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

Structural Biology Laboratory, School of Life
Sciences, Jawaharlal Nehru University,
New Delhi 110 067, India

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

Received 21 July 2012
Accepted 9 September 2012

Purification, crystallization and preliminary X-ray crystallographic analysis of the ETS domain of human Ergp55 in complex with the *cfos* promoter DNA sequence

The Ergp55 protein belongs to the Ets family of transcription factors. The Ets transcription factors are involved in various developmental processes and the regulation of cancer metabolism. They contain a highly similar DNA-binding domain known as the ETS domain and have diverse functions in oncogenesis and physiology. The Ets transcription factors differ in their DNA-binding preference at the ETS site and the mechanisms by which they target genes are not clearly understood. To understand its DNA-binding mechanism, the ETS domain of Ergp55 was expressed and purified. The ETS domain was crystallized in the native form and in complex forms with DNA sequences from the *E74* and *cfos* promoters. An X-ray diffraction data set was collected from an ETS-*cfos* DNA complex crystal at a wavelength of 0.9725 Å on the BM14 synchrotron beamline at the ESRF, France. The ETS-*cfos* DNA complex crystal belonged to space group C222₁, with four molecules in the asymmetric unit. For structure analysis, initial phases for the ETS-*cfos* DNA complex were obtained by the molecular-replacement technique with *Phaser* in the *CCP4* suite using the coordinates of Fli-1 protein (PDB entry 1fli) and *cfos* DNA (PDB entry 1bc7) as search models. Structure analysis of the ETS-*cfos* DNA complex may possibly explain the DNA-binding specificity and its mechanism of interaction with the ETS domain of Ergp55.

1. Introduction

The Ets transcription factors activate the transcription of multiple genes involved in cellular growth and development (Sharrocks *et al.*, 1997). The genes that encode Ets transcription factors are the targets of chromosomal translocation, which results in various types of leukaemia, Ewing's sarcoma and other human diseases (Dittmer & Nordheim, 1998). The Ets transcription factors consist of a highly conserved winged-helix-turn-helix DNA-binding domain and bind the consensus DNA core sequence 5'-GGA(A/T)-3' (Sharrocks, 2001).

Ergp55 is a transcription factor from the Ets family. The Ergp55 gene is rearranged in human myeloid leukaemia (Shimizu *et al.*, 1993) and in 5–10% of patients with Ewing's sarcoma (Sorenson *et al.*, 1994). Ergp55 is essential for definitive haematopoiesis, adult haematopoietic stem-cell function and the maintenance of normal peripheral blood platelet numbers (Loughran *et al.*, 2008). TMPRSS2-Ergp55 fusion oncogene transcripts in prostate cancer cells have been significantly associated with aggressive cancer, metastatic spread and increased probability of death (Tomlins *et al.*, 2005).

Crystal structures of the ETS domains of several Ets transcription factors have been determined using X-ray crystallography and nuclear magnetic resonance techniques (Garvie *et al.*, 2001; Kodandapani *et al.*, 1996; Mo *et al.*, 1998; Pufall *et al.*, 2005; Wang *et al.*, 2005; Lamber *et al.*, 2008; Agarkar *et al.*, 2010; Batchelor *et al.*, 1998). A genome-wide analysis of Ets-family DNA binding *in vitro* and *in vivo* has recently been performed (Wei *et al.*, 2010).

To understand the DNA-binding specificity and the mechanism of interaction of the ETS domain of Ergp55, we have purified and crystallized the native ETS domain and its complexes with the DNA sequences of the *E74* promoter (5'-TACCGGAAGT-3') and the *cfos* promoter (5'-GACAGGATGTG-3'). An X-ray data set for the

© 2012 International Union of Crystallography
All rights reserved

ETS-*cfos* DNA complex was collected to 3.1 Å resolution and structure analysis is currently in progress. The ETS-*cfos* DNA complex structure may explain the promoter DNA-binding mechanism and the specificity of the ETS domain of Ergp55. The current structure analysis may contribute to drug development against prostate cancer.

2. Materials and methods

2.1. Expression and purification of the ETS domain

The gene encoding residues 307–399 of Ergp55 (the ETS domain) was obtained by polymerase chain reaction using full-length Ergp55 cDNA purchased from Open Biosystems. The forward primer 5'-GATCGCTAGCGGCAGTGGCCAGATCCAGCTTTGG-3' and the reverse primer 5'-CATGCTCGAGGAGGCCTGGGCGAT-3' were used for ETS gene amplification. The PCR product was digested and inserted into a pET21a(+) vector (Novagen) using *Nhe*I and *Xho*I restriction sites. The resulting plasmid contains bases for three residues from the vector at the N-terminus, the ETS domain and eight residues from the vector at the C-terminus.

