Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural and biochemical characterization of N^5-carboxyaminoimidazole ribonucleotide synthetase and N^5-carboxyaminoimidazole ribonucleotide mutase from *Staphylococcus aureus*

Pedro Brugarolas,[a] Erica M. Duguid,[b] Wen Zhang,[a] Catherine B. Poor[b] and Chuan He[a]*

[a]Department of Chemistry and Institute for Biophysical Dynamics, The University of Chicago, 929 East 57th Street, GCIS E321, Chicago, IL 60637, USA, and [b]Department of Biochemistry and Molecular Biology, The University of Chicago, 929 East 57th Street, Chicago, IL 60637, USA

Correspondence e-mail: chuanhe@uchicago.edu

Received 28 January 2011
Accepted 17 June 2011

PDB References: PurK complexed with MgADP, 3orq; PurK, 3orr; PurE, 3ors.

With the rapid rise of methicillin-resistant *Staphylococcus aureus* infections, new strategies against *S. aureus* are urgently needed. *De novo* purine biosynthesis is a promising yet unexploited target, insofar as abundant evidence has shown that bacteria with compromised purine biosynthesis are attenuated. Fundamental differences exist within the process by which humans and bacteria convert 5-aminoimidazole ribonucleotide (AIR) to 4-carboxy-5-aminoimidazole ribonucleotide (CAIR). In bacteria, this transformation occurs through a two-step conversion catalyzed by PurK and PurE; in humans, it is mediated by a one-step conversion catalyzed by class II PurE. Thus, these bacterial enzymes are potential targets for selective antibiotic development. Here, the first comprehensive structural and biochemical characterization of PurK and PurE from *S. aureus* is presented. Structural analysis of *S. aureus* PurK reveals a nonconserved phenylalanine near the AIR-binding site that occupies the putative position of the imidazole ring of AIR. Mutation of this phenylalanine to isoleucine or tryptophan reduced the enzyme efficiency by around tenfold. The K_m for bicarbonate was determined for the first time for a PurK enzyme and was found to be ~18.8 m*M*. The structure of PurE is described in comparison to that of human class II PurE. It is confirmed biochemically that His38 is essential for function. These studies aim to provide foundations for future structure-based drug-discovery efforts against *S. aureus* purine biosynthesis.

1. Introduction

About one third of the world's population carry *Staphylococcus aureus* innocuously in their nares; under certain circumstances, *S. aureus* can breach natural barriers and cause infections ranging from minor skin abscesses to life-threatening conditions such as toxic shock syndrome and bacteremia (Lowy, 1998). *S. aureus* infections, including those caused by methicillin-resistant *S. aureus* (MRSA), are the main cause of hospital-acquired infections and are becoming prevalent in community settings (Deleo *et al.*, 2010). It is estimated that 0.8% of all hospital discharges in the US arise from *S. aureus* infections, with an estimated annual burden of $9.5 billion and 12 000 deaths (Lodise & McKinnon, 2007; Noskin *et al.*, 2005). Alarmingly, 95% of all *S. aureus* clinical isolates are resistant to penicillin, once the drug of choice, and 20–50% are resistant to methicillin (Alekshun, 2005; Appelbaum, 2006). Despite these frightening statistics, antibiotic discovery and development is declining, with only three new classes of antibiotics, lipopeptides, oxazolidinones and mutilins, being brought to clinic in the last five decades (Fischbach & Walsh, 2009). An ideal yet unexploited target for

