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

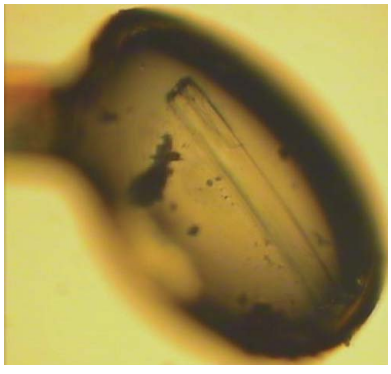
Expression, purification, crystallization, data collection and preliminary biochemical characterization of methicillin-resistant *Staphylococcus aureus* Sar2028, an aspartate/tyrosine/phenylalanine pyridoxal-5′-phosphate-dependent aminotransferase

Jaldappagari Seetharamappa,[a,b] Muse Oke,[a] Huanting Liu,[a] Stephen A. McMahon,[a] Kenneth A. Johnson,[a] Lester Carter,[a] Mark Dorward,[a] Michal Zawadzki,[a] Ian M. Overton,[c] C. A. Johannes van Niekirk,[c] Shirley Graham,[a] Catherine H. Botting,[a] Garry L. Taylor,[a] Malcolm F. White,[a] Geoffrey J. Barton,[c] Peter J. Coote[a] and James H. Naismith[a]*

[a]Scottish Structural Facility and Centre for Biomolecular Sciences, The University, St Andrews KY16 9ST, Scotland, [b]Department of Chemistry, Karnatak University, Pavate Nagar, Dharwad 580 003, Karnataka State, India, and [c]Scottish Structural Facility and School of Life Sciences Research, University of Dundee, Dow Street, Dundee DD1 5EH, Scotland

Correspondence e-mail: naismith@st-andrews.ac.uk

Received 22 March 2007
Accepted 19 April 2007



Sar2028, an aspartate/tyrosine/phenylalanine pyridoxal-5′-phosphate-dependent aminotransferase with a molecular weight of 48 168 Da, was overexpressed in methicillin-resistant *Staphylococcus aureus* compared with a methicillin-sensitive strain. The protein was expressed in *Escherichia coli*, purified and crystallized. The protein crystallized in a primitive orthorhombic Laue group with unit-cell parameters $a = 83.6$, $b = 91.3$, $c = 106.0$ Å, $\alpha = \beta = \gamma = 90^\circ$. Analysis of the systematic absences along the three principal axes indicated the space group to be $P2_12_12_1$. A complete data set was collected to 2.5 Å resolution.

1. Introduction

In recent years, the emergence of antimicrobial resistance in previously susceptible common pathogenic bacteria has been a major concern in public health. *Staphylococcus aureus* is amongst the most significant of such pathogens; it is found widely in hospitals and, if untreated, can lead to death through sepsis. Methicillin-resistant *S. aureus* (MRSA) demonstrates resistance not only to methicillin but also to other penicillin-type antibiotics (Foster, 1996). MRSA is now widespread in UK hospitals (Johnson *et al.*, 2001) and is termed a 'superbug' by the popular press. In addition to its presence in hospitals, community-acquired MRSA strains are now being reported around the world (Holden *et al.*, 2004; Pesavento *et al.*, 2007; Boyle-Vavra & Daum, 2007). The increase in the rate of methicillin resistance was slow and gradual in the UK during the 1980s. The subsequent surge in infections has been attributed in part to community-based infections taking hold (Herold *et al.*, 1998). Many such community cases occur in otherwise healthy individuals with few risk factors for MRSA (Herold *et al.*, 1998) and the infections are often centred on wounds. The resistance of MRSA to β-lactams arises from a number of changes: impaired β-lactam entry into bacteria, instability of β-lactams to bacterial serine- or metallo-β-lactamases or the loss of β-lactam binding affinity to penicillin-binding protein (Edwards *et al.*, 2004; Bush *et al.*, 1995). Finding novel targets to treat staphylococcal infections, particularly MRSA, is a pressing problem. We have used two-dimensional gel proteomics to identify expression differences between MRSA strain 252 and the methicillin-sensitive *S. aureus* strain 476 (MSSA). We have identified that the expression of Sar2028 is upregulated in MRSA252 compared with MSSA476. MRSA252 Sar2028 is an acidic protein (MW 48 168 Da; pI 5.25; Holden *et al.*, 2004) previously annotated as a hypothetical protein belonging to the pyridoxal 5′-phosphate (PLP) dependent aminotransferase superfamily.

