Expression, purification, crystallization and preliminary X-ray analysis of the PaaI-like thioesterase SAV0944 from *Staphylococcus aureus*

Yogesh B. Khandokar, Noelia Roman, Kate M. Smith, Parul Srivastava and Jade K. Forwood*

School of Biomedical Sciences, Charles Sturt University, Boorooma Street, Wagga Wagga, NSW 2650, Australia

Correspondence e-mail: jforwood@csu.edu.au

Received 23 September 2013
Accepted 7 January 2014

Staphylococcus aureus is the causative agent of many diseases, including meningitis, bacteraemia, pneumonia, food poisoning and toxic shock syndrome. Structural characterization of the PaaI-like thioesterase SAV0944 (*SaPaaI*) from *S. aureus* subsp. *aureus* Mu50 will aid in understanding its potential as a new therapeutic target by knowledge of its molecular details and cellular functions. Here, the recombinant expression, purification and crystallization of *SaPaaI* thioesterase from *S. aureus* are reported. This protein initially crystallized with the ligand coenzyme A using the hanging-drop vapour-diffusion technique with condition No. 40 of Crystal Screen from Hampton Research at 296 K. Optimal final conditions consisting of 24% PEG 4000, 100 mM sodium citrate pH 6.5, 12% 2-propanol gave single diffraction-quality crystals. These crystals diffracted to beyond 2 Å resolution at the Australian Synchrotron and belonged to space group $P12_11$, with unit-cell parameters $a = 44.05$, $b = 89.05$, $c = 60.74$ Å, $\beta = 100.5^\circ$. Initial structure determination and refinement gave an R factor and R_{free} of 17.3 and 22.0%, respectively, confirming a positive solution in obtaining phases using molecular replacement.

1. Introduction

Staphylococcus aureus is a Gram-positive bacterium which is both pathogenic and commensal to humans. Whilst it resides with no harm in ~20% of the general population, it is also the causative agent of many diseases ranging from minor to life-threatening, with the latter including meningitis, bacteraemia, pneumonia, food poisoning and toxic shock syndrome. It is also a notorious causative of nosocomial infections in hospitals around the world (Klevens *et al.*, 2008). *S. aureus*-mediated bacteraemia, particularly involving methicillin-resistant (MR) strains, has increased in recent years even in developed countries such as the United States, European countries, Australia and New Zealand (Collignon *et al.*, 2005; Shorr & Lodise, 2006; European Antimicrobial Resistance Surveillance System, 2006).

Prior to the introduction of antibiotics, the mortality rate owing to *S. aureus* infection was 90%. Even in the post-antibiotic era, fatality rates have reached 30%, which is mostly attributable to the emergence of methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *S. aureus* (VRSA) infections (World Health Organization, 2013). The increasing emergence of MRSA and VRSA demonstrates the urgent need for the development of new antimicrobial agents.

Understanding molecular details and cellular function through protein X-ray crystallography provides the potential for the development of new therapeutics. Here, we report the crystallization of PaaI-like thioesterase (*SaPaaI*), a protein that belongs to the amino-acid transport and metabolism gene cluster from *S. aureus* subsp. *aureus* Mu50. PaaI thioesterase catalyses the cleavage of CoA from phenylacetyl-CoA and hydroxyphenylacetyl-CoA substrates formed in the phenylacetate pathway. The functionality of the phenylacetate pathway of *Burkholderia cenocepacia* proved to be responsible for full pathogenicity when studied in a *Caenorhabditis elegans* host model (Law *et al.*, 2008). The general functions of acyl-CoA thioesterase involve the release of the carboxylate unit for degradation, regulation of cellular pathways and release of CoA for use in other reactions (Hunt & Alexson, 2002; Kirkby *et al.*, 2010). PaaI thio-

esterase is a member of the hotdog-fold thioesterase family, which includes 4-hydroxybenzoyl-CoA thioesterase, but is able to recognize phenylacetyl-CoA as substrate rather than benzoyl-CoA (Song *et al.*, 2006). The crystallization and preliminary diffraction analysis of SaPaaI will enable its high-resolution structure determination and will enable the characterization of its involvement in possible roles in amino-acid transport and metabolism, which in turn has the potential to assist in rational drug design against this pathogen.