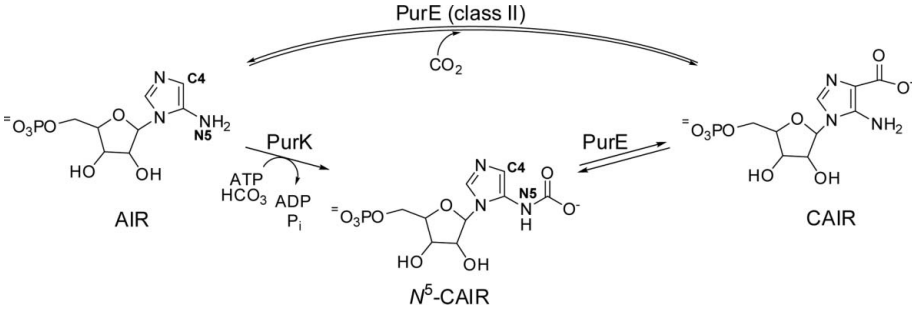


Figure 1
Conversion of AIR to CAIR. Mammals use a single enzyme (PurE class II), while bacteria, yeast and plants use two enzymes (PurE and PurK).

the development of new antibiotics is *de novo* purine biosynthesis. Numerous genetic and functional studies have shown that the virulence of many bacteria and fungi is attenuated by compromised *de novo* purine biosynthesis (McFarland & Stocker, 1987; Kirsch & Whitney, 1991; Perfect *et al.*, 1993; Drazek *et al.*, 1995; Wang *et al.*, 1996; Polissi *et al.*, 1998; Donovan *et al.*, 2001; Cersini *et al.*, 2003; Samant *et al.*, 2008). Most relevant for this study is the recent report that *Bacillus anthracis* (another Gram-positive human pathogen) *purE* and *purK* mutants are attenuated in human blood (Samant *et al.*, 2008) and the report that Gram-positive *Streptococcus pneumonia purE* and *purK* mutants are attenuated in a mouse model of infection (Polissi *et al.*, 1998). In addition, recent research in our laboratory has shown that *Staphylococcus aureus* purine-biosynthesis mutants display increased pigment production and decreased survival in a murine-abscess model of infection (Lan *et al.*, 2010). Finally, inhibitors of purine biosynthesis such as 6-thioguanine (6TG) have been shown to suppress the virulence of *S. aureus* (Lan *et al.*, 2010). However, these substrate analogs have strong immunosuppressive activity in humans and therefore are not suitable antibiotics.

De novo purine biosynthesis begins with phosphoribosyl pyrophosphate (PRPP), on which the purine base is built stepwise to generate inosine monophosphate (IMP), the common precursor of AMP and GMP (Supplementary Scheme 1[1]). This pathway is composed of 11 steps in bacteria, yeasts and plants, but only ten in animals (Tiedeman *et al.*, 1989; Watanabe *et al.*, 1989; Meyer *et al.*, 1992; Firestine & Davisson, 1994; Firestine *et al.*, 1994; Chung *et al.*, 1996; Schmuke *et al.*, 1997; Sørensen & Dandanell, 1997). This difference derives from the conversion of AIR to CAIR. In bacteria, this is a two-step conversion catalyzed by two different enzymes. Firstly, PurK (N^5-CAIR synthetase) catalyzes the conversion of bicarbonate, ATP and AIR to ADP and N^5-CAIR. Subsequently, PurE catalyzes an unusual mutase reaction in which N^5-CAIR is converted to CAIR (Fig. 1; Mueller *et al.*, 1994). In humans this conversion is catalyzed by a single enzyme (class II PurE), which directly adds CO_2 to the C4 position of AIR to generate CAIR (Li *et al.*, 2007). These differences at the molecular level make PurK and PurE promising targets for the development of structure-based therapies.

