Structure of *Staphylococcus aureus* guanylate monophosphate kinase

Kamel El Omari,[a] Balvinder Dhaliwal,[a] Michael Lockyer,[b] Ian Charles,[b] Alastair R. Hawkins[c] and David K. Stammers[a]*

[a]Division of Structural Biology, The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, England, [b]Arrow Therapeutics, Britannia House, 7 Trinity Street, London SE1 1DA, England, and [c]Institute of Cell and Molecular Biosciences, Catherine Cookson Building, Medical School, Framlington Place, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne NE2 4HH, England

Correspondence e-mail: daves@strubi.ox.ac.uk
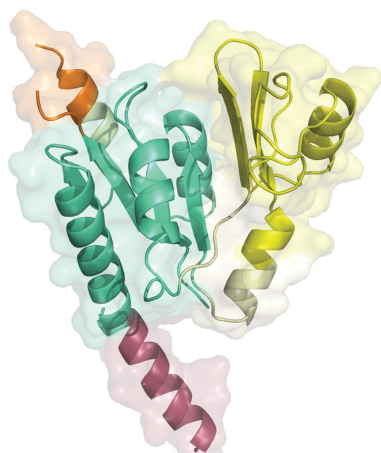
Nucleotide monophosphate kinases (NMPKs) are potential antimicrobial drug targets owing to their role in supplying DNA and RNA precursors. The present work reports the crystal structure of *Staphylococcus aureus* guanylate monophosphate kinase (SaGMK) at 1.9 Å resolution. The structure shows that unlike most GMKs SaGMK is dimeric, confirming the role of the extended C-terminus in dimer formation as first observed for *Escherichia coli* GMK (EcGMK). One of the two SaGMK dimers within the crystal asymmetric unit has two monomers in different conformations: an open form with a bound sulfate ion (mimicking the β-phosphate of ATP) and a closed form with bound GMP and sulfate ion. GMP-induced domain movements in SaGMK can thus be defined by comparison of these conformational states. Like other GMKs, the binding of GMP firstly triggers a partial closure of the enzyme, diminishing the distance between the GMP-binding and ATP-binding sites. In addition, the closed structure shows the presence of a potassium ion in contact with the guanine ring of GMP. The potassium ion appears to form an integral part of the GMP-binding site, as the Tyr36 side chain has significantly moved to form a metal ion–ligand coordination involving the lone pair of the side-chain O atom. The potassium-binding site might also be exploited in the design of novel inhibitors.

1. Introduction

Staphylococcus aureus is one of the most common causes of nosocomial infections found in intensive-care units (McGahee & Lowy, 2000). The ability of *S. aureus* to colonize and infect almost any tissue makes it responsible for a broad range of human diseases (Lowy, 1998). Many such infections can give rise to life-threatening conditions including bacteraemia, endocarditis, sepsis or toxic shock syndrome (Lowy, 1998). Worryingly, the emergence of methicillin-resistant *S. aureus* (MRSA) strains makes the treatment of infections arising from this pathogen increasingly difficult. Multidrug-resistant MRSA has become common in developed countries (Fluit *et al.*, 2001), requiring treatment by vancomycin. Unfortunately, vancomycin-resistant strains have also been reported (Hiramatsu, 2001). New and well chosen drug targets are thus necessary in order to develop novel antibacterials to overcome the current problems associated with the treatment of MRSA.

One such potential target for novel antibacterial drugs is *S. aureus* guanylate monophosphate kinase (SaGMK; EC 2.7.4.8), the structure of which we describe in this report. SaGMK is a 24 kDa protein which belongs to the nucleoside monophosphate kinase (NMPK) superfamily and catalyses reversible phosphoryl transfer from a nucleotide donor to a nucleotide acceptor. GMKs are responsible for phosphorylation of (d)GMP to (d)GDP using ATP as phosphoryl donor (Agarwal *et al.*, 1978). GMKs are involved in the synthesis of nucleotide precursors, indirectly modulating the synthesis of DNA and RNA; thus, inhibition of SaGMK would affect bacterial growth. Additional roles of GMKs that may have potential in other therapeutic areas derive from their involvement in the recycling of cGMP (Hall & Kuhn, 1986) as well as in the activation of some guanosine-analogue prodrugs used to treat certain cancers and viral infections

(Miller *et al.*, 1992). Recently, the structure of *Escherichia coli* guanylate kinase in complex with the monophosphate of the anti-herpes drug ganciclovir has been reported (Hible, Daalova *et al.*, 2006).