2. Methods

2.1. Cloning

The 375 bp open reading frame (*sav0944*) encoding SaPaaI was amplified by polymerase chain reaction using *S. aureus* subsp. *aureus* Mu50 genomic DNA as a template (ATCC 700699D-5). The PCR-amplified product was then cloned into expression vector pMCSG21 using ligation-independent cloning. Briefly, this involved amplification of the *sav0944* gene with the forward and reverse primers **TACTTCCAATCCAATGCCATGACTCATTTATTAGAGACATTGAG** and **TTATCCACTTCCAATGTTATTATTTTAAAGGTTTAAATAGCAACTGTAC**, respectively, where the bold nucleotides are required for the ligation-independent cloning (LIC) procedure and the underlined nucleotides are complementary to the gene nucleotide sequence. Following the purification of amplified PCR product using a Qiagen PCR purification kit, the gene was cloned into *Ssp*I- and T4 polymerase-treated pMCSG21 expression vector. Confirmed clones by colony PCR were sent to the Australian Genome Research Facility (AGRF) for DNA sequencing to ensure the fidelity of the gene sequence. *BioEdit* and *ClustalW* were used to analyse the nucleotide sequences.

2.2. Expression and purification

The cloned SaPaaI allowed the recombinant expression of an N-terminal His₆-SaPaaI fusion protein in *Escherichia coli* BL21 (DE3) pLysS cells. A single colony of transformed *E. coli* BL21 (DE3) pLysS cells was grown to saturation in a 5 ml primary culture in Luria–Bertani (LB) broth overnight. Subsequently, 1 ml of this primary culture was used to inoculate 500 ml auto-induction medium (Studier, 2005) containing 100 µg ml⁻¹ spectinomycin, and the culture was incubated in a shaker incubator at 298 K for 24 h. Following centrifugation, the cell pellet was resuspended in 1/10 the volume of His buffer A (50 mM phosphate buffer pH 8.0, 300 mM NaCl, 20 mM imidazole) and stored at 253 K. Three freeze–thaw cycles and the addition of lysozyme (1 ml of a 20 mg ml⁻¹ stock) was used to lyse the cells, and centrifugation at 20 000g was used to remove cellular debris. The protein supernatant extract was loaded onto a 5 ml His column (HisTrap HP, GE Healthcare) using FPLC at room temperature, washed in His buffer A and eluted using a ten-

column-volume gradient to 100% His buffer B (50 mM phosphate buffer pH 8.0, 300 mM NaCl, 500 mM imidazole). *Tobacco etch virus* (TEV) protease was incubated for 12 h at 277 K with the protein (60 µg protease per milligram of protein) to remove the His₆ tag. Following TEV protease treatment, size-exclusion chromatography (HiLoad 26/60 Superdex 200 column, GE Healthcare) was performed with 10 ml TEV protease-treated unconcentrated protein to remove minor impurities and aggregates, and to exchange the buffer to 50 mM Tris pH 8.0, 125 mM NaCl. The protein was concentrated using Amicon Ultra-15 concentrators (Millipore) to 25 mg ml⁻¹ (as determined by UV spectroscopy at 280 nm with a predicted extinction coefficient of 6990 M⁻¹ cm⁻¹), aliquoted and stored at 193 K.

2.3. Crystallization

Initial crystallization trials were carried out using the hanging-drop vapour-diffusion method in 48-well plates with a 1:1 ratio (total drop volume of 3 µl) of protein solution with commercially available crystallization screens from Hampton Research. Plates were kept in a vibration-free incubator at 296 K and crystallization drops from plates were observed after 24 h intervals. Conditions that gave rise to crystals were optimized by iteratively varying the pH and the concentrations of precipitants and salts and the protein.

2.4. X-ray diffraction and structure solution

Large diffraction-quality crystals (800 × 100 × 100 µm) were cooled in cryoprotectant comprised of 25% glycerol in crystallization solution. Diffraction data were collected at 100 K at a wavelength of 0.95370 Å on the MX2 beamline at the Australian Synchrotron. The collected diffraction data were indexed and integrated using *iMosflm* (Battye *et al.*, 2011) and scaled in *AIMLESS* (Evans, 2006, 2011). Phases were determined by molecular replacement using *E. coli* YdiI (PDB entry 1sbk; Northeast Structural Genomics Consortium, unpublished work; 38% identity; Fig. 1) as a search model in *Phaser* (McCoy *et al.*, 2007) as distributed in the *CCP4* program suite v.6.3.0 (Winn *et al.*, 2011). Model building with *Coot* (v.0.7; Emsley *et al.*, 2010) and refinement with *REFMAC5* (Murshudov *et al.*, 2011) and *PHENIX* (Echols *et al.*, 2012) were performed to produce a final model with an *R* factor and *R*_{free} of 17.3 and 22.0%, respectively. Full structural and functional characterization including mutagenesis is currently being undertaken.

