electronic reprint

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091
Editors: **H. M. Einspahr and J. M. Guss**

Crystallization and preliminary X-ray crystallographic studies of *Mycobacterium tuberculosis* CRP/FNR family transcription regulator

Mohd. Akif, Yusuf Akhter, Seyed E. Hasnain and Shekhar C. Mande

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site provided that this cover page is retained. Republication of this article or its storage in electronic databases or the like is not permitted without prior permission in writing from the IUCr.

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

Crystallization and preliminary X-ray crystallographic studies of *Mycobacterium tuberculosis* CRP/FNR family transcription regulator

Mohd. Akif,[a] Yusuf Akhter,[a] Seyed E. Hasnain[a,b,c] and Shekhar C. Mande[a]*

[a]Centre for DNA Fingerprinting and Diagnostics, ECIL Road, Nacharam, Hyderabad 500076, India, [b]University of Hyderabad, Hyderabad 500046, India, and [c]Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560064, India

Correspondence e-mail: shekhar@cdfd.org.in

Received 26 June 2006
Accepted 19 July 2006

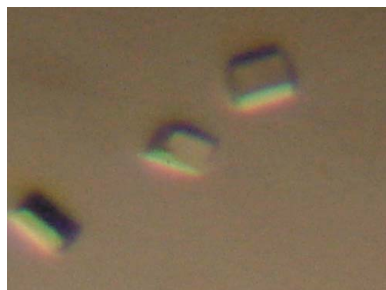
CRP/FNR family members are transcription factors that regulate the transcription of many genes in *Escherichia coli* and other organisms. *Mycobacterium tuberculosis* H37Rv contains a probable CRP/FNR homologue encoded by the open reading frame Rv3676. The deletion of this gene is known to cause growth defects in cell culture, in bone marrow-derived macrophages and in a mouse model of tuberculosis. The mycobacterial gene Rv3676 shares ~32% sequence identity with prototype *E. coli* CRP. The structure of the protein might provide insight into transcriptional regulation in the pathogen by this protein. The *M. tuberculosis* CRP/FNR transcription regulator was crystallized in space group $P2_12_12_1$, with unit-cell parameters a = 54.1, b = 84.6, c = 101.2 Å. The crystal diffracted to a resolution of 2.9 Å. Matthews coefficient and self-rotation function calculations reveal the presence of two monomers in the asymmetric unit.

1. Introduction

Mycobacterium tuberculosis, the causative agent of tuberculosis, is one of the most dreaded human pathogens and causes ~3 million deaths every year (World Health Organization, 2006). The characteristic feature of *M. tuberculosis* is reactivation from a latent phase, which causes the disease (Ulrichs & Kaufmann, 2006). *M. tuberculosis* possesses intricate mechanisms for survival inside the hostile environment of the host. It can persist for a long period of time in a dormant or non-replicating state, which may become active in replicating bacilli after several years when the host becomes immunocompromized (Ulrichs & Kaufmann, 2006). The outcome of *M. tuberculosis* infection solely depends upon its interactions with the environment provided by the host (Kaufmann, 2001). The reactivation may be caused by a variety of environmental signals and is mediated through several transcriptional regulators. A family of transcriptional regulators belonging to the CRP/FNR class is actively associated with low oxygen stress and starvation, including perception of various other environments (Korner *et al.*, 2003). This regulator may therefore play an important role in reactivating the dormant bacilli.

The cyclic AMP (cAMP) receptor protein (CRP) is a well known transcription regulator that regulates the transcription of many genes in *Escherichia coli* (Schultz *et al.*, 1991). It is one of the best studied transcription regulators and is also referred to as CAP (catabolite activator protein). CRP is a 45 kDa dimeric protein that has both cAMP- and DNA-binding domains (Aiba *et al.*, 1982). cAMP is required for the regulatory and DNA-binding activity of CRP (Passner *et al.*, 2000). The DNA-binding domain has a well conserved helix–turn–helix motif. The archetypical cAMP-binding domain has evolved to accommodate different functional specificities in signal detection, DNA binding and interaction with RNA polymerase to allow different family members to respond to a wide range of signals (Green *et al.*, 2001).