Like most NMPKs, GMKs consist of three domains: the core, the lid and the NMP-binding domains (Yan & Tsai, 1999). The core domain consists of a four-stranded β-sheet flanked by α-helices and also contains the P-loop which binds the ATP β-phosphate. The lid domain forms interactions with ATP when the protein is in the closed conformation. GMKs and other NMPKs are known to undergo conformational changes from an open and unbound form to a closed conformation bound to one ligand and then to a fully closed conformation bound to both ligands (Sekulic *et al.*, 2002; Kotaka *et al.*, 2006; Vonrhein *et al.*, 1995). The closure of the lid and the NMP-binding domain onto the core domain allows the assembly of the catalytic machinery and the formation of an associative transition state (Yan & Tsai, 1999).

To date, GMK crystal structures from *E. coli* (Hible *et al.*, 2005), *Mycobacterium tuberculosis* (Hible, Christova *et al.*, 2006), *Saccharomyces cerevisiae* (Stehle & Schulz, 1992) and mouse (Sekulic *et al.*, 2002) have been solved. We report here the structures of unliganded SaGMK and SaGMK complexed with GMP at 1.9 Å resolution.

2. Materials and methods

2.1. Protein purification

The *S. aureus* GMK-encoding DNA sequence was PCR-amplified using the following primers: GTCGTAACATATGGATAATG and GTGGATCCAACATTATTTTTTAG. The PCR product was subcloned into the *E. coli* expression vector pET15b using *Nde*I and *Bam*HI restriction sites to yield the recombinant plasmid pMUT67. The *E. coli* expression strain Codon⁺, transformed with pMUT67, was grown in an orbital incubator at 303 K in LB medium supplemented with 100 µg ml⁻¹ ampicillin and 35 µg ml⁻¹ chloramphenicol until an attenuance at 550 nm of 0.6 was reached. IPTG was then added to a final concentration of 0.2 mg ml⁻¹ and the culture was incubated for a further 5 h at 303 K. Cells were then harvested by centrifugation and the pellet was resuspended in buffer *A* (50 m*M* potassium phosphate pH 7.2, 1 m*M* DTT) and 1 m*M* benzamidine. The cells were disrupted by sonication and centrifuged at 10 000*g* for 42 min. The cell-free supernatant was applied onto an IMAC column (9 × 5 cm) charged to one-third capacity with Zn^{2+} and pre-equilibrated with buffer *A*. The column was eluted with a linear 0.0–0.3 *M* imidazole gradient in buffer *A* and SaGMK-containing fractions (identified by SDS–PAGE) were pooled, dialysed against buffer *A* and loaded onto a hydroxyapatite column (9 × 5 cm) pre-equilibrated with buffer *A*. SaGMK was eluted with a 1 l linear gradient of 50–400 m*M* potassium phosphate buffer pH 7.2 containing 1 m*M* DTT. On average, the purification protocol produced 400 mg pure protein from 25 g cell paste. Finally, SaGMK was buffer-exchanged into 20 m*M* Tris pH 7.2, 40 m*M* KCl and 0.1% azide and then concentrated to 18 mg ml⁻¹.

2.2. Crystallization and data collection

GMP (5 m*M*) and ATPγS (5 m*M)* were added to SaGMK for initial screening of 768 crystallization conditions, including those from Hampton Research, Wizard and Emerald kits. A Cartesian Technologies pipetting robot was used to set up 100 + 100 nl sitting drops in Greiner 96-well plates, which were placed in a TAP storage vault equipped with an automated imaging system at the Oxford Protein Production Facility (Walter *et al.*, 2005). After optimization,

Table 1
Statistics for crystallographic structure determination.

Values in parentheses are for the outermost resolution shell.