3. Results

3.1. Expression and purification

The overexpression of recombinant SaPaaI protein in *E. coli* BL21 (DE3) pLysS cells using auto-induction medium yielded approximately 30 mg of protein per litre of culture. The protein was highly soluble as determined by SDS–PAGE analysis of the whole cell

Figure 1

Amino-acid sequence alignment of SaPaaI with the search model used in molecular replacement using *ClustalX2*. The sequence identity between the two proteins is approximately 38%.

lysate, the soluble fraction and the purified protein from affinity chromatography (Fig. 2). Elution of the protein occurred at approximately 300 mM imidazole and resulted in 95% purity. Removal of the His₆ affinity tag by TEV protease treatment reduced the size of the protein by 3 kDa, consistent with the size of the tag, and left three N-terminal vector-encoded amino acids (SDA). Size-exclusion chromatography resulted in a single symmetrical peak, indicating a homogeneous protein preparation (Fig. 3), and greater than 99% purity as judged by SDS-PAGE analysis. The SaPaaI protein was concentrated to 25 mg ml⁻¹ and used for crystallization screening trials.

Table 1

Data-collection statistics.

Values in parentheses are for the outer resolution shell.

Source	MX2 beamline, Australian Synchrotron
Wavelength (Å)	0.9537
Data-collection temperature (K)	100
Crystal system	Monoclinic
Space group	<i>P</i> 12 ₁ 1
Unit-cell parameters (Å, °)	<i>a</i> = 44.05, <i>b</i> = 89.05, <i>c</i> = 60.74, <i>β</i> = 100.53
Resolution range (Å)	30.41–2.00 (2.072–2.000)
No. of measured reflections	125958 (8506)
No. of unique reflections	30911 (3089)
Completeness (%)	99.20 (99.39)
Multiplicity	3.5 (3.5)
Mean <i>I</i> /σ(<i>I</i>)	9.6 (2.4)
<i>R</i> _{merge} [†]	0.060 (0.407)

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual measurement of the reflection with Miller indices *hkl* and $\langle I(hkl) \rangle$ is the mean intensity of that reflection.

Figure 2

SDS-PAGE analysis (reducing gel). Lane A, Precision Plus Protein Standards (labelled in kDa). Lane B, whole cell lysate. Lane C, supernatant after centrifugation of the whole cell lysate. Lane D, flowthrough fraction after loading the supernatant onto the His column. Lane E, elution fraction after performing affinity chromatography (95% pure sample); the lower band represents the monomeric form of SaPaaI, whereas the upper band represents the dimeric form of SaPaaI. Lane F, elution fraction from gel-filtration chromatography (99% pure sample) following digestion with TEV protease to remove the His affinity tag.

3.2. Crystallization and diffraction

Initial crystallization conditions were screened through sparse-matrix screening using commercially available screens from Hampton Research. Ten conditions (Crystal Screen condition Nos. 4, 9, 14, 15, 18, 22, 33, 39, 40 and 46) gave rise to small crystals at 296 K after 24 h, all of which had the common precipitants of PEG 400, PEG 4000 or PEG 8000, with the exception of condition No. 4. Four conditions (Nos. 4, 9, 22 and 40) were optimized to produce diffraction-quality crystals with and without ligand (coenzyme A). Optimized condition No. 40 consisting of 24% PEG 4000, 100 mM sodium citrate pH 6.5, 12% 2-propanol produced large diffraction-quality crystals in the presence of 1 mM coenzyme A (CoA) ligand after 24 h at 296 K (Fig. 4). A single crystal was cooled in a cryoprotectant incorporating 25% glycerol in the crystallization solution and X-ray diffraction data were collected at the Australian Synchrotron. The crystal diffracted X-rays to beyond 2.0 Å resolution (Fig. 5). The crystal belonged to space group *P*12₁1, with unit-cell parameters *a* = 44.05, *b* = 89.05,

Figure 3

(a) Purification of SaPaaI by affinity chromatography. (b) Size-exclusion chromatography of affinity-purified SaPaaI to remove minor impurities, aggregates and the 3 kDa His₆ tag and to replace the buffer with 50 mM Tris, 125 mM NaCl pH 8.0. In both figures, the blue line corresponds to the UV absorbance, enabling monitoring of the protein; in (a) the green line corresponds to gradient elution and in (b) the brown line is conductivity, indicating elution of the salt peak from the affinity-chromatography step.

