

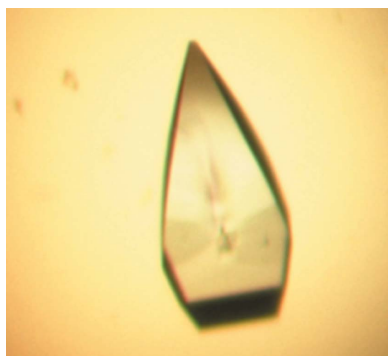
Abimael Cruz-Migoni,^a
Sergey N. Ruzheinikov,^a
Svetlana E. Sedelnikova,^a
Barbara Obeng,^a Sylvia Chieng,^b
Rahmah Mohamed,^{b,c} Sheila
Nathan,^{b,c} Patrick J. Baker^a and
David W. Rice^{a*}

^aKrebs Institute for Biomolecular Research,
Department of Molecular Biology and
Biotechnology, The University of Sheffield,
Sheffield S10 2TN, England, ^bSchool of
Biosciences and Biotechnology, Faculty of
Science and Technology, Universiti Kebangsaan
Malaysia, 43600 Bangi, Selangor DE, Malaysia,
and ^cMalaysia Genome Institute, UKM-MTDC
Technology Centre, 43600 Bangi, Selangor DE,
Malaysia

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

Received 20 June 2011

Accepted 18 July 2011



© 2011 International Union of Crystallography
All rights reserved

Cloning, purification and crystallographic analysis of a hypothetical protein, BPSL1549, from *Burkholderia pseudomallei*

Burkholderia pseudomallei BPSL1549, a putative protein of unknown function, has been overexpressed in *Escherichia coli*, purified and subsequently crystallized by the hanging-drop vapour-diffusion method using PEG as a precipitant to give crystals with overall dimensions of $0.15 \times 0.15 \times 0.1$ mm. Native data were collected to 1.47 Å resolution at the European Synchrotron Radiation Facility (ESRF). The crystals belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 37.1$, $b = 45.4$, $c = 111.9$ Å and with a single polypeptide chain in the asymmetric unit.

1. Introduction

Burkholderia pseudomallei is a Gram-negative, facultative, intracellular pathogen of both animals and humans that grows in hot wet soil and stagnant water (Dance, 1991; Galyov *et al.*, 2010). This bacterium is the causative agent of melioidosis (Brett & Woods, 2000), a disease endemic to tropical and subtropical areas, particularly Southeast Asia and Northern Australia (Gan, 2005). Infection by *B. pseudomallei* occurs by inhalation, inoculation or ingestion of contaminated soil or water, with the size of the inoculum influencing the pattern and the severity of the disease (Cheng & Currie, 2005). Commonly, melioidosis has a complex spectrum of clinical manifestations, including abscesses, pneumonia and, at worst, fatal septicaemia. The formation of abscesses can occur in any organ, although in patients with chronic melioidosis the lungs are commonly affected (Dhensiri *et al.*, 1988; Currie *et al.*, 2000). As a result of these symptoms mimicking those of other common diseases, such as malaria, typhoid fever, leptospirosis and tuberculosis (Leelarasamee & Bovornkitti, 1989), melioidosis is often misdiagnosed, making it difficult to treat in a timely fashion (Aldhous, 2005). There is currently no protective vaccine against melioidosis (Dance *et al.*, 1989; Bossi *et al.*, 2004) and treatment requires intensive antibiotic therapy for many months. Moreover, many patients have a high risk of relapse after a dormant state of many decades (Chaowagul *et al.*, 1993; Ngaay *et al.*, 2005) and the emergence of drug-resistant strains is reducing the efficacy of conventional therapies such as ceftazidime or imipenem (Simpson *et al.*, 1999).

