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

Cloning, purification and preliminary crystallographic analysis of a putative DNA-binding membrane protein, YmfM, from *Staphylococcus aureus*

Ling Xu, Svetlana E. Sedelnikova, Patrick J. Baker and David W. Rice*

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

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

Received 10 April 2008
Accepted 2 June 2008

The *Staphylococcus aureus* protein YmfM contains a helix–turn–helix motif and is thought to be a putative DNA-binding protein which is associated with the membrane through a C-terminal hydrophobic transmembrane anchor. Truncation of the protein by the removal of this C-terminal hydrophobic segment has enabled the overexpression of a soluble domain of *S. aureus* YmfM (ΔYmfM) in *Escherichia coli*, which has been purified and subsequently crystallized. Crystals of ΔYmfM diffract to beyond 1.0 Å resolution and belong to one of the pair of enantiomorphic tetragonal space groups $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = b = 45.5$, $c = 72.9$ Å and one molecule in the asymmetric unit. The crystals of ΔYmfM have an unusually low V_M of 1.6 Å^3 Da^{-1}, which is one of the lowest values observed for any protein to date. A full structure determination is under way in order to provide insights into the function of this protein.

1. Introduction

In *Staphylococcus aureus*, the gene SA1125 (hereinafter referred to as *ymfM*) encodes a 130-residue protein of which the C-terminal 25 residues are very hydrophobic and are predicted to form a transmembrane helical anchor. In the *S. aureus* genome *ymfM* forms part of an operon (Ermolaeva *et al.*, 2001) that encodes two putative membrane proteins, one of which (YmfL) is of unknown function, whilst the other, phosphatidylglycerophosphate synthase (PgsA), is an essential enzyme for cell survival in a range of organisms (Gerdes *et al.*, 2003; Kobayashi *et al.*, 2003; Martin *et al.*, 1999). This pattern of gene organization is very similar across a range of bacteria in which *ymfM* is located between *ymfL* and *pgsA*.

As a contribution towards understanding the structure–function relationship of *S. aureus* YmfM, we have initiated the determination of its three-dimensional structure. In this paper, we describe the cloning, overexpression, purification, crystallization and data collection of a soluble construct of the protein from which the putative transmembrane anchor (residues 105–130) has been removed.

2. Materials and methods

2.1. Cloning, overexpression and purification

Initial attempts to overexpress full-length YmfM in *Escherichia coli* resulted in slow cell growth and little evidence of expression. Therefore, it was decided to make a construct of the protein by removing the 25 C-terminal hydrophobic residues starting at Ile106. The *ymfM* gene fragment covering the putative soluble domain of YmfM (residues 1–105) was PCR-amplified directly from genomic DNA of *S. aureus* strain SH1000 with the primers TTGAAAACGGTCGGTGAAG (forward) and TTATGGCTCTTTTGATTTACTCTTATAATC (reverse). The purified DNA fragment (318 bp) was inserted into pETBlue1 vector using an AccepTor vector kit (Novagen). The positive clones were confirmed by blue/white selection and colony PCR and the extracted plasmid was transformed into *E. coli* Tuner (DE3) (Novagen). The transformed *E. coli* Tuner strain was grown in LB medium at 310 K with vigorous aeration until an OD_{600} of 0.6 was attained, at which point overexpression was induced

with 1 m*M* IPTG and growth continued for 5 h. The cells were harvested by centrifugation at 4500*g* for 20 min at 277 K.

For purification, cells were disrupted by sonication in 50 m*M* Tris–HCl pH 8.0. The cell debris and denatured proteins were removed by centrifugation at 20 000*g* for 20 min. Analysis of the soluble fraction by SDS–PAGE showed a large overexpression band corresponding to the expected molecular weight of the protein (12 kDa). The supernatant was collected and loaded onto a DEAE-Sepharose Fast Flow column (Amersham Biosciences) and the protein was eluted with a linear gradient of 0–0.5 *M* NaCl in 50 m*M* Tris–HCl pH 8.0. The fractions containing ΔYmfM were combined and subjected to gel-filtration chromatography using a Hi-Load Superdex 200 column (Amersham Biosciences) equilibrated with 0.5 *M* NaCl in 50 m*M* Tris–HCl pH 8.0 and eluted with the same buffer. The gel-filtration analysis shows that ΔYmfM runs with an approximate molecular weight of 15 kDa, suggesting that the protein is a monomer in solution. Peak fractions corresponding to ΔYmfM were concentrated to 18–20 mg ml^{-1} in a Vivaspin concentrator with a 5000 Da molecular-weight cutoff, filtered and buffer-exchanged to 10 m*M* sodim phosphate pH 5.6. Approximately 20 mg pure protein was obtained from 1 l culture, with the purity of the protein being estimated at greater than 95% as determined by SDS–PAGE. The molecular weight of ΔYmfM was confirmed by electrospray mass spectrometry and by conventional protein sequencing of the first 20 residues.

2.2. Crystallization and preliminary X-ray analysis

Preliminary crystallization conditions were screened robotically by the sitting-drop vapour-diffusion method using crystallization kits from both Hampton Research and Nextal (Qiagen). Initial small bipyramidal crystals were observed using 0.01 *M* zinc sulfate, 0.1 *M* MES pH 6.5 and 25% PEG 550 MME as the precipitant. Manual optimization by increasing the drop size using the same precipitant conditions led to larger crystals of overall dimensions 100 × 50 × 50 µm.