Here, we present the first comprehensive enzymatic and structural characterization of PurE and PurK from *S. aureus*. Several structures of PurK and PurE from other organisms have previously been described and enzymatically characterized (Meyer *et al.*, 1992; Firestine *et al.*, 1994; Mueller *et al.*, 1994; Li *et al.*, 2007; Mathews *et al.*, 1999; Thoden *et al.*, 1999, 2008, 2010; Settembre *et al.*, 2004; Boyle *et al.*, 2005; Constantine *et al.*, 2006; Hoskins *et al.*, 2007). Therefore, we chose to forgo an extensive description. Instead, we present a comparison of the newly obtained *S. aureus* PurK structure and the archetype *E. coli* PurK. We also compare *S. aureus* PurE with human PurE. We believe that the characterization of PurK and PurE from the major human pathogen *S. aureus* will be valuable for future drug-development efforts and that their unique features presented here may be exploitable when trying to design specific antibiotics against *S. aureus*, in particular considering the recent results on purine biosynthesis affecting virulence in *S. aureus* (Lan *et al.*, 2010).

2. Materials and methods

2.1. Cloning, expression and purification

S. aureus PurK, PurE and PurC were cloned from *S. aureus* Newman genomic DNA into the target vector pMCSG19 by the ligation-independent cloning method (Donnelly *et al.*, 2006). The resulting plasmid was delivered into *Escherichia coli* BL21 Star (DE3) competent cells containing pRK1037 (Science Reagents Inc.) by heat shock. Positive colonies were selected on plates containing 100 µg ml^{-1} ampicillin and 30 µg ml^{-1} kanamycin. A 10 ml overnight culture grown from a single colony was used to inoculate 1 l autoclaved LB medium containing 100 µg ml^{-1} ampicillin and 30 µg ml^{-1} kanamycin. Cells were grown at 310 K and 250 rev min^{-1} to an $OD_{600\,nm}$ of 0.6 and the temperature was then reduced to 289 K. 15 min later, protein expression was induced with 1 m*M* IPTG and the cells were grown overnight at 289 K. The next morning, the cells were harvested at 277 K by centrifugation at 6000 rev min^{-1} for 8 min. All subsequent steps were performed at 277 K. The cell pellet was suspended in 20 ml buffer *A* (10 m*M* Tris–HCl pH 7.4, 300 m*M* NaCl, 1 m*M* $MgCl_2$, 10 m*M* β-mercaptoethanol) with 5% glycerol and 10 m*M* PMSF. The cells were lysed by sonication and centrifuged at 12 000 rev min^{-1} for 25 min. The supernatant was filtered through a 0.45 µm filter and applied onto an Ni–NTA column. The column was washed with 10% buffer *B* (10 m*M* Tris–HCl pH 7.4, 500 m*M* imidazole, 300 m*M* NaCl, 1 m*M* $MgCl_2$, 10 m*M* β-mercaptoethanol) and eluted with a linear

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: EA5142). Services for accessing this material are described at the back of the journal.