The plasmid was transformed into *Escherichia coli* BL21 (DE3) cells for protein expression. The cells were grown at 310 K in Luria-Bertani medium supplemented with ampicillin (100 µg ml⁻¹) until the OD₆₀₀ reached 0.5–0.6. The culture was induced with 1 mM IPTG

at 310 K and grown for a further 4 h. The ETS domain was over-expressed in the soluble fraction of the cells. The cells were harvested by centrifugation at 10 000g for 15 min at 277 K and washed again with 20 mM Tris-HCl pH 7.5. The cells were pelleted by centrifugation. The cell pellet was resuspended in lysis buffer [20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidinium-HCl, 3 mM β-mercaptoethanol, 5% glycerol, 0.1% Triton X-100, 0.2 mg ml⁻¹ lysozyme] and kept on ice for 1 h. The cells were lysed by sonication (eight pulses with 1 min cooling on ice), centrifuged at 12 000g for 20 min and the supernatant was collected.

For immobilized metal-affinity chromatography, the supernatant was mixed with pre-equilibrated His-Bind Fractogel (Novagen) and incubated for 2 h on a 360° rocker at 277 K. The protein mixture was loaded into an empty column and washed with buffer consisting of 25 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM PMSF, 1 mM benzamidinium-HCl, 3 mM β-mercaptoethanol, 5% glycerol, 40 mM imidazole. After washing, the protein was eluted in buffer consisting of 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM PMSF, 0.5 mM benzamidinium-HCl, 3 mM β-mercaptoethanol, 5% glycerol, 300 mM imidazole. The eluted fractions were pooled and concentrated using an Amicon ultracentrifugal device (3 kDa cutoff). The concentrated protein was loaded onto a Superdex 75 column (HiLoad 16/60) pre-equilibrated with buffer consisting of 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 3 mM β-mercaptoethanol, 5% glycerol. The peak

Figure 1

(a) Elution profile of the purified ETS domain from a Superdex 75 (16/60) column. The major peak corresponds to monomeric ETS domain. (b) SDS-PAGE analysis of the purified ETS domain after size-exclusion chromatography. Lane M, molecular-mass marker (labelled in kDa). Lane 1 corresponds to the top fraction of the eluted peak.

Figure 2

Typical crystals of (a) native ETS domain, (b) the ETS-E74 DNA complex and (c) the ETS-*cfos* DNA complex. These crystals were used in X-ray diffraction experiments.

fractions were pooled and concentrated to 10–15 mg ml⁻¹ using a 3 kDa cutoff centrifugal ultrafiltration device (Millipore, USA). The protein concentration was measured using the absorbance at 280 nm. The purified protein had a purity of greater than 98% as determined by SDS-PAGE and mass-spectrometric analyses (Fig. 1). The final recombinant ETS domain contained 104 residues: three residues from the vector at the N-terminus, 93 residues of the ETS domain of Ergp55 and eight residues from the vector at the C-terminus, including a 6×His tag. Part of the ETS-domain purification has also been published in our recent paper (Gangwar *et al.*, 2012).

2.2. Crystallization

The ETS domain was concentrated to 8 mg ml⁻¹ in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 5% glycerol, 3 mM β -mercaptoethanol before crystallization. Initial crystallization conditions were screened using Crystal Screen, Crystal Screen 2, PEG/Ion and Index from Hampton Research. The hanging-drop vapour-diffusion technique was used for crystallization experiments. 1 μ l protein solution was mixed with 1 μ l precipitant solution and equilibrated against 500 μ l reservoir solution. All crystallization experiments were performed at 277 and 289 K.

2.2.1. Crystallization of the native ETS domain. Crystals of the native ETS domain were obtained by the vapour-diffusion technique at 277 K. 1 μ l protein solution at 7–8 mg ml⁻¹ was mixed with 1 μ l reservoir solution consisting of 100 mM sodium acetate buffer pH 4.5, 22%(w/v) PEG 4000 and equilibrated against 500 μ l reservoir solution. Rectangular-shaped crystals appeared in 10–15 d and grew to maximum dimensions of 0.3 $\times$ 0.2 $\times$ 0.15 mm (Fig. 2a).

2.2.2. Crystallization of the ETS-E74 DNA (5'-TACCGGAAGT-3') complex. HPLC-grade purified E74 DNA (5'-TACCGGAAGT-3') was obtained from Sigma-Aldrich. Prior to crystallization, 0.2 mM ETS domain was mixed with 0.25 mM duplex E74 DNA and incubated for 2 h at 277 K. 1 μ l ETS-DNA complex was mixed with 1 μ l reservoir solution consisting of 20%(w/v) PEG 4000, 200 mM MgCl₂ and equilibrated against 500 μ l reservoir solution. The hanging-drop

Table 1

Data-collection, processing and molecular-replacement statistics for the ETS-*cfos* DNA complex.

Values in parentheses are for the highest resolution shell.

