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

Cloning, purification, crystallization and preliminary X-ray analysis of ESX-1-secreted protein regulator (EspR) from *Mycobacterium tuberculosis*

Shanti P. Gangwar, Sita R. Meena and Ajay K. Saxena*

Rm-403/440, Structural Biology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

Correspondence e-mail: ajaysaxena@mail.jnu.ac.in

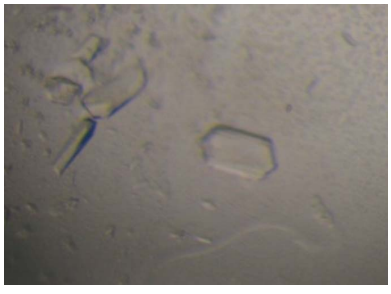
Received 9 June 2010
Accepted 9 November 2010

ESX-1-secreted protein regulator (EspR; Rv3849) is a key regulator in *Mycobacterium tuberculosis* that delivers bacterial proteins into the host cell during infection. EspR binds directly to the *Rv3616c-Rv3614c* promoter and activates transcription and secretes itself from the bacterial cell by the ESX-1 system. The three-dimensional structure of EspR will aid in understanding the mechanisms by which it binds to the *Rv3616c-Rv3614c* promoter and is involved in transcriptional activation. This study will significantly aid in the development of EspR-based therapeutics against *M. tuberculosis*. The full-length EspR gene from *M. tuberculosis* (H37Rv strain) was cloned and overexpressed as a soluble protein in *Escherichia coli*. The protein was purified by affinity chromatography using His-tagged protein followed by size-exclusion chromatography. EspR was crystallized using polyethylene glycol 3350 as precipitant. The crystals diffracted to 3.2 Å resolution using synchrotron radiation of wavelength 0.97625 Å. The crystal belonged to space group $P3_121$ and contained three monomers in the asymmetric unit. Native and heavy-atom-derivatized data sets were collected from EspR crystals for use in *ab initio* structure-solution techniques.

1. Introduction

The ESX-1 protein-secretion system of *Mycobacterium tuberculosis* delivers virulence factors into host macrophages and disarms them during infection (Stanley *et al.*, 2003; Hsu *et al.*, 2003; Pathak *et al.*, 2007). The ESX-1 secretion system is involved in innate immune modulation after infection in macrophages (Stanley *et al.*, 2003, 2007; MacGurn & Cox, 2007; Volkman *et al.*, 2004). Despite the essential role of the ESX-1 system in virulence, the mechanism of ESX-1 secretion is not known. Recently, a locus containing the gene *Rv3616c-Rv3614c* required for ESX-1 secretion-system activity has been identified (Fortune *et al.*, 2005; MacGurn *et al.*, 2005). Rv3616c is a secreted substrate of the secretion-system pathway. In addition, two other substrates, ESTAT-6 and CFP-10, are also involved in the secretion pathway; however, their roles in virulence are not clear (Hsu *et al.*, 2003; Pathak *et al.*, 2007; de Jonge *et al.*, 2007; Singh *et al.*, 2003). The secretion of one substrate is dependent on the secretion of the other substrate (Fortune *et al.*, 2005). EspR (Rv3849) is a new substrate of the ESX-1 system, like ESAT-6, CFP-10 and EspA, and is required for the function of the entire ESX-1 system (Fortune *et al.*, 2005). EspR is induced upon phagocytosis and activates the expression of downstream ESX-1 components. Depletion of EspR affects *M. tuberculosis* gene expression, including loci that are critical for ESX-1 function (Raghavan *et al.*, 2008).

EspR is a DNA-binding transcriptional regulator; it binds 520 bp of the *Rv3616c* promoter and is involved in transcriptional activation (Raghavan *et al.*, 2008). The protein consists of 132 residues and has a molecular weight of ∼14.7 kDa. A structural homology search revealed close homology to the *Bacillus subtilis* transcription factor SinR, which is a helix–turn–helix DNA-binding protein (Lewis *et al.*, 1996). The N-terminal domain of EspR harbours a DNA-binding region and point-mutation analysis in this domain showed reduced DNA-binding affinity (Raghavan *et al.*, 2008). The C-terminal domain of EspR is required for transcriptional activity; deletion of ten amino acids from this domain completely abolished activity (Raghavan *et al.*, 2008).