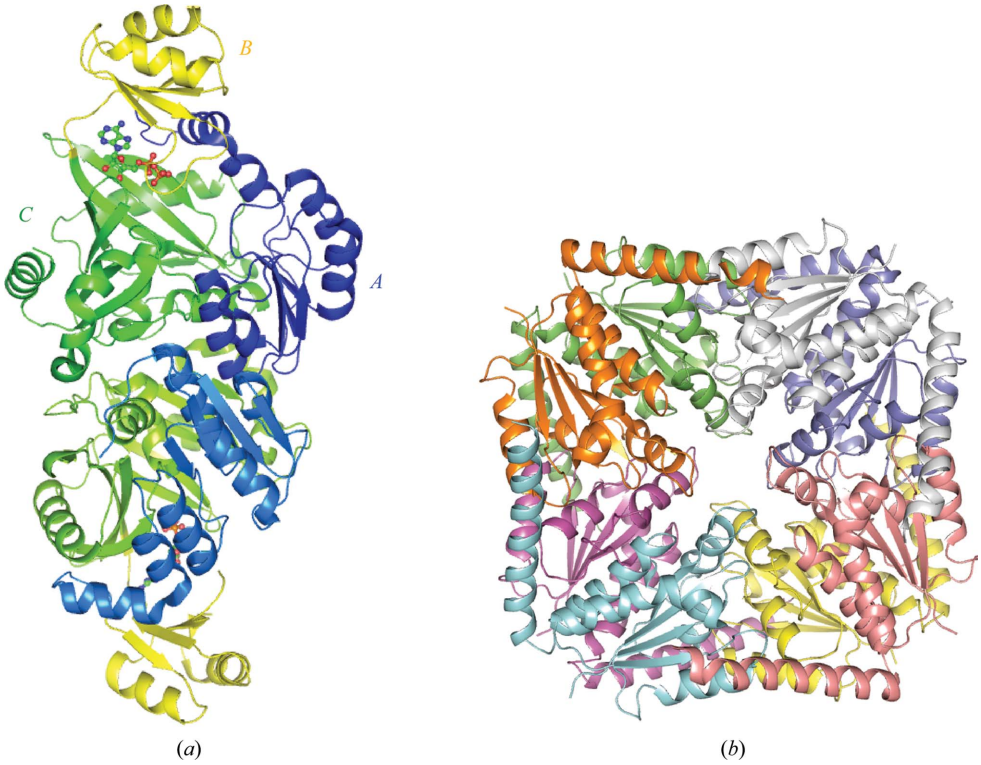
gradient from 10% to 100% buffer *B* over 40 ml. Peak fractions were pooled and kept at 277 K. The purity of the proteins was verified by SDS–PAGE analysis. Fractions containing pure protein were combined, concentrated and run on a desalting column with buffer *A*. The polyhistidine tag was removed by protease digestion using 0.01 mg TEV protease per milligram of protein overnight at 277 K. A small fraction of the protein was run on a Superdex 200 gel-filtration column to determine the associated state of the protein in solution. The rest of the protein was concentrated to ∼0.2 m*M* and stored at 193 K with 33% glycerol. The selenomethionine form of PurK was expressed under the same conditions except that it was grown in minimal medium (per litre: 12.8 g $Na_2HPO_4{\cdot}7H_2O$, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl, 4 g glucose, 0.5 mg thiamine, 0.1 m*M* $CaCl_2$, 2 m*M* $MgSO_4$). When the culture reached an $OD_{600\,nm}$ of 0.6, the following amino acids were added: 100 mg lysine, 100 mg threonine, 100 mg phenylalanine, 50 mg leucine, 50 mg valine and 60 mg selenomethionine. Protein expression was induced 15 min later with 1 m*M* IPTG and the temperature was reduced to 289 K. Every step thereafter was identical to those used for the wild-type protein.

2.2. Site-directed mutagenesis

Protein variants were generated using an Agilent/Stratagene site-directed mutagenesis kit using the following primers: PurEH38F, 5′-GAAAAACAAGTAGTATCCGCATTTCGTACGCCAAAAATGATGG-3′ and 5′-CCATCATTTTTGGCGTACGAAATGCGGATACTACTTGTTTTTC-3′; PurEH38N, 5′-ACAAGTAGTATCCGCAAATCGTACGCCAAAAATG-3′ and 5′-CATTTTTGGCGTACGATTTGCGGATACTACTTGT-3′; PurEH38W, 5′-GAAAAACAAGTAGTATCCGCATGGCGTACGCCAAAAATGATGGTTC-3′ and 5′-GAACCATCATTTTTGGCGTACGCCATGCGGATACTACTTGTTTTTC-3′; PurKF78W, 5′-GATGTGATTACTTATGAATGGGAAAACATTTCAGCCCAAC-3′ and 5′-GTTGGGCTGAAATGTTTTCCCATTCATAAGTAATCACATC-3′; PurKF78I, 5′-GTGATGTGATTACTTATGAAATTGAAAACATTTCAGCCCAAC-3′ and 5′-GTTGGGCTGAAATGTTTTCAATTTCATAAGTAATCACATCAC-3′.

