Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic analysis of chorismate synthase from *Mycobacterium tuberculosis*

Marcio Vinicius Bertacine Dias,[a] Fernanda Ely,[b] Fernanda Canduri,[a,b] José Henrique Pereira,[a] Jeverson Frazzon,[b] Luiz Augusto Basso,[b] Mário Sérgio Palma,[c] Walter Filgueira de Azevedo Jr[a,b]* and Diógenes Santiago Santos[d]*

[a]Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil, [b]Rede Brasileira de Pesquisas em Tuberculose Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil, [c]Laboratory of Structural Biology and Zoochemistry, CEIS/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil, and [d]Centro de Pesquisas em Biologia Molecular e Funcional/ Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS 90619-900, Brazil

Correspondence e-mail: walterfa@df.ibilce.unesp.br, diogenes@pucrs.br

Received 15 July 2004
Accepted 10 August 2004

The enzymes of the shikimate pathway are potential targets for the development of new therapies because they are essential for bacteria but absent from mammals. The last step in this pathway is performed by chorismate synthase (CS), which catalyzes the conversion of 5-enolpyruvylshikimate-3-phosphate to chorismate. Optimization of crystallization trials allowed the crystallization of homogeneous recombinant CS from *Mycobacterium tuberculosis* (*Mt*CS). The crystals of *Mt*CS belong to space group $P6_422$ (or $P6_222$) and diffract to 2.8 Å resolution, with unit-cell parameters $a = b = 129.7$, $c = 156.8$ Å. There are two molecules in the asymmetric unit. Molecular-replacement trials were not sucessful. Heavy-atom derivative screening is in progress.

1. Introduction

Tuberculosis is the second leading cause of deaths worldwide, killing nearly 2 million people each year. Most cases are in under-developed countries; over the past decade, tuberculosis incidence has increased in Africa, mainly as a result of the burden of HIV infection, and in the former Soviet Union, owing to socioeconomic change and decline of the health-care system (Frieden *et al.*, 2003). In 1993, the gravity of the situation led the World Health Organization (WHO) to declare tuberculosis a global emergency in an attempt to heighten public and political awareness (Cole *et al.*, 1998). Thus, newer and more efficient anti-tuberculosis drugs are needed. Potential targets for the development of new therapies are the enzymes of the shikimate pathway, because they are essential for bacteria, fungi and apicomplexan parasites, but absent from mammals (Bentley, 1990; Roberts *et al.*, 1998). In microorganisms, the shikimate pathway is used to synthesize the three proteinogenic aromatic amino acids phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp), the folate coenzymes benzoid and naphtoid quinones and a broad range of mostly aromatic secondary metabolites, including siderophores (Dosselaere & Vanderleyden, 2001). This pathway consists of seven enzymes that catalyse the sequential conversion of erythrose-4-phosphate and phosphoenol pyruvate to chorismate, the common precursor of aromatic compounds (Herrmann & Weaver, 1999). The last step in this pathway is performed by chorismate synthase, which catalyzes the conversion of 5-enolpyruvylshikimate 3-phosphate to chorismate *via* the 1,4-*anti*-elimination of phosphate and a proton, a reaction that is unique in nature. The enzyme has an absolute requirement for reduced FMN as a cofactor, although the 1,4-*anti*-elimination of phosphate and the C(6-*pro-R*) H atom does not involve a net redox change. The role of the reduced FMN in catalysis has long been elusive. However, recent detailed kinetic and bioorganic approaches have fundamentally advanced our understanding of the mechanism of action of chorismate synthase (Macheroux *et al.*, 1999).