M. tuberculosis H37Rv contains a probable CRP/FNR homologue encoded by the open reading frame Rv3676 (Cole *et al.*, 1998). It is reported as a specific transcription factor, deletion of which is known to cause growth defects in laboratory medium, in bone marrow-



© 2006 International Union of Crystallography
All rights reserved

electronic reprint

derived macrophages and in a mouse model of tuberculosis (Rickman *et al.*, 2005). Although Rv3676 shares 32% sequence identity with *E. coli* CRP, it exhibits wide divergence at the N-terminal region. The lack of conserved residues in this region might suggest different interactions of Rv3676 with RNA polymerase. Moreover, only four of the six residues that are involved in cAMP binding in *E. coli* CRP are conserved in Rv3676. It has been reported that the CRP/FNR homologue is closer to the COOA branch represented by the CO sensor protein from *Rhodosprillium rubrum* (Korner *et al.*, 2003). Considering the fact that Rv3676 shares heterogeneity in both the DNA-binding and cAMP-binding sequences compared with other prototypes, its structural properties should be interesting to address. To understand the molecular mechanism of transcription regulation of the CRP/FNR family regulator in *M. tuberculosis*, Rv3676 has been crystallized for structure determination. This paper describes the preliminary crystallographic characterizations of the *M. tuberculosis* CRP/FNR transcription regulator.

2. Material and methods

2.1. Expression and purification

E. coli BL21 (DE3) cells harbouring the expression vector pET28a with the gene Rv3676, cloned in *Nde*I and *Hin*dIII sites with a six-His tag at the C-terminus, were grown in 200 ml LB supplemented with 30 µg ml^{-1} kanamycin. The culture was induced at an OD_{600} of 0.4 with 0.4 m*M* IPTG at 300 K and 200 rev min^{-1} to allow protein expression. The cells were harvested by centrifugation and resuspended in lysis buffer (PBS; 50 m*M* phosphate buffer pH 8 and 155 m*M* NaCl) supplemented with 0.1 m*M* PMSF. After sonication, the supernatant was applied onto a Talon cobalt-affinity resin column (Clontech, USA) pre-equilibrated with lysis buffer, followed by washing with five bed volumes with lysis buffer supplemented with 10 m*M* imidazole. The recombinant protein was eluted with lysis buffer supplemented with 250 m*M* imidazole.

2.2. Crystallization

The purified protein was concentrated to 9 mg ml^{-1} and dialyzed against 10 m*M* Tris pH 8 containing 20 m*M* NaCl using a Centricon concentrator (Amicon; 3 kDa molecular-weight cutoff). The purified protein precipitated when stored at 253 K. Thus, freshly purified protein was used for crystallization trials with a variety of random conditions using a Magic 96 matrix. The hanging-drop vapour-diffusion technique was used for random screening of crystallization conditions. Crystals were obtained when 2 µl protein solution was mixed with 2 µl well solution and allowed to equilibrate against 500 µl well solution at 277 K. 0.2 *M* Li_2SO_4 and 15% ethanol in 0.1 *M* sodium citrate buffer pH 5.5 were used in the well solution.

2.3. Data collection

Crystals of CRP/FNR were soaked in artificial mother liquor (0.1 *M* sodium citrate buffer pH 5.5, 0.2 *M* Li_2SO_4 and 21% ethanol) supplemented with 20% ethylene glycol as the cryoprotectant. The diffraction data were collected from a single crystal at 100 K at the XRD1 beamline at the ELETTRA synchrotron facility, Trieste, Italy using a MAR CCD 165 detector. The crystal-to-detector distance was maintained at 200 mm with oscillations of 1°, covering up to 180° in order to obtain complete data. Determination of unit-cell parameters and integration of reflection data was performed by *DENZO*/*SCALEPACK* (Otwinowski & Minor, 1997). The intensities were then converted to structure-factor amplitudes by the *TRUNCATE* program from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The self-rotation function was calculated using *POLARRFN* as available in the *CCP*4 suite to identify the noncrystallographic twofold symmetry.