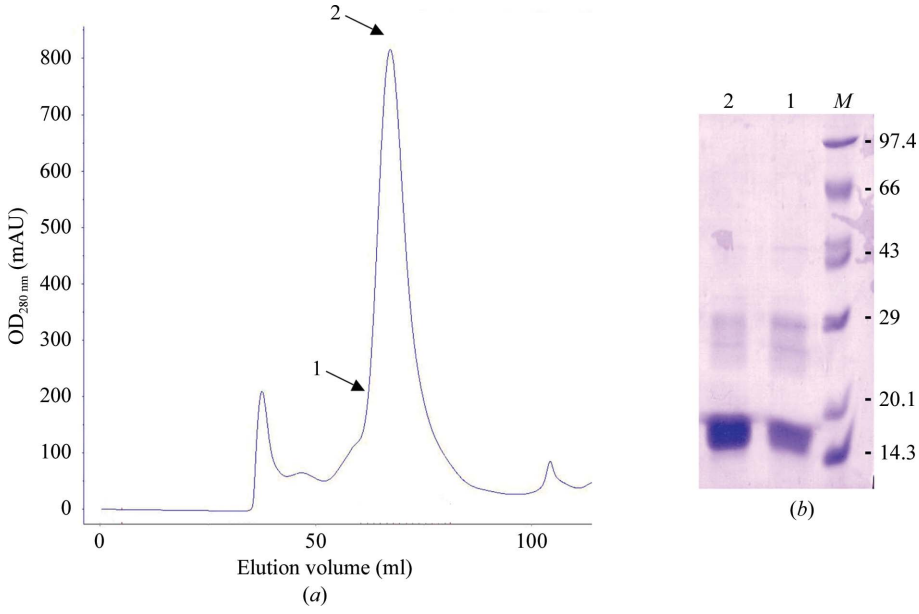


Figure 1
(*a*) FPLC elution profile of the purification of EspR protein by size-exclusion chromatography using a Superdex 75 (16/60) column. The major peak corresponds to EspR protein. (*b*) SDS–PAGE analysis after size-exclusion chromatography of purified EspR protein. Lane *M*, molecular-weight markers (kDa). Lanes 1 and 2, SDS–PAGE analysis of the eluted fractions containing purified EspR protein.

Here, we report the cloning, purification, crystallization and preliminary X-ray crystallographic study of EspR. Structural studies of EspR will contribute significantly to understanding the mechanism of DNA binding and transcriptional activation, as well as its involvement in the ESX-1 secretion pathway.

2. Materials and methods

2.1. Expression and purification

The gene encoding EspR (Met1–Ala132) was amplified from *M. tuberculosis* (H37Rv strain) by polymerase chain reaction and cloned into pET-28a(+) vector (Novagen) containing a 6×His tag and a thrombin cleavage site at the N-terminus. The following primers were used for PCR: forward primer 5′-GATCGCTAGCATGCAACCGATGACCGCT-3′ and reverse primer 5′-CATGCTCGAGCTAATCGTCGATCCCTTC-3′. The resulting construct was transformed into *Escherichia coli* BL21 (DE3) cells.

The cells were grown in 2 l Luria–Bertani (LB) medium containing 50 µg ml^{-1} kanamycin at 310 K until the OD_{600} reached 0.6, followed by induction with 125 µ*M* IPTG at 310 K for 4 h. The cultured cells were harvested by centrifugation, resuspended in 50 ml lysis buffer consisting of 25 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 1 m*M* benzamidine–HCl, 0.1% Triton X-100, 5% glycerol, 2 m*M* β-mercaptoethanol, 1 m*M* phenylmethylsulfonyl fluoride and 0.5 mg ml^{-1} lysozyme and disrupted by sonication at 277 K. The protein was purified using affinity chromatography and size-exclusion chromatography at 277 K (Fig. 1) and characterized by N-terminal sequencing and mass spectrometry. The purified EspR protein contained a total of 150 amino-acid residues: six residues from the 6×His tag, 12 residues from the thrombin cleavage site and 132 residues of EspR protein.

2.2. Crystallization

The EspR protein was concentrated to 8 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.5 for crystallization experiments. The initial crystallization conditions were screened using Structure Screens I and II from Molecular Dimensions and Crystal Screen, Crystal Screen 2 and PEG/Ion Screen from Hampton Research. All crystallization experiments were performed at 277 K using the sitting-drop vapour-diffusion technique. In each trial, 1 µl EspR protein solution was mixed with 1 µl precipitant solution and equilibrated against a reservoir containing 100 µl precipitant solution.

