crystallization papers

Expression, purification, crystallization and preliminary crystallographic analysis of a putative GTP-binding protein, YsxC, from *Bacillus subtilis*

Sanjan K. Das, Svetlana E. Sedelnikova, Patrick J. Baker, Sergey N. Ruzheinikov, Simon J. Foster and David W. Rice*

The Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, England

Correspondence e-mail: d.rice@sheffield.ac.uk

Bacillus subtilis YsxC has been putatively identified as a member of the GTP-binding protein family. Gene-knockout/deletion analysis has suggested that this protein is essential for survival of the microorganism and hence may represent a target for the development of a novel anti-infective agent. The *B. subtilis ysxC* gene was cloned and the protein was overexpressed in *Escherichia coli* and subsequently purified. Using hanging-drop vapour-diffusion crystallization techniques, two different crystal forms of YsxC were obtained in the presence and absence of GDP and which have one and two copies of YsxC in the asymmetric unit, respectively. Both crystal forms diffract to beyond 2.0 Å resolution and are suitable for structure determination.

1. Introduction

The alarming rise of new antibiotic resistant strains of pathogenic bacteria (Kotra *et al.*, 2000) is creating a clear need for the development of new effective antibacterial agents for therapeutic use. The completion of the genome sequence of number of important pathogenic bacteria (Cole *et al.*, 1998) has opened up new opportunities for the development of new anti-infective agents using structure-based drug design (Kunst *et al.*, 1997) against proteins whose role is known to be essential for bacterial survival. As a contribution to this area, we have established a programme of structure determination on a number of essential bacterial gene products.

The *Bacillus subtilis* gene *ysxC*, also referred to as *orfX* (Riethdorf *et al.*, 1994), was originally identified as one of a number of putative essential genes of unknown function following a functional genomic analysis involving a genome-wide gene-inactivation program of this microorganism (Pragai & Harwood, 2000). Thus, whilst the gene could not be inactivated directly, a conditional lethal *ysxC* construct was recovered in which viability was inducer-dependent (Pragai & Harwood, 2000).

The *ysxC* gene encodes a 195 amino-acid 21.9 kDa molecular-weight protein and homologues can be identified in various pathogenic bacteria such as S*taphylococcus aureus*, *Pseudomonas aeruginosa* and *Neisseria meningitidis*. Comparison of the sequence of YsxC from a range of microorganisms has revealed four blocks of very significant sequence conservation. More widely, a global comparison of the YsxC sequence using BLAST (Altschul *et al.*, 1997) reveals that the protein shares limited sequence similarity to the superfamily of GTP-binding proteins. Structural and biochemical studies have shown that these diverse guanine nucleotide-binding proteins commonly function as GTPases and share a core functional domain formed by a central β-sheet (Sprang, 1997). Sequence alignment of family members has identified five major regions of sequence similarity which are associated with aspects of GTP binding and catalysis (G1, G2, G3, G4 and G5; Bourne *et al.*, 1991). Whilst the overall level of sequence similarity between these GTPases and YsxC is low, sequence comparisons have shown that the four blocks of conservation in the family of the latter clearly align with regions G1–G4, including a perfect consensus sequence for the G1 region (GX_4GKS) which defines the P-loop responsible for the recognition of the phosphate moiety of the nucleotide. This suggests that YsxC can be tentatively identified as a GTP-binding protein/GTPase (Pragai & Harwood, 2000). More recently, the family of putative prokaryotic GTPases has been extended by the addition of new members which differ in the number and size of the GTP-binding domains (Caldon *et al.*, 2001). Nevertheless, comparisons of the sequences prokaryotic GTPases reveal that their relationship to their eukaryotic homologues is remote, as the overall level of sequence identity is low (Pragai & Harwood, 2000).

In many living organisms, GTPases act as a major signalling mechanism, operating as molecular switches *via* the hydrolysis of GTP (Sprang, 1997). They are involved in many critical functions, including translation of mRNA, cell cycling, cell division and signal transduction or hormone signalling. This suggests that YsxC might be involved in a critical signalling pathway, although at present

the identity of the pathway is unknown. In *B. subtilis*, *ysxC* is encoded as part of a bicistronic operon which also includes the *lonA* gene that encodes an ATP-dependent serine endopeptidase (Riethdorf *et al.*, 1994) and there are strong indications that the transcription of *ysxC* and *lonA* are coupled, as the start codon of *ysxC* overlaps the *lonA* coding sequence. This suggests that YsxC may be involved in an intracellular signalling process which controls protein turnover in response to changing environmental conditions. In this paper, we report the cloning, purification, crystallization and initial X-ray analysis of *B. subtilis* YsxC.