Structure solution was attempted by molecular replacement using *AMoRe* (Navaza, 1994) as well as *Phaser* (McCoy *et al.*, 2005) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The homologous *E. coli* CRP complex with cAMP (PDB code 1i5z; Weber & Steitz, 1987) having 32% sequence identity was taken as a search model for molecular replacement. Molecular replacement was also attempted with the reduced CO-sensing protein from *R. rubrum* as the search model (PDB code 1ft9; Lanzilotta *et al.*, 2000).

3. Results and discussion

M. tuberculosis CRP/FNR (Rv3676) is a protein of 224 amino acids, which was purified to homogeneity from an *E. coli* heterologous expression system. The use of His_6 from the pET28a vector led to the introduction of ten additional amino acids in addition to the six histidines into the protein. The protein purity was observed to be better then 99% on SDS–PAGE. The yield of protein was 45 mg from 1 l culture. Upon gel filtration, the protein eluted at a Stokes radius consistent with a dimer, which is in keeping with the quaternary structures observed for other transcription factors of this family.

The protein was stable at room temperature, but was found to precipitate at 253 K. The freshly purified protein was used for crystallization trials and crystals were obtained after one week when the crystallization plate was incubated at a constant temperature of 277 K. The crystals were very unstable at room temperature and often dissolved even while being inspected under an optical microscope. Surprisingly, when the plates were re-incubated at 277 K, good diffraction-quality crystals reappeared in the plates (Fig. 1). These crystals were therefore quickly mounted in cryoloops, frozen and used for data collection.

The completeness of the data was found to be 99.5% (Table 1). The 2.9 Å data were processed using the *HKL*-2000 program suite and the crystal was found to belong to space group $P2_12_12_1$. The data-collection statistics are shown in Table 1. A twinning test (http://nihserver.mbi.ucla.edu/Twinning/) showed that the data were not twinned. Assuming the presence of two monomers in the asymmetric unit, a value of the Matthews coefficient of 2.58 $Å^3$ Da^{-1} was

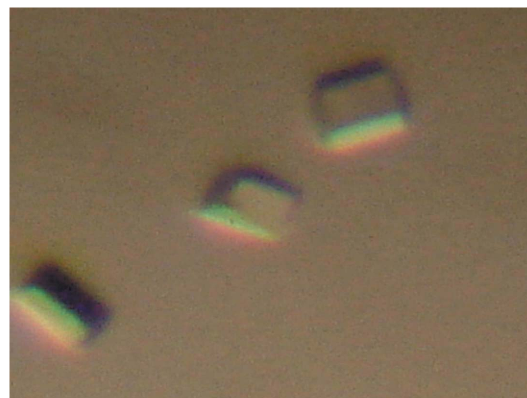


Figure 1
Crystals of CRP/FNR transcription factor grown using 15% ethanol, 0.2 *M* Li_2SO_4 and 0.1 *M* sodium phosphate citrate pH 5.5. The size of each crystal is approximately 0.1 × 0.15 × 0.1 mm.

electronic reprint

Table 1
Diffraction data statistics.

Values in parentheses are for the last resolution shell (3.00–2.90 Å).

Wavelength (Å)	1.0
Resolution (Å)	50–2.9
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 54.1, b = 84.6, c = 101.2
Completeness (%)	99.5 (99.4)
$R_{\rm merge}$† (%)	8 (27.3)
$I/\sigma(I)$	16.2 (5.1)
Unique reflections	10769 (1056)
Total reflections	168417
Redundancy	15.6
Matthews coefficient (Å^3 Da^{-1})	2.58
No. of molecules in ASU	2

† $R_{\rm merge} = \sum |I_{hkl} - \langle I_{hkl} \rangle| / \sum |I_{hkl}|$, where I_{hkl} are the intensities of symmetry-redundant reflections and $\langle I_{hkl} \rangle$ is the average over all reflections.

obtained, which corresponds to a solvent content of 52.4% (Matthews, 1968). Because of the presence of two molecules in the asymmetric unit, the self-rotation function was used to determine noncrystallographic symmetry. The two peaks at φ = 90° and ω = 30 and 60° from the c^* axis in a self-rotation function plot show the expected twofold noncrystallographic symmetry. However, a satisfactory solution using molecular replacement could not be obtained with either *E. coli* CRP or the CO-sensing protein from *R. rubrum*.