The importance of *B. pseudomallei* as a disease-causing organism has now resulted in the completion of the sequencing of the genomes of a range of strains, including the most well characterized, K96243 (Holden *et al.*, 2004). Genomic analysis has shown that *B. pseudomallei* contains two chromosomes of size 5.23 and 9.73 Mb with a high G + C content. The large chromosome contains a higher proportion of coding sequences involved in core functions associated with central metabolism and cell growth, whereas the small chromosome carries genes encoding accessory functions associated with adaptation and survival in different environments (Holden *et al.*, 2004). Whilst some virulence factors involved in the pathogenesis of *B. pseudomallei* have been proposed (Adler *et al.*, 2009; Galyov *et al.*, 2010), the molecular basis of the pathogenicity of the bacterium is poorly understood (Stone, 2007). Proteomic studies have led to the identification of a number of proteins whose levels differ between the pathogenic *Burkholderia* strain *B. pseudomallei* and the nonpathogenic strain *B. thailandensis*. These proteins therefore constitute a list

of possible pathogenicity factors (Wongtrakongate *et al.*, 2007). One of the proteins identified in this study was the product of the *bpsl1549* gene, a 211-residue protein of unknown function whose sequence shows no similarities to any protein of known three-dimensional structure and which does not appear to have any homologues outside *Burkholderia*. In this paper, we report the cloning, overexpression, purification and initial X-ray analysis of BPSL1549.

2. Materials and methods

2.1. Cloning and overexpression

The coding sequence of the putative full-length BPSL1549 gene was PCR-amplified directly from genomic DNA of *B. pseudomallei* strain D286, a pathogenic strain isolated from a patient with melioidosis at Kuala Lumpur General Hospital (Lee *et al.*, 2007), using the primers ATGCCCAACTCACTCGAAGC (forward) and CTATTGCTTGCGCTGCTGGA (reverse). The sequence of the encoded protein is identical in this strain and the more well characterized strain K96243. The purified DNA fragment (636 bp) was inserted into a pETBlue-1 vector using an AccepTor Vector Kit (Novagen). The positive clones were confirmed by blue/white selection and colony PCR and the plasmids were isolated and purified. The plasmids were sequenced and checked against the sequence from strain K96243 in the NCBI database and transformed into *Escherichia coli* Tuner (DE3) cells (Novagen) for overexpression. In order to produce wild-type BPSL1549 protein, a 250 ml flask containing 50 ml LB medium with 50 µg ml⁻¹ carbenicillin and chloramphenicol was inoculated with a single colony of the transformed strain and grown overnight at 310 K on a shaking tray at 250 rev min⁻¹. 8 ml of this culture was used to inoculate 2 l flasks each containing 450 ml LB medium supplemented with carbenicillin and chloramphenicol as above. Growth was carried out at 310 K with vigorous aeration until an OD₆₀₀ of 0.6 was attained, at which point overexpression was induced by the addition of 1 mM IPTG and the culture was grown for an additional 4 h. The cells were harvested by centrifugation at 5000 rev min⁻¹ for 25 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 (23 kDa).

To produce selenomethionine-containing protein, the transformed *E. coli* Tuner cells were grown in LB medium until an OD₆₀₀ of 0.6 was reached. The cells were then harvested by centrifugation at 5000 rev min⁻¹ for 20 min, the medium was decanted and the cells were resuspended in minimal medium containing 10.5 g l⁻¹ K₂HPO₄, 1 g l⁻¹ ammonium sulfate, 4.5 g l⁻¹ KH₂PO₄, 0.5 g l⁻¹ trisodium citrate, 5 g l⁻¹ glycerol, 0.5 g l⁻¹ adenine, guanosine, thymine and uracil, 1 mg l⁻¹ MgSO₄, 4 mg l⁻¹ thiamine, 40 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. The cultures were incubated with shaking for 10 min; IPTG was then added to a final concentration of 1 mM and the cultures were grown for an additional 4 h.