Aminotransferases are ubiquitous metabolic enzymes that catalyze the transfer of an amino group from an amino acid to either a keto acid or an aldehyde. PLP, an active form of vitamin B_6, is the essential cofactor for catalysis. During the aminotransferase reaction, PLP serves as a transitory acceptor of the amino group from the donor substrate and is converted to pyridoxamine 5′-phosphate (PMP). PMP then serves as a donor of the amino group to the acceptor

substrate, with accompanying regeneration of PLP (Hirotsu *et al.*, 2005). Most PLP-dependent enzymes contain a conserved lysine residue which forms a Schiff base with the coenzyme in the resting state; the species is referred to as the internal aldimine. Sar2028 has relatively low sequence homology (<25%) to known PLP enzymes and its function is not known. At a 60% identity threshold, homologues of the protein can be identified in several other organisms, including *Bacillus* spp. As part of our program on MRSA, we have cloned, expressed, purified and crystallized this protein.

2. Materials and methods

2.1. Proteomics

Cells were prepared as described by Oku *et al.* (2004) with minor modifications. Briefly, 5 l of late-exponential cells were grown in TSB media (Oxoid, UK) at 310 K and harvested; pellets were then washed twice with ice-cold H_2O. Pellets were resuspended in 20 m*M* Tris–HCl pH 6.8 and incubated with 200 µg ml^{-1} lysostaphin (Sigma) at 277 K for 30 min. Cells were then sonicated for 2 × 1 min on ice, with 1 min on ice in between, prior to centrifugation at 15 000*g* for 15 min at 277 K. The supernatant was retained and centrifuged at 100 000*g* for 60 min at 277 K. The pellet was then resuspended in 20 m*M* Tris–HCl and 1 m*M* 2-mercaptoethanol before centrifugation at 100 000*g* for 60 min at 277 K and resuspension of the pellet in lysis buffer (Lawrence *et al.*, 2004).

Comparative two-dimensional electrophoresis was performed using a MultiphorII apparatus for isoelectric focusing (IEF; GE Healthcare, Sweden) and a DALT-6 apparatus (GE Healthcare, Sweden) for the second dimension as described by Lawrence *et al.* (2004). Protein samples were run on preparative gels (0.1 mg total protein) and stained with silver stain as described by the manufacturer (GE Healthcare). Protein spots that showed reproducible changes in protein abundance (present on both gel sets from duplicate experiments) were catalogued. Gels were scanned using an Image Scanner (GE Healthcare). Protein spots with altered abundance were excised from these gels and identified by LC-MS/MS. Images were exported into Microsoft *PhotoDraw* 2000 v.2.0 for annotation and presentation.

Each gel spot was cut into 1 mm cubes and subjected to in-gel digestion using standard protocols (Gharahdaghi *et al.*, 1999; Shevchenko *et al.*, 1996). The peptides were extracted with 10% formic acid and concentrated to 20 µl (SpeedVac, ThemoSavant). They were then separated using an UltiMate nanoLC (LC Packings, Amsterdam) equipped with a PepMap C18 trap and column. The eluent was sprayed into a Q-Star Pulsar XL tandem mass spectrometer (Applied Biosystems, Foster City, CA, USA) and analysed in information-dependent acquisition (IDA) mode. MS/MS data for doubly and triply charged precursor ions were converted to centroid data, without smoothing, using the *Mascot Daemon* 2.1 (Matrix Science, London) data-import filter for *Sciex Analyst*. The MS/MS datafile generated was analysed using the *Mascot* search engine against the MSDB. The data were searched with tolerances of 0.2 Da for the precursor and fragment ions, trypsin as the cleavage enzyme, one missed cleavage, carbamidomethyl modification of cysteines as a fixed modification and methionine oxidation selected as a variable modification. A side-by-side image of two-dimensional gels with the spot identified as Sar2028 is shown in Fig. 1.