The structural details of interactions between transcription factors and a specific DNA sequence is well established for the cAMP receptor (CRP) family of transcription factors. As in the catabolic gene activator CAP (Busby & Ebright, 1999) and the CO-sensing protein, the binding of cAMP switches the protein from an off-state conformation (refractile to DNA binding) to an on-state conformation (allowing DNA binding). In CAP, the binding of cAMP alters the DNA-binding domain. The alteration of the conformation of the DNA-binding domain, which is more than 20 Å away, is not well understood, primarily because the available structures of CAP are either in the presence of cAMP alone or as a cAMP and DNA complex (Schultz *et al.*, 1991; Weber & Steitz, 1987). Therefore, the nature of conformational changes that take place upon cAMP binding will be better understood if the structure of CAP is known in the absence of the activator cAMP.

In conclusion, we have crystallized the CRP/FNR family transcription factor from *M. tuberculosis*. However, the low sequence homology with other known CRPs and the non-availability of the uncomplexed CRP structure meant that we could not obtain molecular-replacement solutions. This suggests the need to obtain a structure solution using experimental phasing techniques such as multi-wavelength anomalous dispersion (MAD) or multiple isomorphous replacement (MIR).

We thank the staff of the XRD1 beamline at the ELETTRA synchrotron, Trieste, Italy for assistance during data collection and the Indo–Italian POC-DST Action for financial assistance in performing experiments at ELETTRA. This work is supported by grants from the Department of Biotechnology and the Council of Scientific and Industrial Research (CSIR), New Delhi, India. MA and YA gratefully acknowledge financial support from the CSIR for a Senior Research Fellowship and Junior Research Fellowship, respectively; SCM is an International Senior Research Fellow of the Welcome Trust, UK.

References

Aiba, H., Fujimoto, S. & Ozaki, N. (1982). *Nucleic Acids Res.* **10**, 1345–1361.
Busby, S. & Ebright, R. H. (1999). *J. Mol. Biol.* **293**, 199–213.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Green, J., Scott, C. & Guest, J. (2001). *Adv. Microb. Physiol.* **44**, 1–34.
Kaufmann, S. H. (2001). *Nature Rev. Immunol.* **1**, 20–30.
Korner, H., Soab, H. J. & Zumft, W. G. (2003). *FEMS Microbiol. Rev.* **27**, 559–592.
Lanzilotta, W. N., Schuller, D. J., Thorsteinsson, M. V., Kerby, R. L., Roberts, G. P. & Poulos, T. L, (2000). *Nature Struct. Biol.* **7**, 876–880.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Passner, J. M., Schultz, S. C. & Steitz, T. A. (2000). *J. Mol. Biol.* **304**, 847–859.
Rickman, L., Scott, C., Hunt, D. M., Hutchinson, T., Menendez, M. C., Whalan, R., Hinds, J., Colston, M. J., Green, J. & Buxton, R. S. (2005). *Mol. Microbiol.* **56**, 1274–1286.
Schultz, S. C., Shields, G. C. & Steitz, T. A. (1991). *Science*, **253**, 1001–1007.
Ulrichs, T. & Kaufmann, S. H. (2006). *J. Pathol.* **208**, 261–269.
Weber, I. T. & Steitz, T. A. (1987). *J. Mol. Biol.* **198**, 311–326.
World Health Organization (2006). *Tuberculosis Fact Sheet No. 104: Global and Regional Incidence.* Geneva: World Health Organization.

electronic reprint