2. Materials and methods

The *ysxC* gene from *B. subtilis* 168 was PCR amplified from genomic DNA using appropriate primers designed using the *WEBCUTTER* program (http://www.firstmarket.com/cutter/cut2.html). The PCR product of 585 base pairs was purified (Qiagen) and a parallel restriction digest (310 K overnight) of the PCR product and the plasmid vector pTB361 was performed to produce linear fragments with sticky ends suitable for ligation. The restriction products were purified again (Qiagen) and ligated together using a gene:plasmid DNA ratio of 4:1 at 287 K. After transformation of the ligation mix into *Escherichia coli* DH5α, a positive clone (pSKD02) was verified by performing both PCR screening and a restriction-enzyme digest.

To obtain selenomethionine-labelled protein, YsxC was overexpressed in *E. coli* strain B834 transformed with plasmid pSKD02. A seed batch culture of 50 ml LB with 12.5 µg ml^{-1} tetracycline was inoculated with a single colony of bacteria from a freshly made agar plate which had previously been streaked with bacteria and the culture was grown overnight at 310 K with vigorous aeration. This seed culture was used to inoculate two batches of 500 ml sterile LB media and grown at 310 K to an OD$_{600}$ of 1.15. The culture was pelleted at room temperature by centrifugation at 4000g and resuspended in minimal media (M9 salts, 1 mM MgSO$_4$, 20% glucose and 1 mM CaCl$_2$). This cell suspension was then used to innoculate 3.5 l of minimal media containing 150 µl of antifoam in a 5 l benchtop fermenter with external aeration to stop excessive foaming and grown at 310 K to an OD$_{600}$ of approximately 0.3. At this point of growth, the media was supplemented with 100 µg ml^{-1} lysine, phenylalanine and threonine, 50 µg ml^{-1} valine, isoleucine and leucine and 40 µg ml^{-1} selenomethionine.

The cells were grown for a further 15 min and 1 mM IPTG was then added to induce expression of YsxC. The culture was then fermented for a further 16 h. The cells were then harvested by centrifugation at 5000g for 20 min at 277 K and the pellets were frozen immediately and stored at 253 K.

In order to purify YsxC, about 3 g cell paste harvested from 1 l culture was suspended in 30–40 ml buffer *A* (50 mM Tris–HCl pH 8.0) and the cells were disrupted by ultrasonication at 16 µm amplitude for 3 × 20 s. The lysate was centrifuged to remove the debris at 70 000g for 10 min. The supernatant fraction was loaded onto a column packed with DEAE-Sepharose Fast Flow (Pharmacia) equilibrated with buffer *A*. YsxC does not bind to the column and was collected as a pass-through fraction. 4 *M* ammonium sulfate solution was then added to the sample to bring it to a concentration of 1.5 *M*. The protein solution was then loaded onto a column packed with Phenyl-Toyopearl 650S (Tosoh) equilibrated with 1.2 *M* ammonium sulfate in buffer *A* and proteins were eluted by a reverse gradient of ammonium sulfate concentration from 1.2 to 0 *M*. The protein concentration was checked in each second fraction by the method of Bradford (1976) using Bio-Rad Dye Reagent and analysed by SDS–PAGE (4–12% bis-tris NuPAGE gel, Novex). The fractions containing YsxC were combined and concentrated on a Viva Spin concentrator with a 10 kDa molecular-weight cutoff (Viva Science) to a volume of 1.5 ml and loaded on a Hi-Load Superdex-200 column (Pharmacia) equilibrated with buffer *B* (0.75 *M* NaCl, 50 mM sodium phosphate pH 7.0) and eluted at a flow rate of 1 ml min^{-1}.

For crystallization, the protein was concentrated to 30 mg ml^{-1} using a Viva Spin concentrator and was diluted threefold with Milli-Q water before setting up crystallization trials. Trials were set up with YsxC alone and in complex with GDP using Crystal Screens I and II (Hampton Research). X-ray diffraction data were collected in all cases from crystals flash-frozen in a stream of nitrogen gas at 100 K using an Oxford Cryosystems Cryostream device and 15 and 20% glycerol as a cryoprotectant for the crystals of the apo protein and the GDP complex, respectively. X-ray diffraction data were collected by the rotation method using a MAR345 detector with dual mirror focused Cu Kα X-rays produced by a Rigaku RU-200 rotating-anode generator with a 0.3 × 3 mm filament running at 50 kV and 100 mA. The data sets were processed using the *DENZO/*

Table 1
X-ray data collection for the apo YsxC and its binary complex with GDP.