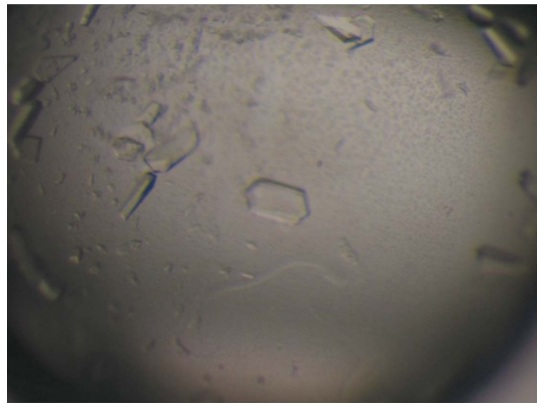


Figure 2
Trigonal crystals of EspR protein. The crystals grew as rectangular bars with typical dimensions of 0.4 × 0.3 × 0.2 mm.

2.3. Data collection

For intensity data collection, a single crystal of EspR was transferred into a solution consisting of 30% PEG 3350, 200 m*M* sodium malonate and 100 m*M* bis-tris propane pH 6. These crystals were directly frozen in liquid nitrogen as 30% PEG 3350 was suitable as a cryoprotectant for diffraction measurements at cryogenic temperature.

A native intensity data set was collected from an EspR crystal at 100 K using a MAR225 image-plate detector on the BM14 beamline

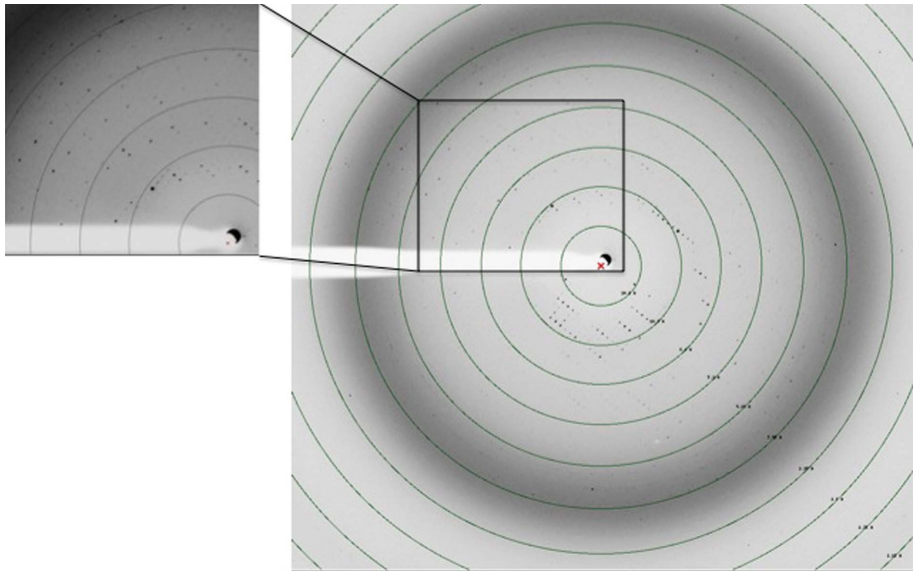


Figure 3
Typical X-ray diffraction pattern of a native EspR crystal (oscillation width 1°). The edge of the frame corresponds to 2.7 Å resolution.

Table 1
X-ray data-collection statistics.

Values in parentheses are for the last resolution shell.

Resolution (Å)	50–3.2 (3.26–3.20)
X-ray source	BM14, ESRF
Wavelength (Å)	0.97625
Space group	$P3_121$
Unit-cell parameters (Å, °)	$a = b = 83.9$, $c = 131.0$, $\alpha = \beta = 90$, $\gamma = 120$
Observed reflections	84807
Unique reflections	8730 (309)
Completeness (%)	97.3 (76.3)
Multiplicity (%)	9.7
$R_{\rm merge}$ (%)†	13.3 (66.9)
Average $I/\sigma(I)$	14.1 (1.8)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith intensity measurement of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of that reflection.

at the ESRF, France. Indexing and integration of the images were performed using the *DENZO* program and scaling and merging were performed using the *SCALEPACK* program (Otwinowski & Minor, 1997). $F_{\rm obs}$ values were produced using the *SCALEPACK2MTZ* program from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

EspR crystals were obtained from a drop comprising 1 µl protein solution and 1 µl reservoir solution consisting of 20% PEG 3350, 200 m*M* sodium malonate and 100 m*M* bis-tris propane pH 6.5. The EspR crystals usually appeared after 7–8 d and grew to maximum dimensions of 0.4 × 0.3 × 0.2 mm (Fig. 2). The crystals often grew with a rectangular shape. In a few drops the crystals grew as clusters and a unique fragment was separated by touching the cryoloop on the surface of the crystals.