Data collection	
Data-collection site	ESRF ID14.1
Detector	MAR CCD
Wavelength (Å)	0.93
Resolution range (Å)	30–1.9 (1.97–1.90)
Redundancy	3.6 (2.9)
Completeness (%)	99.3 (95.6)
Average $I/\sigma(I)$	31.6 (2.9)
R_{merge}†	0.037 (0.275)
Refinement statistics	
No. of reflections used	75428
Total No. of atoms	6201
No. of water molecules	471
Resolution range (Å)	29.4-1.9
R factor‡ (R_{work}/R_{free})	20.7/23.9
R.m.s. bond-length deviation (Å)	0.013
R.m.s. bond-angle deviation (°)	1.32
Mean B factors (Å²)	
Main chain	23.2
Side chain and water	27.3
GMP	22.8
All atoms	25.6

† $R_{merge} = \sum |I - \langle I \rangle| / \sum \langle I \rangle$. ‡ R factor $= \sum |F_o - F_c| / \sum F_o$.

crystals grew in 20% PEG 3350, 0.2 *M* LiSO₄ and 0.1 *M* Tris–HCl pH 8.

X-ray diffraction data were collected at the ESRF ID14.1 beamline from a crystal cryoprotected by the addition of 20% ethylene glycol. Images were indexed and integrated with *DENZO* and data were merged using *SCALEPACK* (Otwinowski & Minor, 1996). The space group was $P2_1$ and the unit-cell parameters were $a = 70.0$, $b = 93.9$, $c = 83.9$ Å, $\beta = 110.6°$, with four monomers in the asymmetric unit. Detailed statistics for X-ray data collection and refinement are given in Table 1. The structure was solved using the automated molecular-replacement program *MrBUMP* (Keegan & Winn, 2006) with the coordinates of *E. coli* guanylate kinase (PDB code 2anb; Hible *et al.*, 2005) as a search model. Refinement was carried out with *REFMAC*5 (Murshudov *et al.*, 1997) using TLS (Winn *et al.*, 2001). Ramachandran plots generated from *PROCHECK* indicate that the model exhibits good stereochemistry, with 100% of the residues in the allowed regions. Structural superpositions were performed with *SHP* (Stuart *et al.*, 1979) and figures were drawn using *PyMOL* (DeLano, 2002).

3. Results and discussion

3.1. Overall structure of SaGMK

SaGMK crystallized with two dimers in the asymmetric unit. Each monomer is bound through its P-loop to a sulfate ion (Fig. 1*b*). Sulfate at 200 m*M* is a component of the crystallization medium and competition with ATPγS (5 m*M*) may account for the lack of binding of this ATP analogue in the crystal. In turn, the absence of bound ATP may explain why the lid domain is disordered and not observed in our structure. The sulfate ion has been proposed to mimic the ATP or ADP β-phosphate (Hible *et al.*, 2005). Indeed, the structure of mouse GMK in a closed conformation (PDB code 1lvg) shows that the P-loop interacts with an ADP β-phosphate. The binding of the sulfate ion to the P-loop may not be sufficient to fully order the lid domain. Indeed, a sulfate ion has been observed in this position in other NMPKs (Blaszczyk *et al.*, 2001), yet does not lead to the complete closure of the lid domain (Hible *et al.*, 2005). Electron density for GMP is clearly seen in only three of the four SaGMK

molecules; thus, the present structure allows study of the domain movement triggered by the binding of GMP.

SaGMK forms a dimer which is similar to that reported in the crystal structure of EcGMK (PDB code 1s96; Fig. 1*b*). The monomers are related by a twofold axis perpendicular to the helices responsible for the dimer stabilization. Surprisingly, in our structure the crystallographic dimer is also linked by a disulfide bridge between Cys182 of each monomer (Fig. 1*b*). As the cytoplasm is essentially a reducing environment (Kadokura *et al.*, 2003), this disulfide bridge is unlikely to normally exist *in vivo*. Nevertheless, because the SaGMK dimer is very similar to the EcGMK dimer, which does not have a disulfide involved in stabilization, we can infer that this crystallographic dimer is biologically relevant. This is supported by dynamic light-scattering studies showing the presence of a single species corresponding to a dimer (Fig. 2). Nevertheless, it might also be possible that under solution conditions different from ours, SaGMK may form higher order oligomers as reported for EcGMK (Hible *et al.*, 2005). Additionally, it should be noted that structures of *M. tuberculosis* GMKs have been reported with and without intra-

dimer disulfide bonds (Hible, Christova *et al.*, 2006), but the differing oxidation states do not apparently give rise to significant conformational changes.