2.2. Cloning, expression and purification

The Sar2028 gene was amplified using MRSA252 genomic DNA as a template and cloned into pDEST14 as an N-terminal TEV-protease cleavable His-tag fusion using a modified Gateway recombination system (Liu & Naismith, unpublished work). The system adds a hexa-His tag and a TEV cleavage site at the N-terminus of the protein. *Escherichia coli* BL21 (DE3) (Novagen) was transformed with the recombinant plasmid. Recombinant cells were cultivated in 4 l Luria broth supplemented with ampicillin (final concentration 100 µg ml^{-1}) at 310 K until A_{600} = 0.6, at which point protein expression was induced by addition of 0.4 m*M* isopropyl β-D-thiogalactopyranoside (IPTG). After a further 6 h incubation period, cells were harvested by centrifugation at 2500*g* and 277 K for 30 min. Harvested cells were washed in PBS and centrifuged as before; cell pellets were stored at 193 K until required.

For purification, cell pellets with overexpressed Sar2028 were resuspended in lysis buffer (50 m*M* sodium phosphate pH 7.5, 500 m*M* NaCl, 1 mg ml^{-1} lysozyme) with appropriate protease inhibitors and lysed on ice using a Constant Systems cell disrupter at 207 MPa. The crude lysate was centrifuged (15 000*g*, 30 min, 277 K) and the cleared lysate was filtered through a 0.22 µm filter. Protein was purified from cleared lysate by two-step nickel-affinity chromatography. The lysate was batch bound to nickel Sepharose 6 fast flow medium (GE Healthcare), poured into a column and washed with 50

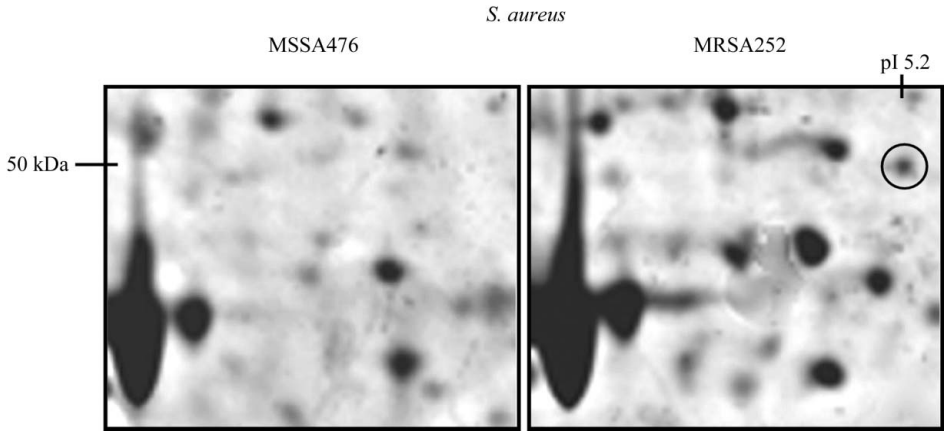


Figure 1
Crops from two-dimensional PAGE gels showing differences in protein expression between *S. aureus* MSSA476 and MRSA252. Crude membrane preparations were separated by IEF over the pH range 4–7 and proteins of interest were identified by LC-MS/MS. The circled protein was identified as Sar2028 (SwissProt Q6GFC0). The experiment was performed in duplicate and a representative result is shown.

column volumes of lysis buffer plus 20 m*M* imidazole. Sar2028 was eluted in ten column volumes of lysis buffer plus 500 m*M* imidazole and immediately desalted into 50 m*M* Tris, 500 m*M* NaCl pH 7.5 on a HiPrep 26/10 desalting column (GE Healthcare) to remove imidazole. His-tagged TEV protease was added to the protein at a 1:10 mass ratio to remove the N-terminal His tag from Sar2028, leaving the native protein. The TEV/target protein mixture was incubated at room temperature for 15 h. Cleaved Sar2028 was separated from His-tagged TEV by difference purification on Ni resin and polished by S-75 gel filtration. Gel filtration is consistent with a dimer, but is not definitive. The purified protein was characterized by SDS–PAGE and mass spectrometry before being concentrated to 10 mg ml^{-1} for crystallization.