2.3. Crystallization and structure determination

Crystallization trials were performed by the hanging-drop vapor-diffusion method using the different proteins at concentrations of 5, 10 and 15 mg ml^{-1}. The protein solutions were concentrated using centrifugal filter units with appropriate cutoffs (Amicon Ultra, Millipore). Protein concentration was determined from the absorption at 280 nm using calculated extinction coefficients (PurK, 24 890 M^{-1} cm^{-1}; PurE, 10 180 M^{-1} cm^{-1}). PurK and its selenomethionine form crystallized at room temperature at a concentration of 10 mg ml^{-1} with a reservoir solution consisting of 0.1 *M* Bis-Tris pH 6.5, 28%(*w*/*v*) polyethylene glycol monomethyl ether 2000 after 24 h. PurK complexed with ADP was obtained by adding ATP to the drop to a final concentration of 10 m*M* for 5 h. PurE crystallized at room temperature at 10 mg ml^{-1} with a reservoir solution consisting of 0.1 *M* sodium acetate trihydrate pH 4.5, 2 *M* ammonium sulfate after 24 h. Crystals were frozen in liquid nitrogen following cryoprotection with reservoir solution containing 20% glycerol. Data were collected at the Advanced Photon Source at Argonne National Laboratory on beamlines 19-ID and 23-ID-B and all images were indexed, integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). The structure of the apo form of PurK was solved at 2.23 Å resolution using the multiple-wavelength anomalous dispersion method in *PHENIX* (Adams *et al.*, 2010). PurK complexed with ATP was solved at 2.23 Å resolution by molecular replacement using the previously obtained PurK structure as the search model. The PurE structure was solved at 1.45 Å resolution by



(*a*) (*b*)

Figure 2
Crystal structures of *S. aureus* PurK and PurE. (*a*) Overall structure of *S. aureus* PurK dimer in complex with MgADP showing domains A, B and C in green, yellow and blue, respectively. ADP is shown in a ball-and-stick representation. Different tones of green, yellow and blue are used for each subunit. (*b*) Overall structure of *S. aureus* PurE showing each monomer in a different color.

Table 1
Statistics for crystallographic structure determination.

Values in parentheses are for the highest resolution bin.

Protein (PDB code)	Apo PurK (3orr)	PurK complexed with MgADP (3orq)	Apo PurE (3ors)
Data collection			
Space group	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters			
a (Å)	66.1	66.4	81.3
b (Å)	53.3	53.4	78.2
c (Å)	107.4	110.4	93.0
$\alpha = \gamma$ (°)	90.0	90.0	90.0
β (°)	95.3	94.7	113.9
Wavelength	0.97951, 0.97937, 0.97196	1.03321	0.97937
Resolution limits (Å)	50–2.23 (2.40–2.23)	50–2.15 (2.23–2.15)	50–1.45 (1.51–1.45)
No. of independent reflections	34628 (3208)	42717 (2693)	187255 (20172)
Completeness (%)	83.8 (45.7)	94.5 (63.9)	99.5 (96.9)
Multiplicity	3.1 (1.9)	3.5 (2.5)	3.7 (3.3)
$\langle I/\sigma(I)\rangle$	24.6 (1.64)	22.9 (2.57)	16.9 (2.70)
$R_{\rm merge}$† (%)	0.057 (0.460)	0.057 (0.361)	0.075 (0.389)
Refinement			
Resolution limits (Å)	40–2.23	44.1–2.23	50–1.45
No. of reflections	34554	35400	177862
$R_{\rm work}/R_{\rm free}$‡	19.4/25.1	17.1/21.6	17.8/20.2
No. of atoms			
Protein	5889	5913§	9919¶
Heteroatoms	217††	255‡‡	955§§
B factors (Å^2)			
Protein	40.5	50.3	14.1
Water	32.4	64.3	34.2
R.m.s. deviations			
Bond lengths (Å)	0.016	0.016	0.008
Bond angles (°)	1.39	1.436	1.164
Ramachandran statistics (%)			
Favored	97.3	97.4	98.9
Allowed	2.7	2.5	1.1
Disallowed	0	0.1	0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl) \times 100$. ‡ R factor $= \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}| \times 100$, where $F_{\rm obs}$ is the observed structure-factor amplitude and $F_{\rm calc}$ is the calculated structure-factor amplitude. § These include multiple conformations for Met31. ¶ These include multiple conformations for Asn21, Ser49, Ile56 and Ser99 in chain *A*, Met44, Ser75 and Ser99 in chain *N*, Lys2, Lys14, Met43, Ile94 and Ser99 in chain *C*, Ser3, Ser11, Lys14, Glu27, Met43, Ile56, Ser99, Ser147 and Glu153 in chain *D*, Met1, Ser11, Met43, Met44, Ile53 and Leu159 in chain *E*, Met44 and Ser75 in chain *F*, Glu17, Met43, Met44, Ile50, Ser99 and Leu140 in chain *G* and Trp13, Ile15, Glu17, Glu27, Met43, Met44, Ser49, Glu53, Thr88, Ser99, Ile114, Ser136, Asn142 and Asn145 in chain *F*. †† The heteroatoms include 217 waters. ‡‡ The heteroatoms include two ADPs, two Mg^{2+} ions, two pyrophosphates, one glycerol and 186 waters. §§ The heteroatoms include 16 sulfates and 865 waters.

