Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary X-ray diffraction analysis of the catalytic domain of adenylyl cyclase Rv1625c from *Mycobacterium tuberculosis*

Amit D. Ketkar,[a] Avinash R. Shenoy,[a] Muppuru M. Kesavulu,[b] Sandhya S. Visweswariah[a] and Kaza Suguna[b]*

[a]Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore 560012, India, and [b]Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

Correspondence e-mail: suguna@mbu.iisc.ernet.in

Received 3 October 2003
Accepted 4 December 2003

The Rv1625c gene product is an adenylyl cyclase identified in the genome of *Mycobacterium tuberculosis* strain H37Rv. It shows sequence similarity to the mammalian nucleotide cyclases and functions as a homodimer, with two substrate-binding sites at the dimer interface. A mutant form of the catalytic domain of this enzyme, K296E/F363R/D365C (KFD→ERC), was overexpressed in *Escherichia coli* cells in a soluble form. Crystals were obtained using the hanging-drop vapour-diffusion method with PEG 8000 as a precipitant. The protein crystallized in space group $P4_1$, with unit-cell parameters $a = b = 71.25$, $c = 44.51$ Å. X-ray diffraction data were collected to a resolution of 3.4 Å and the structure has been solved by the molecular-replacement method using a previously built theoretical model of the protein as the search molecule.

1. Introduction

The nucleotide cyclases are enzymes that convert NTP to cNMP and are classified into different classes based on their amino-acid sequence (Barzu & Danchin, 1994). The class III cyclases are represented in mammals by both soluble and membrane-associated forms of adenylyl and guanylyl cyclases; additional members of this class have also been identified in the genomes of many bacteria (Katayama & Ohmori, 1997; McCue *et al.*, 2000). Some of the class III cyclases from *Mycobacterium tuberculosis* have been characterized biochemically and genetically (Guo *et al.*, 2001; Reddy *et al.*, 2001). Interestingly, the actinobacteria are endowed with multiple putative class III adenylyl cyclase genes that have been identified through large-scale genome analysis (Shenoy *et al.*, 2004). *M. tuberculosis* appears to harbour 16 genes that could encode adenylyl cyclases (Shenoy *et al.*, 2004) and we have begun a systematic analysis of many of these putative gene products both biochemically and structurally. The role of any of these cyclases in the biology or pathogenesis of *M. tuberculosis* has not been described to date.

The crystal structure of a mammalian class III adenylyl cyclase has been described in complex with forskolin and the stimulatory subunit of heterotrimeric G protein, $G_{s\alpha}$ (Tesmer *et al.*, 1997). The structure provided a basis for the requirement of two metal ions for catalysis of the reaction (Tesmer *et al.*, 1999; Zimmermann *et al.*, 1998) and the residues required for catalysis (Liu *et al.*, 1997; Tesmer *et al.*, 1997). It has been shown for the mammalian adenylyl cyclases that the active site is formed at the interface of two domains, C1 and C2, that form a head-to-tail pseudo-heterodimer within the same polypeptide chain (Tang & Hurley, 1998). The C1 and C2 domains are arranged in tandem in the mammalian adenylyl cyclases and are similar in amino-acid sequence but differ in the critical residues that are involved in ATP and metal binding. As a consequence, there is a single active site formed at the C1–C2 interface, with the second site being occupied by forskolin. The crystal structure of the *Trypanosoma* adenylyl cyclase in a monomeric state is also known (Bieger & Essen, 2001). However, no crystal structure of any class III bacterial nucleotide cyclase has been solved to date.

The Rv1625c gene from *M. tuberculosis* shows high sequence similarity to the mammalian adenylyl cyclases (~56% sequence similarity) and cloning and expression of the gene product has conclusively shown that it is an active adenylyl cyclase (Guo *et al.*, 2001; Reddy *et al.*, 2001). Rv1625c is a membrane-associated protein, with six putative transmembrane domains and a single catalytic domain. We have cloned a fragment of the protein encompassing the residues representing the catalytic domain of Rv1625c based on an alignment of the Rv1625c sequence and multiple adenylyl and guanylyl cyclases (Shenoy *et al.*, 2003). The expressed and purified protein showed adenylyl cyclase activity, despite having higher sequence similarity (60% similarity) to the guanylyl cyclases, and the protein formed a dimer constitutively (Shenoy *et al.*, 2004). Interestingly, the critical residues required for substrate specificity classified Rv1625c as an adenylyl cyclase. Mutational studies have shown that it is possible to completely change the nucleotide specificity of