2.3. Crystallization

Crystallization trials for purified Sar2028 in the native form were carried out using a sitting-drop vapour-diffusion screen of commercial sparse-matrix crystallization conditions. Protein solutions at concentrations of 10 mg ml^{-1} with a drop size of 0.2 or 0.3 µl (containing 0.1 or 0.2 µl protein solution and 0.1 µl precipitant solution) and 20 mg ml^{-1} with a drop size of 0.2 µl (containing 0.1 µl protein solution and 0.1 µl precipitant solution) were screened at 293 K using a nanodrop crystallization robot (Cartesian HoneyBee) as a part of the Hamilton–Thermo Rhombix system. The crystal screens employed in the study include JMac, Nextal The Classics, Nextal pH Clear, EBS Wizard, Nextal JCSG and Nextal PEGs. Optimizations of the 12 initially most promising hits were performed. The crystals were confirmed as protein by manual means using the 'click test'. Salt crystals tend to 'click' when broken, while protein crystals tend to be crushed rather than break. Crystallization conditions were varied using a stochastic matrix of 24 conditions with the following stock solutions of different precipitants [40%(*w*/*v*) PEG 8000, 4.5–9.8%; 45%(*w*/*v*) PEG 4000, 18.4–27.0%; 50%(*w*/*v*) PEG 3350, 15–27.5%; 1.0 *M* HEPES buffer pH 7.0, 7.5, 8.0], salt solutions (2.0 *M* $MgCl_2$, 0.02–0.26 *M*; 2.0 *M* Li_2SO_4, 0.23–0.26 *M*) and additives [50%(*w*/*v*) glycerol, 0.6–1.5%; 50%(*w*/*v*) ethylene glycol, 0.7–1.4%] at a protein concentration of 10 mg ml^{-1}. The hanging-drop vapour-diffusion method (1 + 0.5 µl mixture of protein and crystallization solution equilibrated against 450 µl of the latter in 24-well plate) was used for these experiments.

We also systematically varied conditions around 25%(*w*/*v*) PEG 4000 and 1.0 *M* HEPES pH 8.0 with a protein concentration of 10.0 mg ml^{-1} and a final drop concentration of 1 m*M* PLP, which was a suspected cofactor. Adding this cofactor prior to crystallization gave the best crystals, as judged by visual quality. We attribute this to an absence or low occupancy of the cofactor in the as-purified protein. Crystals grown with cofactor were faintly yellow in colour and appeared with a rod-like morphology from precipitate. The rods grew after around 7 d, with well defined edges appearing after a further 3 d incubation period. The optimal condition employed a precipitant of 23.6%(*w*/*v*) PEG 4000, 0.1 *M* HEPES pH 8.0 mixed with 10 mg ml^{-1} protein pre-incubated with PLP at 1.0 m*M* immediately prior to plate set-up. Crystals from hanging-drop experiments (as described above) grew to approximately 0.2 × 0.2 × 0.6 mm (Fig. 2).

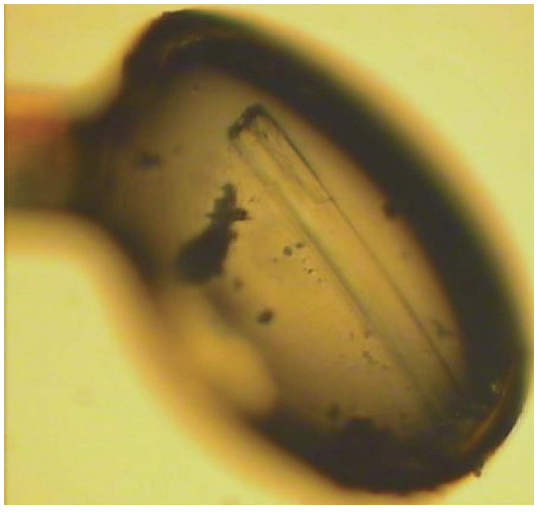


Figure 2
Crystal of Sar2028 mounted in a loop prior to data collection. The crystal dimensions are approximately 0.2 × 0.2 × 0.6 mm.