2.2. Purification

For purification of either the native or SeMet protein, 3 g of cells were disrupted by sonication in 50 mM Tris-HCl pH 8.0. The cell debris was removed by centrifugation at 70 000 rev min⁻¹ for 10 min. The supernatant was collected and loaded onto a DEAE-Sepharose Fast Flow column (GE Healthcare) and proteins were eluted with a linear gradient of 0–0.5 M NaCl in 50 mM Tris-HCl pH 8.0. Elution of BPSL1549 occurred at approximately 80 mM NaCl and fractions containing the highest concentrations of the protein were combined and concentrated using a Vivaspin 20 concentrator. The concentrated samples were then subjected to gel filtration using a Hi-Load Superdex 200 column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 8.0, 0.5 M NaCl and the proteins were eluted with the same buffer. Wild-type BPSL1549 runs on this column with an apparent molecular weight of 20 kDa, suggesting that the protein is a monomer in solution. Peak fractions containing BPSL1549 protein were concentrated to approximately 23 mg ml⁻¹ in a Vivaspin concentrator filter with a 10 kDa molecular-weight cutoff; the buffer was exchanged to 10 mM Tris-HCl pH 8.0 using a diafiltration cup. The final purity was estimated to be greater than 95% as determined by SDS-PAGE (Fig. 1). Approximately 20 mg pure protein was obtained from 1 l culture.

2.3. Crystallization and preliminary X-ray analysis of BPSL1549

Initial crystallization trials were carried out using NeXtal crystallization kits (Qiagen, Germany) on a Matrix Hydra II (Thermo Fisher Scientific, USA) using the standard sitting-drop vapour-diffusion

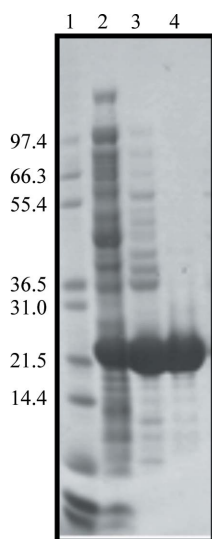


Figure 1
SDS-PAGE analysis of the BPSL1549 purification. Fractions from each purification step were analyzed by SDS-PAGE and Coomassie Brilliant Blue staining. Lane 1, molecular-weight protein markers (Invitrogen; labelled in kDa); lane 2, cell-free extract; lane 3, sample after DEAE Sepharose; lane 4, sample after gel filtration on Superdex 200.

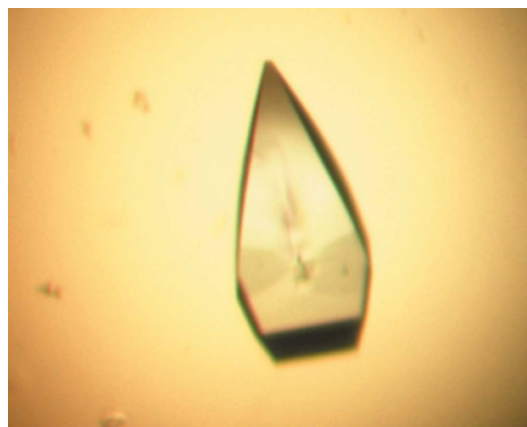
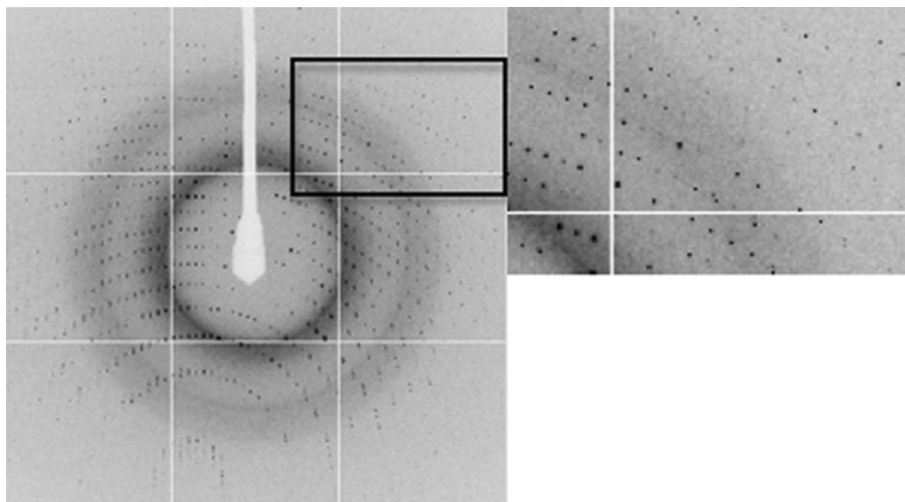


