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

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 27 June 2006
Accepted 14 July 2006

Cloning, purification and preliminary crystallographic analysis of a conserved hypothetical protein, SA0961 (YlaN), from *Staphylococcus aureus*

SA0961 is an unknown hypothetical protein from *Staphylococcus aureus* that can be identified in the Firmicutes division of Gram-positive bacteria. The gene for the homologue of SA0961 in *Bacillus subtilis*, *ylaN*, has been shown to be essential for cell survival, thus identifying the protein encoded by this gene as a potential target for the development of novel antibiotics. SA0961 was cloned and the protein was overexpressed in *Escherichia coli*, purified and subsequently crystallized. Crystals of selenomethionine-labelled SA0961 diffract to beyond 2.4 Å resolution and belong to the monoclinic space group $P2_1$, with unit-cell parameters a = 31.5, b = 42.7, c = 62.7 Å, β = 92.4° and two molecules in the asymmetric unit. A full structure determination is under way to provide insights into the function of this protein.

1. Introduction

In *Bacillus subtilis*, *ylaN* encodes an open reading frame for a hypothetical protein of unknown function compromising 93 residues that has been reported to be essential for cell survival (Kobayashi *et al.*, 2003). In *Staphylococcus aureus*, the YlaN homologue SA0961 consists of 91 residues with 57% sequence identity to the *B. subtilis* protein. Like its counterpart in *B. subtilis*, the sequence of SA0961 (referred to hereafter as YlaN), is particularly rich in leucine (18.7%) and lysine (11%) compared with the average composition of the set of proteins encoded in the *S. aureus* or *B. subtilis* genomes. A *BLAST* search with the sequence of *S. aureus* YlaN against the NCBI database revealed 19 sequences from a range of closely related Gram-positive bacteria that belong to the Firmicutes division, a cluster of Gram-positive microbes with a low G+C content.

In the genome of both *B. subtilis* and *S. aureus*, YlaN is located upstream from *ftsW*, an essential gene in *B. subtilis* (Kobayashi *et al.*, 2003) whose protein product is involved in cell division. However, the intergenic distance of *ylaN* and *ftsW* (about 400 bp) suggests that they are probably not part of the same operon (Wang *et al.*, 2004).

As a contribution towards understanding the structure–function relationships of *S. aureus* YlaN, we have initiated the determination of its three-dimensional structure. In this paper, we describe the cloning, overexpression, purification and crystallization of YlaN and the preliminary analysis of data collected from selenomethionine-containing (SeMet) crystals.

2. Materials and methods

2.1. Cloning, overexpression and purification

The *ylaN* gene fragment was PCR amplified directly from genomic DNA of *S. aureus* strain SH1000 with the primers TTGAAAACGGTCGGTGAAG (forward) and TTAAAATATTAAAACTAACATGATCCATAAC (reverse). The purified DNA fragment (273 bp) was inserted into a 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 *Escherichia coli* Tuner (DE3) (Novagen). In order to produce wild-type or SeMet-incorporated YlaN protein, the

transformed *E. coli* Tuner strain was grown either in LB medium or in minimal medium containing 10.5 g l^{-1} K_2HPO_4, 1 g l^{-1} $(NH_4)_2PO_4$, 4.5 g l^{-1} KH_2PO_4, 0.5 g l^{-1} trisodium citrate·$2H_2O$, 5 g l^{-1} glycerol, 0.5 g l^{-1} adenine, guanine, thymine and uracil, 1 m*M* $MgSO_4$·$7H_2O$, 4 mg l^{-1} thiamine, 50 mg l^{-1} selenomethionine and 100 mg l^{-1} of the amino acids Lys, Phe and Thr in addition to 50 mg l^{-1} Ile, Leu and Val. Growth was carried out 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 was then continued for 5 h. The cells were harvested by centrifugation at 5000 rev min^{-1} for 20 min at 277 K. Analysis of the soluble fraction by SDS–PAGE showed a large overexpression band corresponding to the expected molecular weight of the protein (10 kDa).

For purification of either the wild-type or SeMet protein, 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 24 500 rev min^{-1} for 10 min. 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 YlaN were combined and 4.0 *M* $(NH_4)_2SO_4$ was added to give a final concentration of 1.5 *M*. Precipitated protein was removed by centrifugation at 24 500 rev min^{-1} for 10 min and the supernatant was subsequently loaded onto a column packed with Phenyl Toyo-Pearl650S (Tosoh) and eluted with a reverse gradient of $(NH_4)_2SO_4$ from 1.5 to 0 *M* in 50 m*M* Tris–HCl pH 8.0. The fractions containing YlaN were pooled 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. Gel-filtration analysis shows that wild-type YlaN runs with an approximate molecular weight of 20 kDa, suggesting that the protein is predominantly a dimer in solution, although higher molecular-weight aggregates could also be observed, particularly for the SeMet protein. Peak fractions corresponding to dimeric YlaN were concentrated to 18–20 mg ml^{-1} in a VivaSpin concentrator (5 000 Da molecular-weight cutoff) and buffer-exchanged to 10 m*M* Tris–HCl pH 8.0.