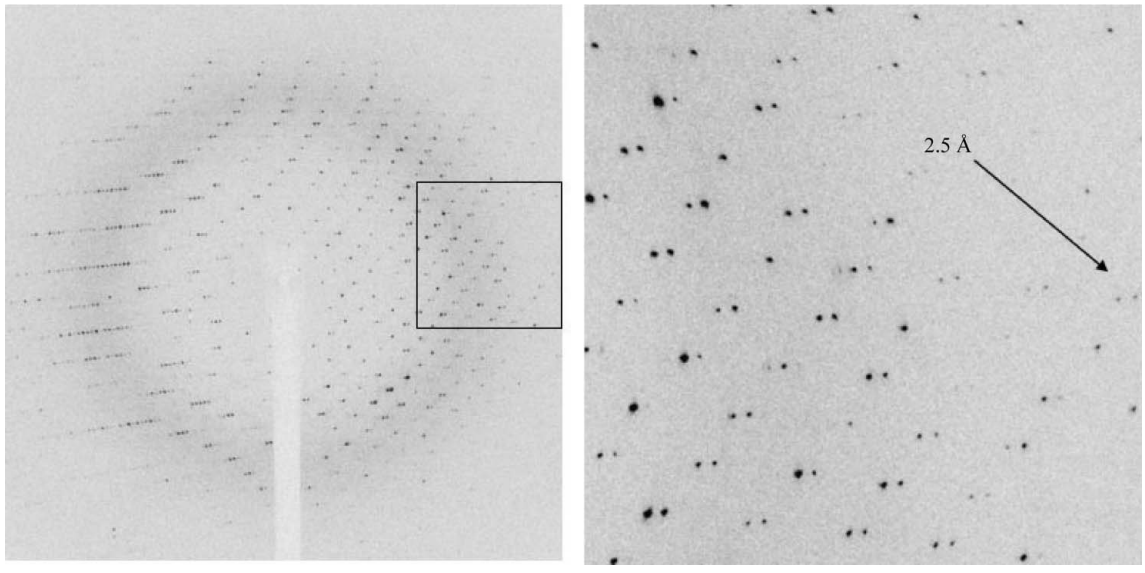


Figure 3
Diffraction pattern of Sar2028 collected at the ESRF. An enlarged portion is shown on the right, with the resolution of the marked spot.

Table 1
Crystal data and data-collection statistics.

Values in parentheses refer to the highest resolution shell.

Wavelength (Å)	0.873
Resolution (Å)	42–2.5 (2.64–2.5)
Space group	$P2_12_12_1$
Temperature (K)	100
Detector	MAR 225 CCD
Unit-cell parameters (Å)	a = 83.6, b = 91.3, c = 106.0
V_M (Å^3 Da^{-1}) for dimer	2.1
Solvent content for dimer (%)	41.5
Unique reflections	28714
$I/\sigma(I)$	25.9 (4.8)
Average redundancy	5.4 (6.1)
Data completeness (%)	99.4 (100)
R_{merge}†	0.058 (0.353)

† $R_{merge} = \sum_i[\sum_j(I_j - \langle I\rangle)]/\sum_i(\sum_j\langle I\rangle)$, where I_j is the intensity of the jth observation of reflection i, $\langle I\rangle$ is the the mean of the intensities of all observations of reflection I, $\sum_i$ is taken over all reflections and $\sum_j$ is taken over all observations of each reflection.