As expected, the overall fold of SaGMK is very similar to those of other GMKs (Figs. 1*a* and 3). The core domain (residues 4–33, 99–126 and 164–191) is a four-stranded β-sheet flanked by four α-helices. The NMP-binding site (residues 39–91) consists of two α-helices and four β-strands (Fig. 1*a*). The lid domain (residues 124–158) is disordered in our structure. In EcGMK, the C-terminus is longer than in other GMKs of known crystal structure and is mainly responsible for the formation of dimers (Hible *et al.*, 2005). In the SaGMK structure, the C-terminus (residues 191–207) is of similar length to that of EcGMK, suggesting that GMKs containing this C-terminal extension (such as *Salmonella typhimurium* or *Vibrio cholerae* GMKs) are very likely to form dimers or oligomers. Indeed, GMKs from *M. tuberculosis*, *S. cerevisiae* and mouse have a shorter C-terminus and are known to behave as monomeric enzymes.

3.2. Active-site structure

The SaGMK active site (Fig. 4*b*) is very similar to that of mouse GMK, indicating that achieving drug selectivity towards a eukaryotic GMK in this region may not be straightforward. Indeed, the main interactions involved in GMP binding are the same as those previously reported in the *S. cerevisiae* GMK structure. The guanine ring makes hydrogen bonds to the conserved Glu74 and Ser39 and also to Glu105. The latter is replaced by an aspartate (Asp103) in the mouse GMK structure. The phosphate hydrogen bonds to Tyr55, Tyr83 and Arg43. A second arginine (Arg46) stabilizes a water molecule which is in contact with the GMP phosphate. The equivalent arginine in *M. tuberculosis* GMK (Arg60) or in mouse GMK (Arg44) hydrogen bonds to the GMP phosphate.

One interesting feature found in the active site is the presence of a potassium ion close to the GMP guanine ring (Fig. 4*b*). This ion has been reported in the mouse GMP structure at a similar position to that observed here. The potassium ion is in contact with the guanine ring N7 (3.1 Å compared with 2.8 Å in the mouse structure; PDB code 1lgv). In the SaGMK and mouse GMK structures the potassium ion is coordinated to the side chain of a serine (Ser39 and Ser37, respectively), to the main-chain carbonyl of an acidic residue (Glu103 and Asp101, respectively) and to three water molecules. In the

(*a*)