Structural information is necessary for structure-based design in drug discovery, for a better understanding of the catalytic mechanism and also for an understanding of protein–ligand interaction (de Azevedo *et al.*, 1996, 1997, 2002). Although structures of CS from some microrganisms have been determined (Ahn *et al.*, 2004; Maclean & Ali, 2003), the crystal structure of CS from *Mycobacterim tuberculosis* has not been described. The crystal structure of shikimate-pathway enzymes will help in the development of new drugs against tuberculosis and other infectious diseases (de Azevedo *et al.*, 2002; Pereira *et al.*, 2003). In this work, we initiated the structure determination of the *aroF*-encoded *Mt*CS, which is composed of 401 residues with a molecular weight of 41 792 Da. We report here the crystallization and the preliminary X-ray crystallographic study of *Mt*CS.

2. Materials and methods

2.1. Cloning, protein expression and purification

Synthetic oligonucleotide primers were designed based on the *aroF* structural gene sequence of *M. tuberculosis* H37Rv (Cole *et al.*, 1998), containing 5′ *Nde*I and 3′ *Bam*HI restriction sites. The PCR product was cloned

Table 1
Summary of data-collection statistics for *Mt*CS.

Values in parentheses are for the highest resolution shell.

X-ray wavelength (Å)	1.427
Unit-cell parameters	
a (Å)	129.74
b (Å)	129.74
c (Å)	156.77
Space group	$P6_422$ or $P6_222$
No. measurements with $I > 2\sigma(I)$	92610
No. independent reflections	19341 (2716)
Completeness in the resolution range 55.8–2.8 Å (%)	97.9 (97.9)
R_{sym}† (%)	5.6 (16.5)
$\langle I/\sigma(I)\rangle$	3.8 (4.0)
Highest resolution shell (Å)	2.94–2.80

† $R_{sym} = \sum_h \sum_i |I(h)_i - \langle I(h)\rangle| / \sum_h \sum_i I(h)_i$, where $I(h)$ is the intensity of reflection h, $\sum_h$ is the sum over all reflections and $\sum_h$ is the sum over i measurements of reflection h.

into pET-23a(+) (Novagen, USA) expression vector and transformed into *Escherichia coli* Rosetta(DE3) host cells. Transformed *E. coli* cells were grown in LB medium at 310 K for 18 h in the absence of IPTG, as used in previously reported expression protocols (Oliveira *et al.*, 2001; Grossman *et al.*, 1998), and harvested by centrifugation. Although it is often argued that the cost of IPTG limits the usefulness of the *lac* promoter to high-added-value products, it has previously been shown that high levels of expression could be obtained with pET vectors as cells entered the stationary phase without the addition of inducer in LB medium. It has been demonstrated that when DE3 hosts are grown to stationary phase in media lacking glucose, cyclic AMP mediated derepression of both the wild type and *lacUV5* promoters occurs (Oliveira *et al.*, 2001). These authors also proposed that cyclic AMP, acetate and low pH are required for high-level expression in the absence of IPTG induction when cells approach the stationary phase in complex media and that derepression of the *lac* operon in the absence of IPTG may be part of a general cellular response to nutrient limitation.

The cell pellet was resuspended in 50 m*M* Tris–HCl pH 7.8 (Sigma Co., USA), sonicated and centrifuged to remove cell debris. The protein-purification protocol included the following steps: streptomycin sulfate precipitation, ammonium sulfate precipitation, Q Sepharose anion-exchange chromatography, phenyl Sepharose hydrophobic interaction chromatography and Mono Q anion-exchange column (Amersham Pharmacia Biotech, UK). A detailed description of the cloning and expression of recombinant *Mt*CS, the protein-purification protocol, N-terminal sequencing, mass spectrometry and determination of the oligomeric state of homogeneous *M. tuberculosis* chorismate synthase will be given elsewhere (manuscript in preparation).

Figure 1
Hexagonal crystals of *Mt*CS. Approximate dimensions are 0.30 × 0.25 × 0.25 mm.