2.4. X-ray data collection

Prior to exposure to X-rays, Sar2028 crystals were soaked for 5–10 s in a cryoprotectant solution [consisting of 0.1 *M* HEPES pH 8, 25%(*w*/*v*) PEG 4000, 20%(*w*/*v*) PEG 400 and 1.0 m*M* PLP]. The soaked crystal was removed from the solution, mounted on a cryo-loop and placed into a stream of nitrogen at 100 K. Diffraction data were collected at the European Scientific Research Facility (ESRF), Grenoble on beamline ID 23-2. The incident X-ray beam had a wavelength of 0.873 Å. Data were recorded on a MAR 225 CCD detector with a crystal-to-detector distance of 252 mm. A typical image is shown in Fig. 3. A total of 102 non-overlapping images were collected using an oscillation angle of 1.5° per frame. All data were integrated using *MOSFLM* (Leslie, 1992) and merged with *SCALA* (Evans, 1997) as implemented in *CCP*4 (Collaborative Computational Project, Number 4, 1994). The initial indexing revealed the crystals to be primitive orthorhombic, with unit-cell parameters a = 83.6, b = 91.3, c = 106.0 Å, $\alpha = \beta = \gamma = 90°$. The program *POINTLESS*, implemented within *CCP*4 (Collaborative Computational Project, Number 4, 1994), was used to confirm Laue group as *Pmmm* and to suggest a space group. Measurement of the intensities of the axial reflections (over 20 for each of three principal axes) is consistent with a twofold screw axis along each, identifying the space group as $P2_12_12_1$. The data-collection statistics are shown in Table 1.

3. Results and discussion

Sar2028 was identified using the *Mascot* search engine (Matrix Science; http://www.matrixscience.com) with three matched peptides with a MOWSE score of 77. The probability of such a match being produced randomly is less than 0.5%. The presence of the Sar2028 in the lysate confirms that it is expressed by MRSA252. Interestingly, the protein was not expressed in MSSA476 under the conditions used in our proteomics study; the reason for this remains to be determined.

Calculation of the Matthews coefficient (V_M = 2.1 Å^3 Da^{-1}) suggested the presence of two molecules of Sar2028 in the asymmetric unit and a solvent content of 41.5%. A monomer and trimer in the asymmetric unit gave Matthews coefficients of 4.2 and 1.4 Å^3 Da^{-1}, with solvent contents of 70.7 and 12.2%, respectively. Analysis of the self-rotation function shows an additional peak on the $\kappa = 180°$ section with a height of 31% of the origin, suggesting a noncrystallographic twofold axis (Fig. 4). The binding of PLP to Sar2028 is suggested from the slight yellow colour of the crystal. The sequence identity between Sar2028 and its closest match in the

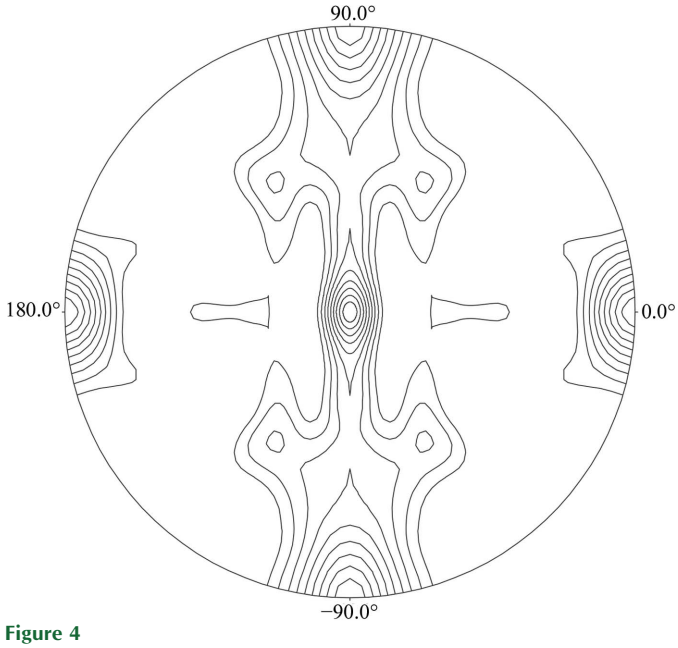


Figure 4
Self-rotation function of Sar2028 produced using *CCP*4 (Collaborative Computational Project, Number 4, 1994). The additional peak can be seen.