(*b*)
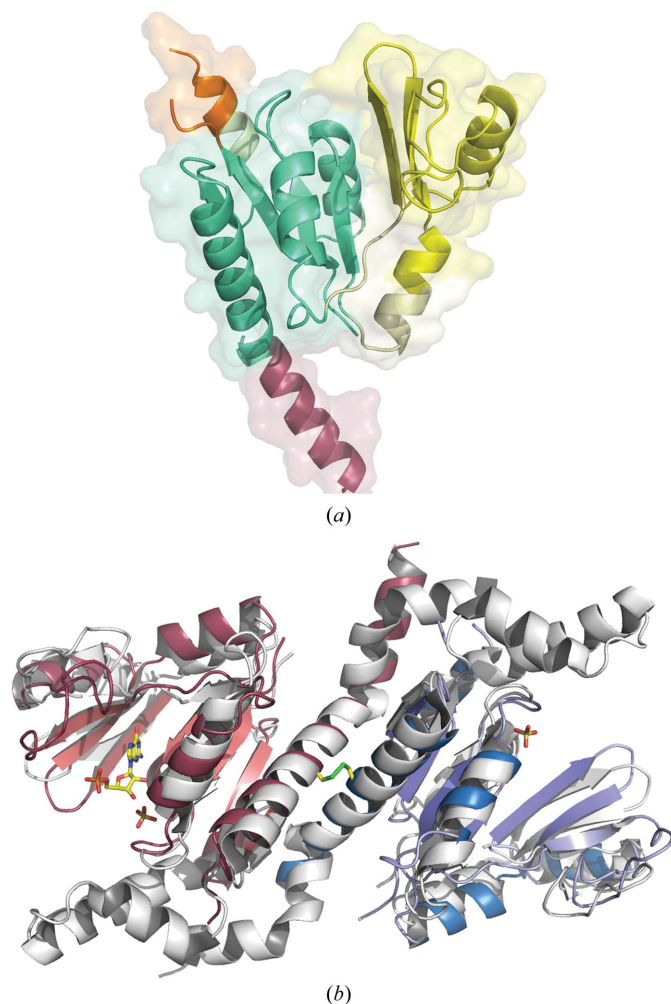
Figure 1
(*a*) Ribbon diagram and surface representation of SaGMK, showing the various domains. The NMP domain is coloured yellow, the core domain green, the visible part of the lid domain orange and the C-terminal domain raspberry. (*b*) Ribbon diagram showing superimposition of the SaGMK and EcGMK dimers. Bound ligands in SaGMK (GMP, sulfate) and the intradimer disulfide bridge are shown in standard atom colours. SaGMK subunits are coloured blue and red, whilst the EcGMK structure is coloured grey.
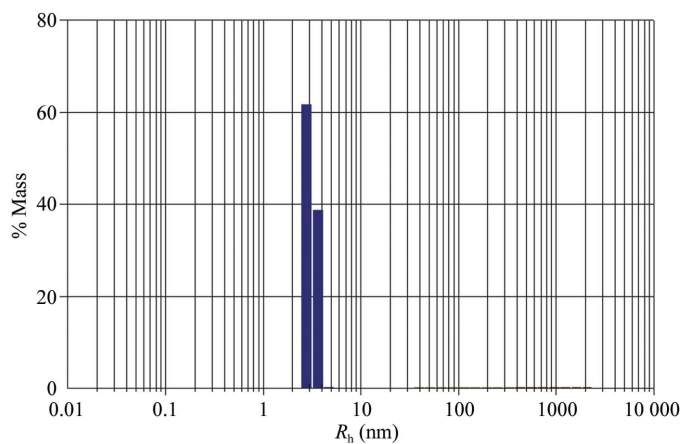
Figure 2
Representative DLS mass particle size distribution of SaGMK (1 mg ml^{-1}) in 20 m*M* Tris pH 7, 150 m*M* NaCl, 1 m*M* DTT. The protein displays a hydrodynamic radius of 3.1 nm, which gives an extrapolated molecular weight of 47.5 kDa corresponding to a dimer of SaGMK.

protein structure communications

SaGMK structure there is an extra metal–ligand contact made between the potassium ion and the Tyr36 side-chain O atom lone pair. This latter contact is absent in mouse GMK since the tyrosine is replaced by phenylalanine. The importance of the potassium for the binding of GMP is not known. However, from our structure we can infer that the presence of this ion is related to the binding of GMP as no potassium is seen in the unliganded SaGMK. The potassium ion, which comes from the enzyme-storage buffer and the crystallization reservoir solution (40 m*M* KCl in each case), may play a role in GMP binding by linking the nucleotide to a nearby residue.

3.3. Domain movements

The presence of liganded and unliganded SaGMK in the same crystal allows the identification of conformational changes induced by the binding of GMP (Figs. 4*a* and 4*b*). As expected, the binding of
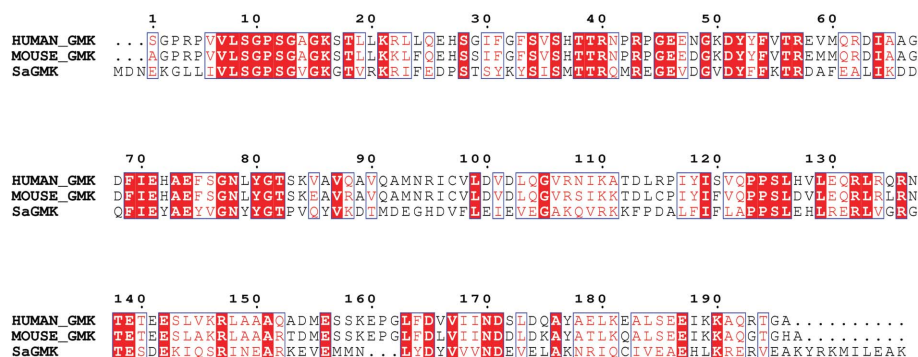
Figure 3
Amino-acid sequence alignment of human GMK, mouse GMK and SaGMK. SaGMK is 33 and 35% identical to human and mouse GMKs, respectively. Human and mouse GMKs share 87% sequence identity.