2.2. Crystallization

The purified *Mt*CS was concentrated and dialyzed against 50 m*M* Tris–HCl buffer pH 7.8 (Hampton Research, USA). The final protein concentration was about 10 mg ml^{-1}. Crystallization was performed by the hanging-drop vapour-diffusion and sparse-matrix methods (Jancarik & Kim, 1991) using tissue-culture multiwell plates with covers (Linbro, ICN Biomedicals, Inc, USA) at a temperature of 293 K. Each hanging drop was prepared by mixing 1 µl each of protein solution and reservoir solution and was placed over 700 µl reservoir solution. Initial conditions were screened using Crystal Screen I and II kits (Hampton Research, USA).

Crystal optimization was carried out by altering the concentration of the salt, precipitant and protein and also the ratio between the protein solution and reservoir solution.

2.3. X-ray data collection

A data set was collected at a wavelength of 1.427 Å using a synchrotron-radiation source (Station PCr, LNLS, Campinas-Brazil; Polikarpov *et al.*, 1998). This wavelength was previously set to optimize the overall data-collection statistics. The data set was collected from a single *Mt*CS crystal using a MAR CCD image-plate system. The crystal was looped out from the drop and flash-cooled. The PEG 400 present in the crystallization conditions served as a cryoprotectant. X-ray diffraction data were collected at a temperature of 100 K under a cold nitrogen stream generated and maintained with an Oxford Cryosystem. The crystal was rotated through a total of 160°, with a 1° oscillation range per frame, a crystal-to-detector distance of 130 mm and an exposure time of 60 s. Data were processed on a Silicon Graphics Octane2 computer using the programs *MOSFLM* (Leslie, 1990) and *SCALA* (Collaborative Computational Project, Number 4, 1994).

2.4. Molecular-replacement trials

Molecular-replacement trials were carried out with the program *AMoRe* (Navaza, 1994). The crystal structures of *Sp*Cs (PDB code 1qxo; Maclean & Ali, 2003) and *Aa*Cs (PDB code 1q1l; Viola *et al.*, 2004) were used as search models.

3. Results and discussion

The initial crystals were obtained with reservoir solution comprising 0.2 *M* magnesium chloride hexahydrate, 0.1 *M* Na HEPES pH 7.5 and 30%(*w*/*v*) PEG 400. The ratio of protein solution to well solution in the drop was 1:1. After optimization of these conditions, better crystals were obtained. These diffracting crystals grew from a reservoir solution containing 0.6 *M* magnesium chloride hexahydrate, 0.1 *M* Na HEPES pH 7.5 and 25%(*w*/*v*) PEG 400 and the protein solution was concentrated to 60 mg ml^{-1} in 50 m*M* Tris–HCl buffer pH 7.8. The ratio of protein solution to well solution in the drop was 2:1. The crystals grew reproducibly to approximate dimensions of 0.3 × 0.25 × 0.25 mm within 2 d (Fig. 1). The crystal diffracted to 2.8 Å with relatively low mosaicity (0.35°). Fig. 2 shows a typical X-ray diffraction pattern. A total of 92 610 measured reflections were merged

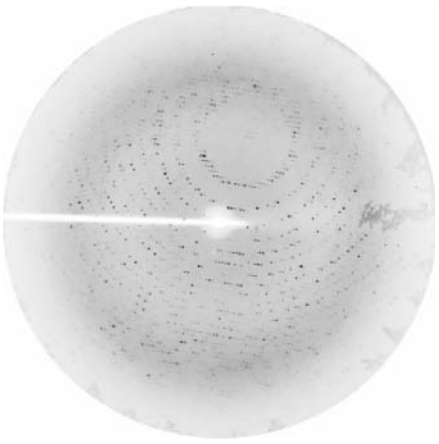

Figure 2
A typical diffraction pattern of the *Mt*CS crystal with 1° oscillation range. The crystal diffracts to 2.8 Å resolution.

into 19 341 unique reflections. The overall R_{sym} was 5.6% and the completeness was 97.9%. The crystal belongs to the hexagonal space group $P6_422$ (or $P6_222$), with unit-cell parameters $a = b = 129.7$, $c = 156.8$ Å. Assuming the asymetric unit content to be two monomers of molecular weight 41 792 Da, the V_M value is 2.28 Å^3 Da^{-1}. Assuming a value of 0.74 cm^3 g^{-1} for the protein partial specific volume, the calculated solvent content in the crystal is 45.96% (Matthews, 1968). Table 1 summarizes the data-collection statistics.