Figure 2
Native crystals of BPSL1549 grown using 0.2 M sodium bromide, 0.1 M bis-tris propane pH 6.5 and 26% (w/v) PEG 3350.


Figure 3

A representative 0.95° oscillation image of data collected from a native BPSL1549 crystal using an ADSC Q315R detector on beamline ID29 at the European Synchrotron Radiation Facility (ESRF). An enlarged view of the region indicated by the square is shown on the right. The diffraction extends to the edge of the plate (1.47 Å).

Table 1

X-ray data-collection statistics for crystals of BPSL1549.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 37.1, b = 45.4, c = 111.9$
Temperature (K)	100
X-ray source	ID29, ESRF
Data-collection wavelength (Å)	0.82
Detector	ADSC Q315R
Resolution range (Å)	25–1.47 (1.56–1.47)
Total No. of observations	155353 (22707)
Unique reflections	32232 (4691)
$R_{\text{merge}}^{\dagger}$	0.066 (0.135)
Completeness (%)	99.0 (99.7)
Multiplicity	4.8 (4.8)
Mean $I/\sigma(I)$	17.3 (9.4)

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

technique by adding 0.2 μl BPSL1549 protein at 23 mg ml^{-1} in 10 mM Tris–HCl pH 8.0 to an equal volume of the precipitant and equilibrating against a 100 μl reservoir of the same precipitant at 290 K. Initial crystals were observed using 0.2 M sodium bromide, 0.1 M bis-tris propane pH 6.5 and 26% (w/v) PEG 3350 as precipitant. Optimization of these conditions using the hanging-drop vapour-diffusion technique led to larger crystals of overall dimensions 0.15 \times 0.15 \times 0.1 mm (Fig. 2). Crystals of the SeMet protein with similar dimensions were obtained under equivalent conditions. For data collection, the crystals were flash-cooled to 100 K in a cryoprotectant solution consisting of 0.2 M sodium bromide, 0.1 M bis-tris propane pH 6.5 and 20% glycerol. Data were collected to 1.47 Å resolution from crystals of the native protein using an ADSC Q315R detector on beamline ID29 at the European Synchrotron Radiation Facility (ESRF) as a total of 125 images using 0.95° oscillation per image (Fig. 3).

3. Results and discussion

Analysis of the diffraction data using the autoindexing routine in *MOSFLM* (Leslie, 1992) showed that the crystals belonged to an orthorhombic space group, with unit-cell parameters $a = 37.1, b = 45.4, c = 111.9$ Å; one subunit in the asymmetric unit gave a calculated

Matthews coefficient of 2.08 Å³ Da⁻¹ with a solvent content of 41% (Matthews, 1977) for a monomer in the asymmetric unit. The data were processed using *MOSFLM* (Leslie, 1992) and *SCALA* (Winn *et al.*, 2011) and analysis of the pattern of systematic absences was consistent with the crystal belonging to space group $P2_12_12_1$. Data-collection and processing statistics are shown in Table 1. Attempts to solve the structure using MAD techniques are currently under way in the hope that the structure of BPSL1549 will contribute to a better understanding of its function and its role in the pathogenesis of *B. pseudomallei*.

