collection at beamline BM14 of the European Synchrotron Radiation Facility, Grenoble, France. This work was supported in part by a grant from the Department of Biotechnology, Government of India.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Burgess, R. R., Travers, A. A., Dunn, J. J. & Bautz, E. K. (1969). *Nature (London)*, **221**, 43–46.
- Cappelli, G., Volpe, E., Grassi, M., Liseo, B., Colizzi, V. & Mariani, F. (2006). *Res. Microbiol.* **157**, 445–455.
- Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
- Dainese, E., Rodrigue, S., Delogu, G., Provvedi, R., Laflamme, L., Brzezinski, R., Fadda, G., Smith, I., Gaudreau, L., Palù, G. & Manganelli, R. (2006). *Infect. Immun.* **74**, 2457–2461.
- Donà, V., Rodrigue, S., Dainese, E., Palù, G., Gaudreau, L., Manganelli, R. & Provvedi, R. (2008). *J. Bacteriol.* **190**, 5963–5971.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Gruber, T. M. & Gross, C. A. (2003). *Annu. Rev. Microbiol.* **57**, 441–466.
- Hahn, M. Y., Raman, S., Anaya, M. & Husson, R. N. (2005). *J. Bacteriol.* **187**, 7062–7071.
- Homerova, D., Halgasova, L. & Kormanec, J. (2008). *FEMS Microbiol. Lett.* **280**, 120–126.
- Hu, Y. & Coates, A. R. (2001). *FEMS Microbiol. Lett.* **202**, 59–65.
- Hu, Y., Kendall, S., Stoker, N. G. & Coates, A. R. (2004). *FEMS Microbiol. Lett.* **237**, 415–423.
- Jaiswal, R. K., Prabha, T. S., Manjeera, G. & Gopal, B. (2013). *Nucleic Acids Res.* **41**, 3414–3423.

- Lonetto, M. A., Brown, K. L., Rudd, K. E. & Buttner, M. J. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 7573–7577.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Park, S. T., Kang, C.-M. & Husson, R. N. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 13105–13110.
- Rodrigue, S., Provvedi, R., Jacques, P.-É, Gaudreau, L. & Manganelli, R. (2006). *FEMS Microbiol. Rev.* **30**, 926–941.
- Sachdeva, P., Misra, R., Tyagi, A. K. & Singh, Y. (2010). *FEBS J.* **277**, 605–626.
- Saïd-Salim, B., Mostowy, S., Kristof, A. S. & Behr, M. A. (2006). *Mol. Microbiol.* **62**, 1251–1263.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
- Sklar, J. G., Makinoshima, H., Schneider, J. S. & Glickman, M. S. (2010). *Mol. Microbiol.* **77**, 605–617.
- Song, T., Dove, S. L., Lee, K. H. & Husson, R. N. (2003). *Mol. Microbiol.* **50**, 949–959.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.