molecular replacement using the structure of *B. anthracis* PurE (Boyle *et al.*, 2005; PDB entry 1xmp) as the search model with *Phaser* from the *CCP*4 suite (Winn *et al.*, 2011). All models were built using the program *Coot* and refined with *REFMAC* (Emsley *et al.*, 2010; Winn *et al.*, 2011). Figures were prepared with *PyMOL* (DeLano, 2002). Data-collection and refinement statistics can be found in Table 1. The atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 3orq, 3ors and 3orr.

2.4. Enzymatic assays

All enzymatic assays were carried out using a Cary 300 UV–Vis spectrophotometer at room temperature. AIR was synthesized according to the procedure of Meyer *et al.* (1992). $NaHCO_3$ solutions were prepared fresh every day prior to experiments.

The enzymatic assay for PurK was adapted from Meyer *et al.* (1992). A schematic view of the assay is shown in Supplementary Scheme 2(*a*). Briefly, 2 µl PurK (10 µ*M*) was added to a microcuvette containing 150 µl 50 m*M* HEPES, 80 m*M* KCl, 20 m*M* $MgCl_2$, 2 m*M* phosphoenolpyruvate, 0.2 m*M* NADH, 1 unit ml^{-1} pyruvate kinase, 2 unit ml^{-1} lactate dehydrogenase, 4 µ*M* PurE and varying amounts of AIR, ATP and $NaHCO_3$. PurE was included in this assay to ensure that no additional AIR was present arising from the spontaneous decomposition of N^5-CAIR. When measuring the $K_{\rm m}$ for $NaHCO_3$ the buffer was freshly prepared and degassed under vacuum for 1 h prior to the experiment. The progress of the reaction was followed by the reduction in absorbance at 340 nm arising from the consumption of NADH ($\varepsilon_{340\,\rm nm}$ = 6220 M^{-1} cm^{-1}) over 5 min and the initial velocities were recorded.

The enzymatic assay for PurE was adapted from Hoskins *et al.* (2007). A schematic view of the assay is shown in Supplementary Scheme 2(*b*). Briefly, 2 µl PurE (3 µ*M*) was added to a microcuvette containing 150 µl 50 m*M* HEPES, 80 m*M* KCl, 20 m*M* $MgCl_2$, 4 m*M* phosphoenolpyruvate, 1 m*M* ATP, 10 m*M* L-aspartic acid, 1/100 of saturated $NaHCO_3$, 1 unit ml^{-1} pyruvate kinase, 4 µ*M* PurK, 4 µ*M* PurC and varying amounts of AIR. The reaction was followed by the increase in absorbance at 282 nm arising from the formation of SAICAR ($\varepsilon_{282\,\rm nm}$ = 8607 M^{-1} cm^{-1}) over 5 min and the initial velocities were recorded.