a guanylyl cyclase and convert an adenylyl cyclase into a non-selective purine nucleotide cyclase (Liu *et al.*, 1997; Sunahara *et al.*, 1998). We therefore mutated the substrate-specificity residues in Rv1625c to the residues that are present in guanylyl cyclases in an attempt to convert the enzyme to a guanylyl cyclase. However, the mutant enzymes showed a dramatic loss in adenylyl cyclase activity with no increase in guanylyl cyclase activity (Shenoy *et al.*, 2003). Interestingly, the dimeric nature of the mutant enzymes was altered, suggesting that substrate-specifying residues also play a role in the juxtapositioning of the two catalytic domains. Computational modelling of the catalytic domain of Rv1625c identified a phenylalanine residue which, when mutated to the arginine found in guanylyl cyclases, would have decreased hydrophobic interactions and would therefore account for the monomeric nature of the mutant protein (Shenoy *et al.*, 2004).

The biochemical properties of the Rv1625c catalytic domain were not identical to those seen in the mammalian enzymes, suggesting that subtle differences could possibly be observed after crystallization and determination of the structure. While attempts have been made to crystallize the wild-type and mutant forms of Rv1625c, we report here the crystallization and preliminary X-ray diffraction analysis of one of the mutants of Rv1625c (KFD→ERC) that is predominantly a monomer in solution.

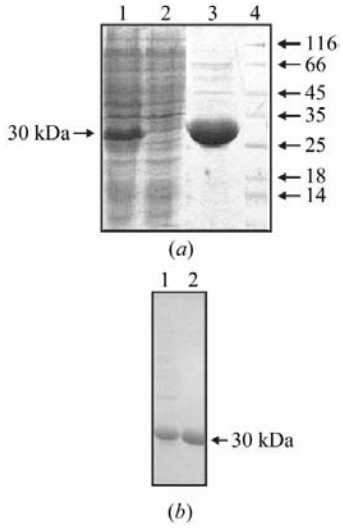

Figure 1
Coomassie-stained 12% SDS gel of different pools during purification of KFD→ERC. (*a*) Lane M, molecular-weight markers (kDa); lane 1, crude lysate (soluble protein band seen at 30 kDa); lane 2, lysate after interaction with Ni–NTA agarose beads (the 30 kDa protein has bound to the column and hence is not seen in the lane); lane 3, purified KFD→ERC after metal-affinity purification. (*b*) Lane 1, affinity-purified KFD→ERC after desalting into the anion-exchange buffer; lane 2, purified (>95%) KFD→ERC after gel filtration.

2. Materials and methods

2.1. Protein expression and purification

Residues encompassing the catalytic domain of Rv1625c (212–443) were cloned as described elsewhere and site-directed mutagenesis (Shenoy & Visweswariah, 2003) was performed to generate a triple mutant K296E/F363R/D365C which has the substrate-specifying residues present in guanylyl cyclases (Shenoy *et al.*, 2004). The plasmid was transformed into the C43 derivative of BL21 (DE3) *Escherichia coli* cells and grown in Terrific Broth; protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside to a concentration of 0.5 m*M*. Cells were harvested following 3 h induction at 310 K, washed with phosphate-buffered saline and resuspended in lysis buffer [20 m*M* Tris–HCl pH 8.0, 0.1 *M* NaCl, 5 m*M* 2-mercaptoethanol (2-ME), 2 m*M* phenylmethylsulfonyl fluoride]. Cells were lysed by sonication and the lysate was centrifuged at 20 000*g* at 277 K for 20 min. The soluble N-terminally hexahistidine-tagged protein in the supernatant was purified by metal-affinity chromatography on an Ni–NTA agarose (Qiagen) column. The column was washed with buffer containing 20 m*M* Tris–HCl pH 8.0, 5 m*M* 2-ME, 0.5 *M* NaCl and subsequently with 20 m*M* Tris–HCl pH 8.0, 5 m*M* 2-ME, 0.1 *M* NaCl, 10 m*M* imidazole and was eluted in 20 m*M* Tris–HCl pH 8.0 buffer containing 0.1 *M* NaCl, 5 m*M* 2-ME, 150 m*M* imidazole and 10% glycerol. The eluted protein, which was 85–90% pure, was desalted in the same buffer but without imidazole and salt and then subjected to anion-exchange chromatography on an ÄKTA-FPLC system (Pharmacia Biotech) using a Q-Sepharose column (Pharmacia Biotech). A linear salt gradient of 0–0.5 *M* NaCl was applied and the pure protein eluted in a single peak at 0.24 *M* NaCl. This was then subjected to gel filtration on a Superdex 200 column (25 × 1 cm, Pharmacia Biotech) in 20 m*M* HEPES–NaOH buffer pH 7.5 containing 10% glycerol and 5 m*M* 2-ME. The pure protein eluted as a monomer. The peak fractions were pooled and concentrated using an Amicon (Millipore) concentrator to 70 mg ml^{-1} and were stored in aliquots at 193 K.