(*a*)

(*b*)
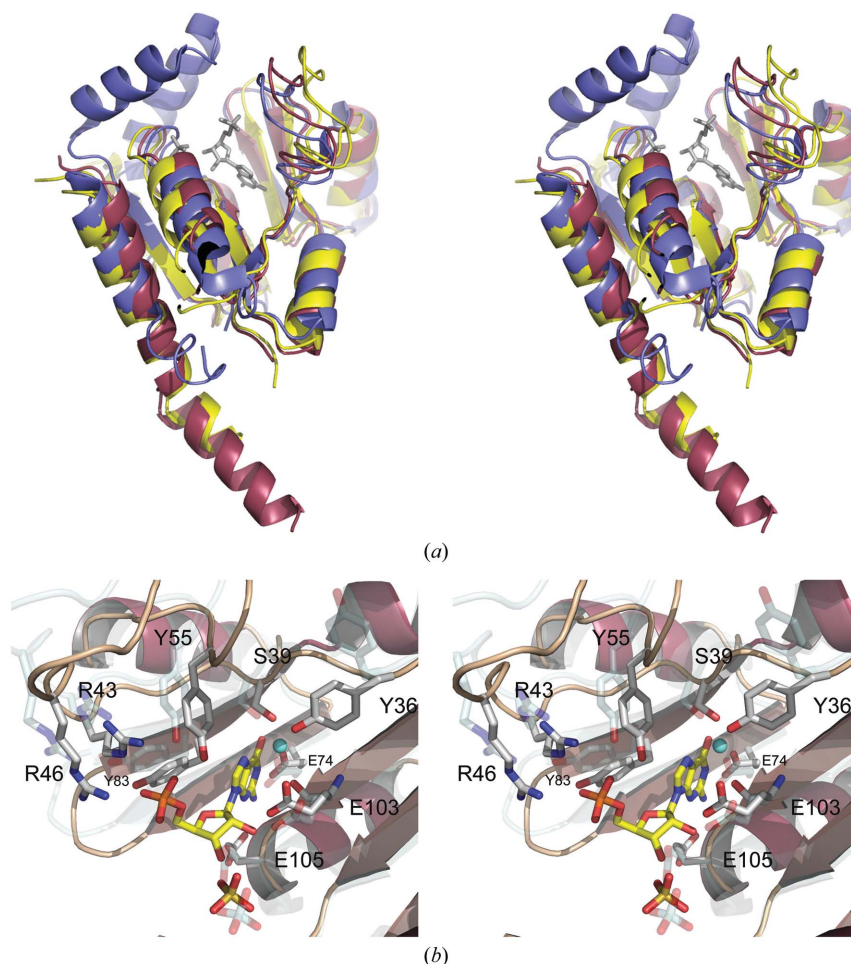
Figure 4
(*a*) Ribbon diagram showing the superimposition of the SaGMK open state (coloured yellow), closed state (coloured raspberry) and the mouse GMK closed state (coloured blue). The SaGMK ligands, sulfate and GMP, are in grey. (*b*) Stereo diagram showing the superimposition of the open conformation of the SaGMK active site (transparent representation) onto the closed conformation. The GMP and sulfate are in standard atom colours; the potassium ion is coloured cyan.

GMP induced domain movements in SaGMK. The NMP-binding domain and the P-loop moved towards each other, bringing GMP and the sulfate ion into close proximity (Fig. 4a).