With the native data of *Mt*CS to 2.8 Å resolution, molecular replacement was used to attempt to solve the structure using *AMoRe* (Navaza, 1994). Various search models, including complete and modified structures of *Sp*CS (PDB code 1qxo) and *Aa*CS (1q1l), did not yield any meaningful results. This was most likely to be because of conformational differences between the search models and the structure under study. Heavy-atom screening is in progress.

This work was supported by grants from FAPESP (SMOLBNet, Proc. 01/07532-0, 02/04383-7, 04/00217-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT) to DSS and LAB. WFA (CNPq, 300851/98-7), MSP (CNPq, 500079/90-0), DSS (CNPq, 304051/1975-06), LAB (CNPq, 520182/99-5) and JF (CNPq, 301131/2003-01) are researchers for the National Research Council.

References

Ahn, H. J., Yoon, H. J., Lee, B. & Suh, S. W. (2004). *J. Mol. Biol.* **27**, 903–915.
Azevedo, W. F. de Jr, Canduri, F. & da Silveira N. J. F. (2002). *Biochem. Biophys. Res. Commun.* **293**, 566–571.
Azevedo, W. F. de Jr, de Oliveira, J. S., Basso, L. A., Palma, M. S., Pereira, J. H., Canduri, F. & Santos, D. S. (2002). *Biochem. Biophys. Res. Commun.* **295**, 142–148.
Azevedo, W. F. de Jr, Leclerc, S., Meijer, L., Havlicek, L., Strnad, M. & Kim, S.-H. (1997). *Eur. J. Biochem.* **243**, 518–526.
Azevedo, W. F. de Jr, Mueller-Dieckmann, H. J., Schulze-Gahmen, U., Worland, P. J., Sausville, E. & Kim, S.-H. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 2735–2740.
Bentley, R. (1990). *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–384.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dosselaere, F. & Vanderleyden, J. (2001). *Crit. Rev. Microbiol.* **27**, 75–131.
Frieden, T. R., Sterling, T. R., Munsiff, S. S., Watt, C. J. & Dye, C. (2003). *Lancet*, **362**, 887–899.
Grossman, T. H., Kawasaki, E. S., Punreddy, S. R. & Osburne, M. S. (1998). *Gene*, **209**, 95–103.
Herrmann, K. M. & Weaver, L. M. (1999). *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* **50**, 473–503.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Leslie, A. G. W. (1990). *Crystallographic Computing*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 27–38. Oxford University Press.
Macheroux, P., Schmid, J., Amrhein, N. & Schaller, A. (1999). *Planta*, **207**, 325–334.
Maclean, J. & Ali, S. (2003). *Structure*, **11**, 1499–1511.
Matthews, B. W. (1968). *J. Mol. Biol* **33**, 491–497.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Oliveira, J. S., Pinto, C. A., Basso, L. A. & Santos, D. S. (2001). *Protein Expr. Purif.* **22**, 430–435.
Pereira, J. H., Canduri, F., de Oliveira, J. S., da Silveira, N. J., Basso, L. A., Palma, M. S., de Azevedo, W. F. Jr & Santos, D. S. (2003). *Biochem. Biophys. Res. Commun.* **312**, 608–614.
Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. C. & Craievich, A. (1998). *J. Synchrotron Rad.* **5**, 72–76.
Roberts, F., Roberts, C. W., Johnson, J. J., Kyle, D. E., Krell, T., Coggins, J. R., Coombs, G. H., Milhous, W. K., Tzipori, S., Ferguson, D. J., Chakrabarti, D. & McLeod, R. (1998). *Nature (London)*, **395**, 801–805.
Viola, C. M., Saridakis, V. & Christendat, D. (2004). *Proteins*, **54**, 166–169.