The native crystal belonged to space group $P3_121$, with unit-cell parameters $a = b = 83.9$, $c = 131.0$ Å, $\alpha = \beta = 90$, $\gamma = 120°$, and contained three monomers in the asymmetric unit. Diffraction data were collected to 3.2 Å resolution (Fig. 3) and details of the data-collection and processing statistics are given in Table 1. Based on the presence of three molecules of EspR in the asymmetric unit, the Matthews coefficient was $V_{\rm M} = 2.6$ Å^3 Da^{-1}, which corresponds to a solvent content of 53.3%. These values lie within the range normally observed in protein crystals (Matthews, 1968).

To obtain phase information, molecular-replacement analysis was performed using the *Phaser* program (McCoy *et al.*, 2005). Two models, (i) residues 613–693 of enterochelin esterase from *Shigella flexneri* (PDB entry 2b20, chain *A*; Y. Kim, N. Maltseva, I. Dementieva, P. Quartey, D. Holzle, F. Collart & A. Joachimiak, unpublished work) with 32% sequence identity and (ii) *B. subtilis* transcription factor SinR (PDB entry 1b0n; Lewis *et al.*, 1998) with 19.7% sequence identity, were obtained using the *MODWEB* homology-modelling web server (http://salilab.org/modeller). Molecular-replacement analysis with both models did not yield useful phases for structure solution of EspR.

Currently, we are collecting heavy-atom derivative data sets, including a selenomethionine derivative, of EspR and expect to solve the structure by *ab initio* methods. The three-dimensional structure of EspR will aid in understanding its structure–function relationship, which will play a key role in therapeutics against *M. tuberculosis*.

AKS is supported by UGC Networking, JNU Capacity Buildup, Council of Scientific and Industrial Research (CSIR) and Department of Science and Technology (DST) grants for research projects. X-ray data from EspR crystals were collected on the BM14 beamline at the ESRF, France as well as at the X-ray diffraction facility of the Advanced Instrumentation Research Facility (AIRF) of Jawaharlal Nehru University, India. SPG is supported by a Junior Research Fellowship from UGC, India and SRM is supported by a Senior Research Fellowship from DBT, India.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Fortune, S. M., Jaeger, A., Sarracino, D. A., Chase, M. R., Sassetti, C. M., Sherman, D. R., Bloom, B. R. & Rubin, E. J. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 10676–10681.
Hsu, T. *et al.* (2003). *Proc. Natl Acad. Sci. USA*, **100**, 12420–12425.

Jonge, M. I. de, Arnaudet, G. P., Fertz, M. M., Romain, F., Bottai, D., Bordin, P., Honoré, N., Marchal, G., Jiskoot, W., England, P., Cole, S. T. & Brosch, R. (2007). *J. Bacteriol.* **189**, 6028–6034.
Lewis, R. J., Brannigan, J. A., Offen, W. A., Smith, I. & Wilkinson, A. J. (1998). *J. Mol. Biol.* **283**, 907–912.
Lewis, R. J., Brannigan, J. A., Smith, I. & Wilkinson, A. J. (1996). *FEBS Lett.* **378**, 98–100.
MacGurn, J. A. & Cox, J. S. (2007). *Infect. Immun.* **75**, 2668–2678.
MacGurn, J. A., Raghavan, S., Stanley, S. A. & Cox, J. S. (2005). *Mol. Microbiol.* **57**, 1653–1663.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pathak, S. K., Basu, S., Basu, K. K., Banerjee, A., Pathak, S., Bhattacharyya, A., Kaisho, T., Kundu, M. & Basu, J. (2007). *Nature Immunol.* **8**, 610–618.
Raghavan, S., Manzanillo, P., Chan, K., Dovey, C. & Cox, J. S. (2008). *Nature (London)*, **454**, 717–722.
Singh, B., Singh, G., Trajkovic, V. & Sharma, P. (2003). *Clin. Exp. Immunol.* **134**, 70–77.
Stanley, S. A., Johndrow, J. E., Manzanillo, P. & Cox, J. S. (2007). *J. Immunol.* **178**, 3143–3152.
Stanley, S. A., Raghavan, S., Hwang, W. W. & Cox, J. S. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 13001–13006.
Volkman, H. E., Clay, H., Berry, D., Chang, J. C. W., Sherman, D. R. & Ramakrishnan, L. (2004). *PLoS Biol.* **2**, e367.