This work was supported by a research grant (07-05-16-MGI-GMB08) from the Ministry of Science, Technology and Innovation, Government of Malaysia and by the PMI-2 initiative of the British Council. AC-M thanks the National Council for Science and Technology of Mexico (CONACyT-Mexico) for scholarship funding. We acknowledge the European Synchrotron Radiation Facility (ESRF) for provision of beam time and thank Christoph Müller-Dieckmann for assistance using station ID29.

References

- Aldhous, P. (2005). *Nature (London)*, **434**, 692–693.
- Bossi, P., Tegnell, A., Baka, A., Van Looock, F., Hendriks, J., Werner, A., Maidhof, H. & Gouvras, G. (2004). *Euro Surveill.* **9**, E3–E4.
- Brett, P. J. & Woods, D. E. (2000). *Acta Trop.* **74**, 201–210.
- Chaowagul, W., Suputtamongkol, Y., Dance, D. A., Rajchanuvong, A., Pattara-arechachai, J. & White, N. J. (1993). *J. Infect. Dis.* **168**, 1181–1185.
- Cheng, A. C. & Currie, B. J. (2005). *Clin. Microbiol. Rev.* **18**, 383–416.
- Currie, B. J., Fisher, D. A., Howard, D. M., Burrow, J. N., Lo, D., Selva-Nayagam, S., Anstey, N. M., Huffam, S. E., Snelling, P. L., Marks, P. J., Stephens, D. P., Lum, G. D., Jacups, S. P. & Krause, V. L. (2000). *Clin. Infect. Dis.* **31**, 981–986.
- Dance, D. A. (1991). *Clin. Microbiol. Rev.* **4**, 52–60.
- Dance, D. A., Wuthiekanun, V., Chaowagul, W. & White, N. J. (1989). *J. Antimicrob. Chemother.* **24**, 295–309.
- Dhiansiri, T., Puapairoj, S. & Sasaengrat, W. (1988). *Radiology*, **166**, 711–715.
- Galyov, E. E., Brett, P. J. & DeShazer, D. (2010). *Annu. Rev. Microbiol.* **64**, 495–517.
- Gan, Y.-H. (2005). *J. Infect. Dis.* **192**, 1845–1850.
- Holden, M. T. *et al.* (2004). *Proc. Natl Acad. Sci. USA*, **101**, 14240–14245.
- Lazar Adler, N. R., Govan, B., Cullinane, M., Harper, M., Adler, B. & Boyce, J. D. (2009). *FEMS Microbiol. Rev.* **33**, 1079–1099.

- Lee, S.-H., Chong, C.-E., Lim, B.-S., Chai, S.-J., Sam, K.-K., Mohamed, R. & Nathan, S. (2007). *Diagn. Microbiol. Infect. Dis.* **58**, 263–270.
- Leelarasamee, A. & Bovornkitti, S. (1989). *Rev. Infect. Dis.* **11**, 413–425.
- Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- Matthews, B. W. (1977). *The Proteins*, 3rd ed., edited by H. Neurath & R. L. Hill, Vol. 3, pp. 468–477. New York: Academic Press.
- Ngauy, V., Lemeshev, Y., Sadkowski, L. & Crawford, G. (2005). *J. Clin. Microbiol.* **43**, 970–972.
- Simpson, A. J., Suputtamongkol, Y., Smith, M. D., Angus, B. J., Rajanuwong, A., Wuthiekanun, V., Howe, P. A., Walsh, A. L., Chaowagul, W. & White, N. J. (1999). *Clin. Infect. Dis.* **29**, 381–387.
- Stone, R. (2007). *Science*, **317**, 1022–1024.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Wongtrakongate, P., Mongkoldhumrongkul, N., Chaijan, S., Kamchonwongpaisan, S. & Tungpradabkul, S. (2007). *Mol. Cell. Probes*, **21**, 81–91.