Figure 4

SaPaaI crystals obtained using 24% PEG 4000, 100 mM sodium citrate pH 6.5, 12% 2-propanol, 1 mM coenzyme A after 24 h at 296 K.

Figure 5

X-ray diffraction image of an SaPaaI protein crystal, showing diffraction beyond 2 Å resolution.

$c = 60.74 \text{ Å}$, $\beta = 100.53^\circ$, and the data frames were integrated and scaled to 2 Å resolution using *iMosflm* v1.0.5 and *AIMLESS*, respectively, with an $R_{p.i.m.}$ of 0.041 and 0.288 across all resolution shells and the outer resolution shell, respectively. The data-collection statistics are summarized in Table 1. The data did not appear to be twinned as determined using *POINTLESS* from the *CCP4* suite (Winn *et al.*, 2011), and a value of $2.18 \text{ Å}^3 \text{ Da}^{-1}$ for the Matthews

coefficient suggested the presence of four molecules in the asymmetric unit (Matthews, 1968). The structure was solved by molecular replacement in *Phaser* (McCoy *et al.*, 2007) using chain A (138 amino acids) of PDB entry 1sbk, with 38% amino-acid identity to SaPaaI, as a search model (LLG and TFZ of 557.4 and 14.5, respectively). Model building and refinement produced a model with an R factor and R_{free} of 17.3 and 22.0%, respectively, and the coordinates and structure factors have been validated and deposited in the Protein Data Bank as entry 4m20. Two CoA molecules and 192 water molecules are present. Six residues (Ser-2, Asp-1, Ala0, Met1, Leu123 and Lys124) are missing in chains A and D and five residues (Ser-2, Asp-1, Ala0, Met1 and Lys124) are missing in chains B and C of the final model. Full structural and functional characterization including activity and specificity assays, mutagenesis and crystallization with different ligands is currently being undertaken.

We thank the beamline scientists and acknowledge the use of the Australian Synchrotron. JKF and YBK are Australian Research Council (ARC) Future and COMPACT scholarship fellows, respectively. Funding has been provided by the Australian government and the School of Biomedical Sciences, Charles Sturt University.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Collignon, P., Nimmo, G. R., Gottlieb, T. & Gosbell, I. B. (2005). *Emerg. Infect. Dis.* **11**, 554–561.
- Echols, N., Grosse-Kunstleve, R. W., Afonine, P. V., Bunkóczi, G., Chen, V. B., Headd, J. J., McCoy, A. J., Moriarty, N. W., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Adams, P. D. (2012). *J. Appl. Cryst.* **45**, 581–586.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- European Antimicrobial Resistance Surveillance System (2006). *EARSS Annual Report 2006*. http://www.ecdc.europa.eu/en/activities/surveillance/ears-net/documents/2006_earss_annual_report.pdf. Bilthoven: EARSS.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
- Hunt, M. C. & Alexson, S. E. H. (2002). *Prog. Lipid Res.* **41**, 99–130.
- Kirkby, B., Roman, N., Kobe, B., Kellie, S. & Forwood, J. K. (2010). *Prog. Lipid Res.* **49**, 366–377.
- Kleves, R. M., Edwards, J. R. & Gaynes, R. P. (2008). *Clin. Infect. Dis.* **47**, 927–930.
- Law, R. J., Hamlin, J. N., Sivo, A., McCorrister, S. J., Cardama, G. A. & Cardona, S. T. (2008). *J. Bacteriol.* **190**, 7209–7218.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Shorr, A. F. & Lodise, T. (2006). *ISMR Update*, **1**, 4–11.
- Song, F., Zhuang, Z., Finci, L., Dunaway-Mariano, D., Kniewel, R., Buglino, J. A., Solorzano, V., Wu, J. & Lima, C. D. (2006). *J. Biol. Chem.* **281**, 11028–11038.
- Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- World Health Organization (2013). *Initiative for Vaccine Research: Staphylococcal Infection*. http://www.who.int/vaccine_research/diseases/soa_bacterial/en/index2.html. Geneva: World Health Organization.