2.2. Crystallization

Crystallization trials were set up using the hanging-drop vapour-diffusion method. Initial screening was performed using sparse-matrix screens from Hampton Research. Prior to setting up trials, the frozen protein solution was thawed and diluted to 10 mg ml^{-1} in dilution buffer (20 m*M* HEPES–NaOH pH 7.5 containing 5 m*M* 2-ME). Needle-shaped crystals were obtained from five different conditions in which polyethylene glycol (PEG) 8000, PEG 6000, PEG 20 000, ethylene glycol and sodium/potassium tartrate were used as the precipitants. Optimization of these conditions to improve the quality of crystals was carried out by varying the protein concentration, precipitant concentration and pH as well as by microseeding. The quality improved upon increasing the protein concentration from 10 to 30 mg ml^{-1}. Diffraction-quality crystals were obtained by mixing 2 µl protein at 30 mg ml^{-1} with 2 µl reservoir solution. The reservoir contained 500 µl of 8% PEG 8000 in 0.1 *M* Tris–HCl pH 8.5. Crystals appeared in 2–3 d and grew to their full size in about 7 d.

2.3. X-ray diffraction and data collection

One of the crystals grown as described above was washed in mother liquor and mounted in a glass capillary for data collection. X-ray diffraction data were collected at room temperature (293 K) with a MAR Research imaging plate mounted on a Rigaku RU-200 rotating anode X-ray generator providing Cu $K\alpha$ radiation. The crystal-to-detector distance was 120 mm. 90 frames, each with a 20 min exposure time, were collected with a 1.0° oscillation range. The images were processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

The catalytic domain of Rv1625c is expressed at high levels in *E. coli* (Shenoy *et al.*, 2003), but we failed to obtain crystals using the wild-type protein. We attribute this to the presence of a mixture of different oligomers of the catalytic domain in the purified protein preparation (Shenoy *et al.*, 2003). However, the mutant (K296E/F363R/D365C) catalytic domain of Rv1625c was also expressed to equivalent levels in a soluble form in *E. coli* and could be purified to near-homogeneity by metal-affinity chromatography, ion exchange and gel filtration (Fig. 1). Interestingly, the Rv1625c mutant was found to exist mainly in a monomeric form. We used it for crystallization in order to gain insights into the structure of the wild-type protein. This highly purified (>95%)

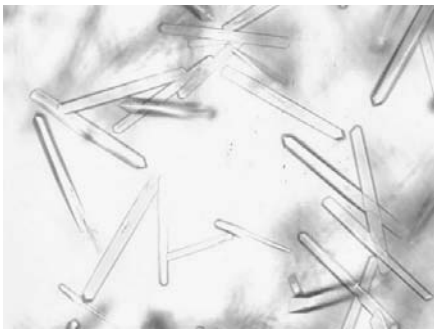

Figure 2
Crystals of KFD→ERC grown by the hanging-drop method. Largest crystal size is 1.0 × 0.1 × 0.1 mm.