Values in parentheses are for the highest resolution shell.

	Apo	YsxC–GDP
Resolution (Å)	15–2.0	15–1.95
	(2.05–2.0)	(2.0–1.95)
Observed reflections	59608	22011
Unique reflections	24474	13194
Completeness (%)	96.6 (93.4)	93.9 (94.8)
R_{merge}† (%)	9.1 (39.2)	6.4 (33.0)
$\langle I/\sigma(I)\rangle$	9.9 (2.2)	10.5 (2.1)

† $R_{merge} = \sum_{hkl}(I_i - I_m)/\sum_{hkl} I_m$, where I_m is the mean intensity of the reflection.

SCALEPACK package (Otwinowski & Minor, 1997) and subsequently handled using *CCP*4 software (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

Analysis of the total cell extract of pre- and post-induction samples using SDS–PAGE indicated a high level of expression of YsxC, with a clear band at approximately 23 kDa on the gel. During gel filtration, YsxC elutes from the column at an approximate molecular weight of 20 kDa, suggesting that YsxC is monomeric in solution. A typical yield of YsxC was about 50 mg of protein from 1 l of culture (about 3 g of cell paste) and the purity of the protein was 85–90% as estimated by SDS–PAGE.

Crystals of the binary complex of YsxC with GDP were grown using the standard vapour-diffusion method by mixing 2 µl drops of the solution of protein at 10 mg ml^{-1} with an equivalent volume of precipitant (20% PEG 4000, 10% 2-propanol and 0.1 *M* Na HEPES buffer pH 7.5 including 5 mM GDP and 2 mM MgCl$_2$) and equilibrating the samples against the same precipitant at 290 K. The crystals grew within 2–3 d and had a plate-like morphology with dimensions up to 0.8 × 0.3 × 0.02 mm. Crystals of the apo protein were grown from crystallization conditions containing 0.2 *M* sodium acetate, 38% PEG 4000, Tris–HCl buffer pH 8.5 with 2 mM MnCl$_2$. These crystals again had a plate-like morphology with dimensions up to 0.3 × 0.1 × 0.02 mm. Analysis of both crystal forms showed that they diffract to high resolution.

Analysis of preliminary data sets collected on crystals of the apo protein and the YsxC–GDP binary complex with the autoindexing routine in the *DENZO/SCALEPACK* package (Otwinowski & Minor, 1997) are consistent with a *C*2 space group in each case but with different unit-cell parameters:

$a = 145.7$, $b = 36.5$, $c = 83.3$ Å, $\beta = 122.2°$ and $a = 46.4$, $b = 68.5$, $c = 61.9$ Å, $\beta = 93.9°$, respectively. Taking the subunit molecular weight to be 22 kDa, this suggests that the crystals of the apo protein contain two molecules in the asymmetric unit, with a V_M of 2.1 Å^3 Da^{-1} (Matthews, 1977), whilst those of the binary complex with GDP contain a monomer, with a V_M of 2.2 Å^3 Da^{-1}. Data-collection statistics are given in Table 1.

Our efforts are currently being directed towards solving the structure of the YsxC–GDP complex by using either multiple isomorphous replacement or MAD-based techniques. Ultimately, this structure may provide clues to the function of the protein and assist in the development of inhibitors as lead compounds in the search for novel antimicrobial agents.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Bourne, H. R., Sanders, D. A. & McCormick, F. (1991). *Nature (London)*, **349**, 117–127.
Bradford, M. M. (1976). *Annu. Rev. Biochem.* **72**, 248–254.
Caldon, C. E., Yoong, P. & March, P. E. (2001). *Mol. Microbiol.* **41**, 289–297.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Kotra, L. P., Golemi, D., Vakulenko, S. & Mobashery, S. (2000). *Chem. Ind.*, pp. 341–344.
Kunst, F. *et al.* (1997). *Nature (London)*, **390**, 249–256.
Matthews, B. W. (1977). *X-ray Structure of Proteins*, 3rd ed., edited by H. Neurath & R. L. Hill, Vol. 3, pp. 468–477. New York: Academic Press.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pragai, Z. & Harwood, C. R. (2000). *J. Bacteriol.* **182**, 6819–6823.
Riethdorf, S., Volker, U., Gerth, U., Winkler, A., Engelmann, S. & Hecker, M. (1994). *J. Bacteriol.* **176**, 6518–6527.
Sprang, S. (1997). *Annu. Rev. Biochem.* **66**, 639–678.