For data collection, a single crystal was flash-cooled at 100 K without further cryoprotection. Two data sets, each consisting of 360 images with 1° rotation per image, were collected to resolutions of 1.5 and 1 Å from the same crystal at 0.9793 Å at ID29, ESRF, Grenoble using an ADSC Q210 two-dimensional detector (Fig. 1). A further single-wavelength anomalous diffraction (SAD) data set was collected to a maximum resolution of 2.0 Å using a MAR CCD165 detector on beamline MAD10.1 at the Daresbury Synchrotron Radiation Source using a wavelength of 2.07 Å, near the theoretical sulfur absorption edge, in order to maximize the f'' component.

Table 1
X-ray data-collection statistics for crystals of the soluble domain of YmfM.

Values in parentheses are for the highest resolution shell.

Data set	Native	Sulfur SAD
Wavelength (Å)	0.9793	2.0700
Unit-cell parameters (Å, °)	$a = b = 45.5$, $c = 72.9$	$a = b = 45.5$, $c = 73.0$
Mosaicity (°)	1	1
Resolution (Å)	20–1.0 (1.05–1.0)	40–2.0 (2.1–2.0)
Reflections measured	265991 (7313)	91989 (10794)
Unique reflections	38933 (3512)	5192 (685)
Completeness (%)	92.5 (58.6)	93.9 (89.4)
Redundancy	6.8 (2.1)	17.1 (15.8)
$\langle I/\sigma(I)\rangle$	17.2 (3.1)	51.4 (30.2)
$R_{\rm merge}$† (%)	9.6 (18.5)	4.3 (8.5)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the observed intensity and mean intensity of related reflections, respectively.

3. Results and discussions

The two data sets were processed using the autoindexing routine in *MOSFLM* (Leslie, 1992) followed by scaling using *SCALA* (Evans, 1997) from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994) and revealed that the crystals belong to a primitive tetragonal system, point group 422, with unit-cell parameters $a = b = 45.5$, $c = 72.9$ Å. Details of the data-collection and processing statistics are presented in Table 1. Analysis of the systematic absences suggested that the space group is one of the enantiomorphic pair $P4_12_12$ or $P4_32_12$. The $V_{\rm M}$ for a monomer in the asymmetric unit is 1.6 Å^3 Da^{-1}, which is one of the lowest values ever observed for a protein and lower than the range observed by Matthews (1977). This indicates that the cell is tightly packed, with the volume of solvent in the crystal being approximately 21%. A full structure determination

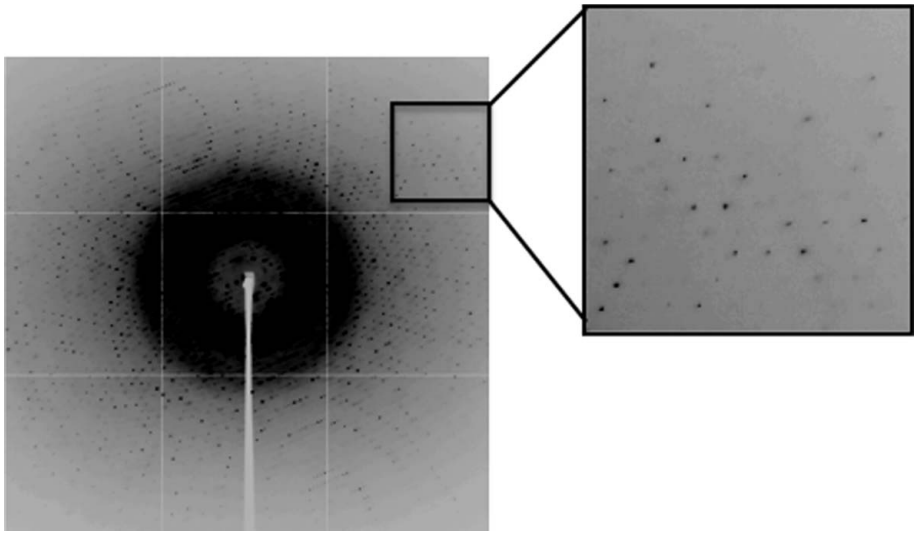

Figure 1
Diffraction image from a native ΔYmfM crystal recorded at ID29, ESRF, Grenoble. The resolution limit at the edge of the detector is 1.0 Å.

is under way in order to provide insights into the structure and possible molecular function of this protein.

This work was supported by BBSRC. LX wishes to thank the ORS scheme and the University of Sheffield for financial support.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Ermolaeva, M. D., White, O. & Salzberg, S. L. (2001). *Nucleic Acids Res.* **29**, 1216–1221.

Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.

Gerdes, S. Y. *et al.* (2003). *J. Bacteriol.* **185**, 5673–5684.

Kobayashi, K. *et al.* (2003). *Proc. Natl Acad. Sci. USA*, **100**, 4678–4683.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Martin, P. K., Li, T., Sun, D., Biek, D. P. & Schmid, M. B. (1999). *J. Bacteriol.* **181**, 3666–3673.

Matthews, B. W. (1977). *The Proteins*, edited by H. Neurath & R. L. Hill, Vol. 3, pp. 468–477. New York: Academic Press.