In all the coupled assays, control runs were performed to ensure that the enzyme under investigation was rate-limiting. Generally, 200–1000-fold excesses of the coupling enzymes were used.

3. Results

3.1. Structures of apo PurK and PurK complexed with ADP at 2.3 Å resolution

S. aureus PurK forms a homodimer in the crystalline lattice (Fig. 2*a*) and in solution as shown by gel-filtration chromatography (Supplementary Fig. 1). This protein belongs to the ATP-grasp superfamily and is composed of three domains (A, Met1–Val123; B, Pro124–Ile188; C, Lys189–Asn374; Fig. 2*a*). Domain A contains a five-stranded parallel β-sheet flanked by two α-helices on each face. A helix–residue–helix motif characteristic of the ATP-grasp superfamily connects domains A and B. Domain B is composed of two α-helices and a four-

stranded antiparallel β-sheet. Domain C contains a C-shaped nine-stranded antiparallel β-sheet flanked by two α-helices on the inside and three on the outside. Domains A and C pack against each other and against domains A and C from the other monomer to form the main body of the enzyme (Fig. 2*a*). Domain B protrudes away from the main body of the enzyme, forming a cleft that closes over domain C in the ADP-bound form. Higher temperature factors indicate the mobility of the B domain (Supplementary Fig. 2*a*), which is also apparent when comparing the relative position of the B domain in the ADP-bound and apo forms. An overlay of the apo and the ADP-bound forms of the enzyme shows that the C^{α} atoms of domains A and C align with a root-mean-square deviation (r.m.s.d.) of only 0.2 Å, while the C^{α} atoms of the B domains align with an r.m.s.d. of 2.9 Å (Supplementary Fig. 2*b*). Another indication of the flexibility of the B domain is the absence of electron density observed for the loop that extends from Gly152 to Gly158 in the apo form of the enzyme, thus indicating its disorder. This loop wraps around the diphosphate in the ADP-bound form. In the ADP-bound form we observed very clear electron density for MgADP in the ATP-binding site, indicating that ATP hydrolysis has occurred.

The electron-density quality for MgADP and the B domain is better for chain *A* of the dimer and the discussion below refers to this chain only. Fig. 3(*a*) shows the $F_o - F_c$ electron-density maps for MgADP and its interactions with the surrounding residues. ADP is anchored by Glu191 from domain C, which forms hydrogen bonds to the 2′-hydroxyl and 3′-hydroxyl groups of the ribose. The adenine base is held in a hydrophobic pocket surrounded by Leu186, Ile188, Phe256 and Tyr185, which forms π-stacking interactions with the base. The N6 position of adenine forms hydrogen bonds to Glu183 and the carbonyl O atom of Lys184, providing base-specificity. The diphosphate group interacts on one side with Arg108, Lys148 and Gln159 from the B domain and on the opposite side with an Mg^{2+} ion in an octahedral geometry coordinated by Glu254 and Glu267 from the C domain.

The binding site for AIR lies between domains A and C (Thoden *et al.*, 2010). Attempts to obtain the AIR-bound form either by cocrystallization or soaking failed to reveal any electron density for AIR. In monomer *B* of the ADP-bound form we observed electron density in the AIR site that we assigned as glycerol. Fig. 4(*a*) shows a close-up view of the AIR-binding site with AIR modeled in based on the structure of *Aspergillus clavatus* PurK complexed with AIR (Thoden *et al.*, 2010; PDB entry 3k5i). To generate this model we simply superimposed the two structures and placed AIR in the *S. aureus* structure. According to this model, AIR forms interactions from its 2′-hydroxyl and 3′-hydroxyl groups to Glu77. The phosphate interacts with Arg347 and Lys340. Ile20 and Phe78 provide a hydrophobic environment for the imidazole base.