protein was used for crystallization and crystals were obtained using the conditions described above (Fig. 2). Crystal data and the details of data collection are given in Table 1. The solvent content of 34.2% and the Matthews coefficient of 1.9 $Å^3\,Da^{-1}$ obtained by assuming the presence of one molecule per asymmetric unit are within the range normally observed for protein crystals (Matthews, 1968). Since the tetragonal cell can accommodate only one cyclase molecule in the asymmetric unit and since 00*l* reflections were not measured, the space group must be either $P4$ or $P4_1$ or $P4_2$ or $P4_3$. Attempts were made to solve the structure by molecular replacement using *AMoRe* (Navaza, 1994) using the coordinates of the monomers of the C2–C2 homodimer (Zhang *et al.*, 1997; PDB code 1ab8; ∼56% sequence similarity) and the C1–C2 heterodimer (Tesmer *et al.*, 1997; PDB code 1azs; ∼56% sequence similarity) of the membrane-bound mammalian adenylyl cyclase as search models. However, a satisfactory solution could not be obtained. Finally, we used the previously generated theoretical model of the wild-type Rv1625c catalytic domain (Shenoy *et al.*, 2003) as the search molecule and a unique solution was obtained which confirmed the presence of one molecule per asymmetric unit. The theoretical model was generated using the coordinates of the C2 homodimer (PDB code 1ab8) and coordinates of both the C1 as well as the C2 subunits (PDB code 1azs). Unlike the mammalian 12-transmembrane adenylyl cyclases, homodimeric cyclases such as Rv1625c have both C1-like and C2-like regions within a single polypeptide chain. The theoretical model, which represents an average of the two crystal structures with the exact sequence incorporated into it, may have therefore provided a better search model than the individual structures. The best values for the correlation coefficient (36.5) and *R* factor (48.3%) were obtained for the space group $P4_1$ using data in the 15–4.5 Å resolution range. There were no short contacts between molecules when the packing was checked for this solution. Therefore, the space group was fixed as $P4_1$. Electron-density maps were calculated after a few cycles of refinement. Further refinement and model building are in progress.

Table 1
Diffraction statistics for KFD→ERC.

Values in parentheses are for the last resolution shell (3.52–3.4 Å).

Space group	$P4_1$
Resolution range (Å)	15–3.4
Unit-cell parameters	
a (Å)	71.25
b (Å)	71.25
c (Å)	44.51
No. observations	13605
No. unique reflections	3192
Completeness (%)	98 (99.7)
$R_{\rm merge}$ (%)	10 (22.7)
Average $I/\sigma(I)$	10.4 (6.4)
No. molecules per AU	1

The intensity data were collected at the X-ray Facility for Structural Biology at the Institute, supported by the Department of Science and Technology (DST) and the Department of Biotechnology (DBT), Government of India. We acknowledge financial assistance from the DST and the DBT, Government of India and the Wellcome Trust, UK. ADK acknowledges financial support from the Council of Scientific and Industrial Research (CSIR).

References

Barzu, O. & Danchin, A. (1994). *Prog. Nucleic Acid Res. Mol. Biol.* **49**, 241–283.
Bieger, B. & Essen, L.-O. (2001). *EMBO J.* **20**, 433–445.
Guo, Y. L., Seebacher, T., Kurz, U., Linder, J. U. & Schultz, J. E. (2001). *EMBO J.* **20**, 3667–3675.
Katayama, M. & Ohmori, M. (1997). *J. Bacteriol.* **179**, 3588–3593.
Liu, Y., Ruoho, A. E., Rao, V. D. & Hurley, J. H. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 13414–13419.
McCue, L. A., McDonough, K. A. & Lawrence, C. E. (2000). *Genome Res.* **10**, 204–219.
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.
Reddy, S. K., Kamireddi, M., Danireddi, K., Young, L., Davis, A. & Reddy, P. T. (2001). *J. Biol. Chem.* **276**, 35141–35149.
Shenoy, A. R., Sivakumar, K., Krupa, A., Srinivasan, N. & Visweswariah, S. S. (2004). In the press.
Shenoy, A. R., Srinivasan, N., Subramaniam, M. & Visweswariah, S. S. (2003). *FEBS Lett.* **545**, 253–259.
Shenoy, A. R. & Visweswariah, S. S. (2003). *Anal. Biochem.* **319**, 335–336.
Sunahara, R. K., Beuve, A., Tesmer, J. J. G., Sprang, S. R., Garbers, D. L. & Gilman, A. G. (1998). *J. Biol. Chem.* **273**, 16332–16338.
Tang, W. J. & Hurley, J. H. (1998). *Mol. Pharmacol.* **54**, 231–240.
Tesmer, J. J., Sunahara, R. K., Gilman, A. G. & Sprang, S. R. (1997). *Science*, **278**, 1907–1916.
Tesmer, J. J., Sunahara, R. K., Johnson, R. A., Gosselin, G., Gilman, A. G. & Sprang, S. R. (1999). *Science*, **285**, 756–760.
Zhang, G., Liu, Y., Ruoho, A. E. & Hurley, J. H. (1997). *Nature (London)*, **386**, 247–253.
Zimmermann, G., Zhou, D. & Taussig, R. (1998). *J. Biol. Chem.* **273**, 19650–19655.