Protein Data Bank (PDB), aspartate aminotransferase from *E. coli* (PDB code 1qir), is less than 20% over the full protein length. However, *3DPSSM* (Kelley *et al.*, 2000) identifies the protein as having the same fold as other aminotransferases. The protein molecules are typically found as a dimer, which is consistent with our density calculation on Sar2028. PDB entry 1qir was used as a model for determining the structure of Sar2028 using molecular replacement as a phasing strategy. We have tried searching using both a monomer and a dimer. In addition, we have varied the model by truncation and by sequence changes. The molecular-replacement programs used were *Phaser* (Storoni *et al.*, 2004; McCoy *et al.*, 2005), *MrBUMP* (Collaborative Computational Project, Number 4, 1994) and *AMoRe* (Navaza, 1994; Collaborative Computational Project, Number 4, 1994). As yet, we have not been able to obtain a robust solution which is capable of refinement. We therefore expect to solve the structure by phasing using selenomethionine-variant protein. We have now purified selenomethionine-substituted protein using the same procedures as outlined above. Ongoing experiments are attempting to address the substrate profile of the protein. The structure and function of this protein may then suggest novel targets for therapeutic intervention and, in addition, new insights into this class of protein.

The protein was targeted as part of the Scottish Structural Proteomics Facility (SSPF), which is funded by the Scottish Higher Education Funding Council (SHEFC) and the Biotechnology and Biological Sciences Research Council UK (BBSRC). JS acknowledges the support of Karnatak University, UGC, New Delhi, ACU and the British Council, UK.

References

Boyle-Vavra, S. & Daum, R. S. (2007). *Lab. Invest.* **87**, 3–9.
Bush, K., Jacoby, G. A. & Medeiros, A. A. (1995). *Antimicrob. Agents Chemother.* **39**, 1211–1233.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Edwards, J. R. & Betts, M. J. (2000). *J. Antimicrob. Chemother.* **45**, 1–4.

Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.
Foster, T. (1996). In *Medical Microbiology*, 4th ed., edited by S. Baron. Galveston, TX, USA: University of Texas Medical Branch.
Gharahdaghi, F., Weinberg, C. R., Meagher, D. A., Imai, B. S. & Mische, S. M. (1999). *Electrophoresis*, **20**, 601–605.
Herold, B. C., Immergluck, L. C., Maranan, M. C., Lauderdale, D. S., Gaskin, R. E., Boyle-Vavra, S., Leitch, C. D. & Daum, R. S. (1998). *J. Amer. Med. Assoc.* **279**, 593–598.
Hirotsu, K., Goto, M., Okamoto, A. & Miyahara, I. (2005). *Chem. Rec.* **5**, 160–172.
Holden, M. T. G. *et al.* (2004). *Proc. Natl Acad. Sci. USA*, **101**, 9786–9791.
Johnson, A. P., Aucken, H. M., Cavendish, S., Ganner, M., Wale, M. C. J., Warner, M., Livermore, D. M. & Cookson, B. D. (2001) *J. Antimicrob. Chemother.* **48**, 143–144.
Kelley, L. A., McCallum, C. M. & Sternberg, M. J. (2000). *J. Mol. Biol.* **299**, 501–522.
Lawrence, C. L., Botting, C. H., Antrobus, R. & Coote, P. J. (2004). *Mol. Cell. Biol.* **24**, 3307–3323.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Oku, Y., Kurokawa, K., Ichihashi, N. & Sekimizi, K. (2004). *Microbiology*, **150**, 45–51.
Pesavento, G., Ducci, B., Comodo, N. & Nostro, A. L. (2007). *Food Control*, **18**, 196–200.
Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. (1996). *Anal. Chem.***68**, 850–858.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.