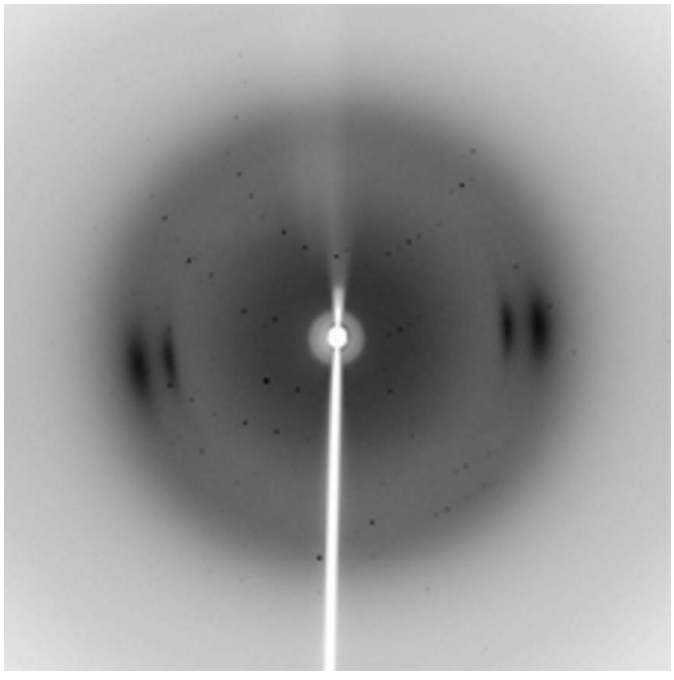

Figure 1
Diffraction image from a YlaN selenomethionine crystal recorded at station 10.1 at SRS Daresbury. The crystal diffracted beyond a resolution of 2.4 Å.

Table 1
X-ray data-collection statistics for YlaN SeMet crystals.

Values in parentheses are for the highest resolution shell.

Data set	Peak (λ_1)	Inflection (λ_2)	Remote (λ_3)
Wavelength (Å)	0.9795	0.9745	0.9802
unit-cell parameters (Å, °)	a = 31.5, b = 42.7, c = 62.7, β = 92.4	a = 31.5, b = 42.7, c = 62.4, β = 92.6	a = 31.4, b = 42.7, c = 62.4, β = 92.4
Resolution (Å)	25–2.4 (2.5–2.4)	25–2.4 (2.5–2.4)	25–2.4 (2.5–2.4)
Reflection measured	23940 (3486)	23369 (3419)	23363 (3424)
Unique reflections	6626 (961)	6582 (959)	6576 (952)
Completeness (%)	99.8 (99.8)	99.7 (99.7)	99.8 (99.6)
Redundancy	3.6 (3.6)	3.6 (3.6)	3.6 (3.6)
$I/\sigma(I)$	10.5 (2.6)	9.8 (2.2)	11.2 (1.8)
R_{merge}† (%)	7.9 (42.9)	8.4 (50.9)	9.0 (58.9)

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

Approximately 6 mg pure protein was obtained from 2 l of culture, with the purity of the SeMet protein being estimated at greater than 95% as determined by SDS–PAGE. The electrospray mass spectrum of the SeMet YlaN protein suggested that the N-terminal methioninie had been cleaved and this was confirmed by conventional protein sequencing. Analysis of the molecular weight by electrospray mass spectroscopy further suggested that the selenium incorporation of the sample was greater than 90%.

2.2. Crystallization and preliminary X-ray analysis

Preliminary crystallization conditions were screened by the hanging-drop vapour-diffusion method using Hampton Research crystallization kits. Initial small cuboid crystals were observed using 0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 8.5 and 30% PEG 4000 as the precipitant. Optimization of these conditions led to larger crystals of overall dimensions 100 × 100 × 100 µm from 0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 8.5 and 20% PEG 4000. SeMet crystals with an approximate dimensions 80 × 80 × 80 µm were obtained under the same conditions.

For data collection, a single crystal was flash-cooled in a cryo-protectant solution consisting of 0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 8.5 and 23% PEG 4000 and 20% glycerol at 100 K. Multiple-wavelength anomalous diffraction (MAD) data were collected from this crystal to a maximum resolution of 2.4 Å (Fig. 1) using a MAR CCD 165 detector on beamline MAD10.1 at the Daresbury Synchrotron Radiation Source (SRS). Three wavelengths were chosen near the selenium-absorption edge based on a fluorescence absorption spectrum obtained from the frozen crystal in order to maximize the f'' component (λ_1, peak), to minimize the f' component (λ_2, inflection) and to maximize $\Delta f'$ (λ_3, remote). A total 180 images, with 1° rotation per image, were collected at all three wavelengths.

3. Results and discussions

Analysis of the diffraction data using the autoindexing routine in *MOSFLM* (Leslie, 1992) and scaling in *SCALA* (Evans, 1997) from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994) shows that the crystals belong to space group $P2_1$, with unit-cell parameters a = 31.5, b = 42.7, c = 62.7 Å, $\alpha = \gamma = 90$, β = 92.4°. The corresponding unit-cell volume is 8.4 × 10^4 $Å^3$, which, assuming the asymmetric unit contains a dimer, gives a V_M value of 2.0 $Å^3$ Da^{-1}, which is within the range observed by Matthews for protein crystals

(Matthews, 1977). Details of the data-collection statistics are presented in Table 1. A full structure determination is under way to provide insights into the structure and possible molecular function of this protein.

This work was supported by the BBSRC. The Krebs Institute is a designated BBSRC Biomolecular Sciences Centre and a member of the North of England Structural Biology Centre. 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.
Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.
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**.
Matthews, B. W. (1977). *The Proteins*, edited by H. Neurath & R. L. Hill, Vol. 3, pp. 468–477. New York: Academic Press.
Wang, L., Trawick, J. D., Yamamoto, R. & Zamudio, C. (2004). *Nucleic Acids Res.* **32**, 3689–3702.