Resolution (Å)	50–3.1 (3.15–3.10)
Wavelength (Å)	0.9725
X-ray source	BM14, ESRF
Space group	C222 ₁
Unit-cell parameters (Å)	<i>a</i> = 45.9, <i>b</i> = 217.8, <i>c</i> = 113.4
Total reflections	128009
Unique reflections	10906 (469)
Completeness (%)	91.9 (81.7)
<i>R</i> _{merge} [†] (%)	20.6 (63.6)
Average <i>I</i> / σ (<i>I</i>)	4.3 (2.3)
Multiplicity	11.7
Molecular-replacement statistics [‡]	
No. of molecules in the asymmetric unit	4
Translation-function Z-score (TFZ)	29.2
No. of C α clashes (PAK)	0
Log-likelihood gain (LLG)	1469

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th intensity measurement of reflection *hkl* and $\langle I(hkl) \rangle$ is the average intensity of reflection *hkl*. [‡] Scores for the best Phaser solution (McCoy *et al.*, 2007).

vapour-diffusion technique was used for crystallization. The complex crystals appeared at 277 K in 6–10 d and grew to maximum dimensions of 0.7 $\times$ 0.3 $\times$ 0.15 mm (Fig. 2b).

2.2.3. Crystallization of the ETS-*cfos* DNA (5'-GACAGGATGTG-3') complex. HPLC-grade purified *cfos* DNA (5'-GACAGGATGTG-3') was obtained from Sigma-Aldrich. Complementary strands were annealed in 20 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA. Tiny needle-shaped crystals of the ETS-*cfos* DNA complex appeared after 7–8 d using precipitant solution consisting of 100 mM bis-Tris propane pH 6.5, 50 mM CaCl₂, 25%(w/v) PEG MME 550. Diffraction-quality crystals of the ETS-*cfos* DNA complex were obtained by microseeding. In the microseeding experiments, 0.2 μ l crushed seeds were added to a 2 μ l drop consisting of 1 μ l protein solution (0.125 mM ETS domain and 0.150 mM duplex *cfos* DNA) and 1 μ l reservoir solution. The 2 μ l drop was equilibrated against 500 μ l reservoir solution consisting of 100 mM sodium cacodylate pH 5.6, 20%(w/v) PEG 4000 in a hanging-drop vapour-diffusion experiment. The ETS-*cfos* DNA complex crystals appeared at 289 K in 7–12 d and grew to maximum dimensions of 0.4 $\times$ 0.3 $\times$ 0.1 mm (Fig. 2c).

2.3. X-ray diffraction data collection and processing

Crystals of the native and complexed ETS domain were checked for diffraction at the home source using Cu K α radiation and on the BM14 synchrotron beamline at the ESRF. The crystals of the native ETS domain and of the ETS-E74 DNA complex only diffracted to 6–8 Å resolution. However, the ETS-*cfos* DNA complex crystals diffracted to 3.1 Å resolution on the BM14 beamline at the ESRF. For intensity data collection, single crystals of the ETS-*cfos* DNA complex were transferred into a solution consisting of 100 mM sodium cacodylate pH 5.6, 35%(w/v) PEG 4000. These crystals were directly cooled in a liquid-nitrogen stream, as 35%(w/v) PEG 4000 acts as a good cryoprotectant for diffraction experiments at cryogenic temperatures.

Intensity data were collected from an ETS-*cfos* DNA complex crystal using a MAR225 CCD detector mounted on the BM14 beamline at the ESRF, France. Indexing and integration of images were performed using *DENZO* and scaling and merging of data sets were performed using *SCALEPACK* (Otwinowski & Minor, 1997). *F*_{obs} values were obtained using *TRUNCATE* from the *CCP4* suite (Winn *et al.*, 2011).

Figure 3

Typical X-ray diffraction pattern of an ETS-*cfos* DNA complex crystal (oscillation width of 1°). The edge of the diffraction frame corresponds to 2.5 Å resolution.

3. Results and discussion

The ETS domain of Ergp55 was expressed in *E. coli* and purified to homogeneity as determined by SDS-PAGE and mass-spectrometric analyses. The ETS domain eluted from a size-exclusion column as a monomer with a molecular mass of 12.2 kDa. Crystals of native ETS and the ETS–E74 DNA complex were obtained and diffracted poorly to 6–8 Å resolution. The rectangular-shaped ETS–*cfos* DNA complex crystals diffracted better than the native crystals and the ETS–E74 DNA complex crystals. A complete X-ray intensity data set was collected on the BM14 beamline at the ESRF.