In the SaGMK active site (Fig. 4b), the essential residues for GMP binding move to form a more compact pocket around the substrate. Interestingly, the side chain of Glu105 is rotated by 180° about the $C^\alpha{-}C^\beta$ bond in order to avoid a clash with the ribose ring. Tyr36 is not in direct contact with GMP; rather, it interacts *via* the potassium ion in the closed form. In the open form, the Tyr36 side chain points approximately 90° away from this direction. Thus, it seems that the potassium ion is responsible for the Tyr36 side-chain movement. It would be interesting to study the effect of the potassium ion on the enzyme reaction in order to determine if it is essential, at least for mouse GMK and SaGMK. For example, if Tyr36 was mutated to phenylalanine, the effects of the predicted loss in potassium binding on catalytic turnover or GMP binding could be assessed. Such information could be relevant for structure-based drug design involving GMKs in a range of therapeutic areas. Thus, inhibitors targeting SaGMK as potential antibacterials could be optimized for potency. Additionally, it could provide the basis for the improvement of compounds metabolized by mammalian GMKs; for example, anticancer prodrugs which undergo activation, including thiopurines (Sekulic *et al.*, 2002).

References

Agarwal, K. C., Miech, R. P. & Parks, R. E. Jr (1978). *Methods Enzymol.* **51**, 483–490.

Blaszczyk, J., Li, Y., Yan, H. & Ji, X. (2001). *J. Mol. Biol.* **307**, 247–257.

DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA, USA.

Fluit, A. C., Wielders, C. L., Verhoef, J. & Schmitz, F. J. (2001). *J. Clin. Microbiol.* **39**, 3727–3732.

Hall, S. W. & Kuhn, H. (1986). *Eur. J. Biochem.* **161**, 551–556.

Hible, G., Christova, P., Renault, L., Seclaman, E., Thompson, A., Girard, E., Munier-Lehmann, H. & Cherfils, J. (2006). *Proteins*, **62**, 489–500.

Hible, G., Daalova, P., Gilles, A. M. & Cherfils, J. (2006). In the press.

Hible, G., Renault, L., Schaeffer, F., Christova, P., Radulescu, A. Z., Evrin, C., Gilles, A. M. & Cherfils, J. (2005). *J. Mol. Biol.* **352**, 1044–1059.

Hiramatsu, K. (2001). *Lancet Infect. Dis.* **1**, 147–155.

Kadokura, H., Katzen, F. & Beckwith, J. (2003). *Annu. Rev. Biochem.* **72**, 111–135.

Keegan, R. M. & Winn, M. D. (2006). In preparation.

Kotaka, M., Dhaliwal, B., Ren, J., Nichols, C. E., Angell, R., Lockyer, M., Hawkins, A. R. & Stammers, D. K. (2006). *Protein Sci.* **15**, 774–784.

Lowy, F. D. (1998). *N. Engl. J. Med.* **339**, 520–532.

McGahee, W. & Lowy, F. D. (2000). *Semin. Respir. Infect.* **15**, 308–313.

Miller, W. H., Daluge, S. M., Garvey, E. P., Hopkins, S., Reardon, J. E., Boyd, F. L. & Miller, R. L. (1992). *J. Biol. Chem.* **267**, 21220–21224.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.

Sekulic, N., Shuvalova, L., Spangenberg, O., Konrad, M. & Lavie, A. (2002). *J. Biol. Chem.* **277**, 30236–30243.

Stehle, T. & Schulz, G. E. (1992). *J. Mol. Biol.* **224**, 1127–1141.

Stuart, D. I., Levine, M., Muirhead, H. & Stammers, D. K. (1979). *J. Mol. Biol.* **134**, 109–142.

Vonrhein, C., Schlauderer, G. J. & Schulz, G. E. (1995). *Structure*, **3**, 483–490.

Walter, T. S., Diprose, J. M., Mayo, C. J., Siebold, C., Pickford, M. G., Carter, L., Sutton, G. C., Berrow, N. S., Brown, J., Berry, I. M., Stewart-Jones, G. B., Grimes, J. M., Stammers, D. K., Esnouf, R. M., Jones, E. Y., Owens, R. J., Stuart, D. I. & Harlos, K. (2005). *Acta Cryst.* D**61**, 651–657.

Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.

Yan, H. & Tsai, M. D. (1999). *Adv. Enzymol. Relat. Areas Mol. Biol.* **73**, 103–134.