The ETS–*cfos* DNA complex crystals belonged to space group C222₁, with four monomers in the asymmetric unit. The crystals diffracted to 3.1 Å resolution (Fig. 3) and the data-collection and processing statistics are given in Table 1. A complete and high-quality data set was collected from an ETS–*cfos* DNA complex crystal. With four ETS–*cfos* DNA complexes in the asymmetric unit, the Matthews coefficient V_M was 2.7 Å³ Da^{−1} and the solvent content was 54.9%. These values lie within the range normally observed for protein crystals (Matthews, 1968).

To obtain phase information, molecular-replacement analysis was performed using the *Phaser* program (McCoy *et al.*, 2007). Molecular-replacement analysis yielded useful phases for structure solution of the ETS–*cfos* DNA complex. A preliminary model of the ETS–*cfos* DNA complex was obtained using the NMR coordinates of human Fli-1 (PDB entry 1fli; Liang *et al.*, 1994) and *cfos* DNA coordinates from the SAP1–*cfos* DNA complex structure (PDB entry 1bc7; Mo *et al.*, 1998) as the initial models. The obtained model was refined using *REFMAC5* (Murshudov *et al.*, 2011) from the *CCP4* suite and the refinement converged to an R_{work} of 0.44 and an R_{free} of 0.45. Structure refinement and model building using *Coot* (Emsley & Cowtan, 2004) are currently in progress. Details of the ETS–*cfos* DNA complex structure will be published elsewhere.

AKS is supported by a grant from the Department of Science and Technology (DST), India for the current work. Grants from UGC Networking, JNU Capacity Buildup and the DST Purse are gratefully acknowledged. We also acknowledge the X-ray diffraction facility at AIRF Jawaharlal Nehru University. SRM is supported by a Senior

Research Fellowship from DBT, India. SPG is supported by a Senior Research Fellowship from UGC, India.

References

- Agarkar, V. B., Babayeva, N. D., Wilder, P. J., Rizzino, A. & Tahirov, T. H. (2010). *J. Mol. Biol.* **397**, 278–289.
- Batchelor, A. H., Piper, D. E., de la Brousse, F. C., McKnight, S. L. & Wolberger, C. (1998). *Science*, **279**, 1037–1041.
- Dittmer, J. & Nordheim, A. (1998). *Biochem. Biophys. Acta*, **1377**, F1–F11.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Gangwar, S. P., Dey, S. & Saxena, A. K. (2012). *PLoS One*, **7**, e39850.
- Garvie, C. W., Hagman, J. & Wolberger, C. (2001). *Mol. Cell*, **8**, 1267–1276.
- Kodandapani, R., Pio, F., Ni, C.-Z., Piccialli, G., Klemsz, M., McKercher, S., Maki, R. A. & Ely, K. R. (1996). *Nature (London)*, **380**, 456–460.
- Lamber, E. P., Vanhille, L., Textor, L. C., Kachalova, G. S., Sieweke, M. H. & Wilmanns, M. (2008). *EMBO J.* **27**, 2006–2017.
- Liang, H., Mao, X., Olejniczak, E. T., Nettesheim, D. G., Yu, L., Meadows, R. P., Thompson, C. B. & Fesik, S. W. (1994). *Nature Struct. Biol.* **1**, 871–875.
- Loughran, S. J., Kruse, E. A., Hacking, D. F., de Graaf, C. A., Hyland, C. D., Willson, T. A., Henley, K. J., Ellis, S., Voss, A. K., Metcalf, D., Hilton, D. J., Alexander, W. S. & Kile, B. T. (2008). *Nature Immunol.* **9**, 810–819.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Mo, Y., Vaessen, B., Johnston, K. & Marmorstein, R. (1998). *Mol. Cell*, **2**, 201–212.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pufall, M. A., Lee, G. M., Nelson, M. L., Kang, H.-S., Velyvis, A., Kay, L. E., McIntosh, L. P. & Graves, B. J. (2005). *Science*, **309**, 142–145.
- Sharrocks, A. D. (2001). *Nature Rev. Mol. Cell Biol.* **2**, 827–837.
- Sharrocks, A. D., Brown, A. L., Ling, Y. & Yates, P. R. (1997). *Int. J. Biochem. Cell Biol.* **29**, 1371–1387.
- Shimizu, K., Ichikawa, H., Tojo, A., Kaneko, Y., Maseki, N., Hayashi, Y., Ohira, M., Asano, S. & Ohki, M. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 10280–10284.
- Sorensen, P. H. B., Lessnick, S. L., Lopez-Terrada, D., Liu, X. F., Triche, T. J. & Denny, C. T. (1994). *Nature Genet.* **6**, 146–151.
- Tomlin, S. A. *et al.* (2005). *Science*, **310**, 644–648.
- Wang, Y., Feng, L., Said, M., Balderman, S., Fayazi, Z., Liu, Y., Ghosh, D. & Gulick, A. M. (2005). *Biochemistry*, **44**, 7095–7106.
- Wei, G. H. *et al.* (2010). *EMBO J.* **29**, 2147–2160.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.