The spatial organization of the ATP-binding and AIR-binding sites agrees with the previously proposed model for catalysis (Thoden *et al.*, 2008). According to the proposed mechanism, ATP reacts with bicarbonate to generate ADP and carboxyphosphate, which migrates to the AIR active site where it is nucleophilically attacked by the 5-amino group of AIR to generate N^5-CAIR. In our structural model the γ-phosphate of ATP would lie ~6 Å apart from the N5 exocyclic amine of AIR, with the binding site for bicarbonate located approximately in the middle. The residues which are proposed to interact with carboxyphosphate, Arg271, Lys348 and Glu155, are conserved among PurK proteins.

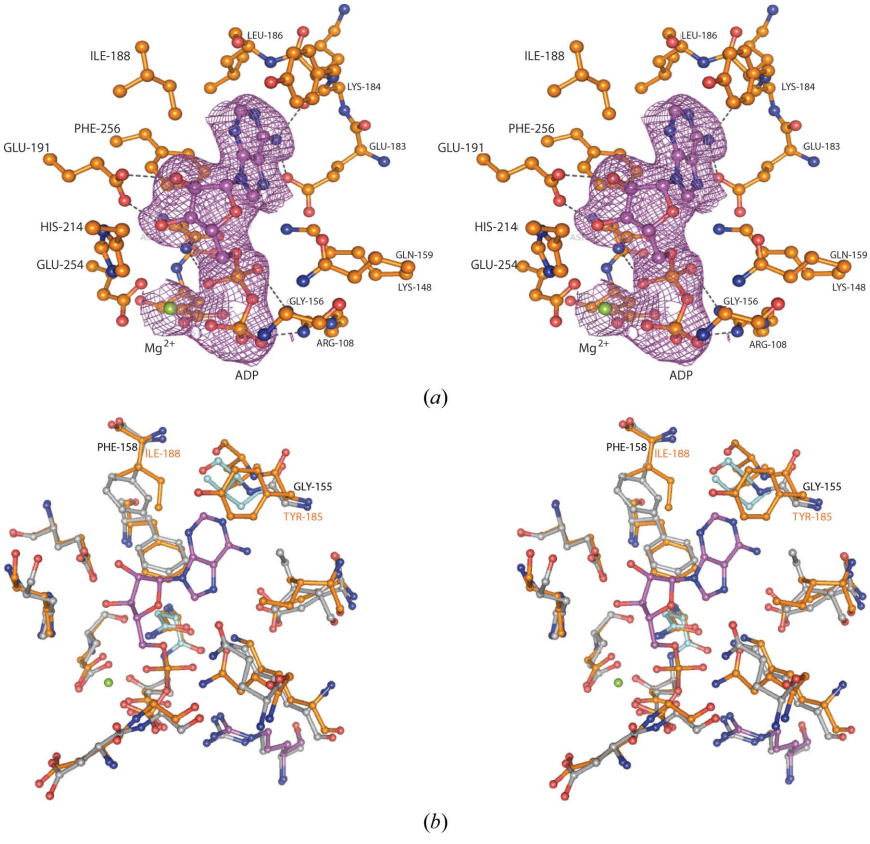


Figure 3
(*a*) Stereoview of the *S. aureus* PurK ATP-binding site showing $F_o - F_c$ electron-density maps at 2σ for MgADP, surrounding residues and interactions. The Mg^{2+} ion is shown in green. (*b*) Overlay of the *S. aureus* PurK ATP-binding site with MgADP (orange) and the *E. coli* PurK (PDB entry 3eth) ATP-binding site (gray). Residues that differ in the two structures are labeled.

3.2. Comparison of *S. aureus* PurK and *E. coli* PurK

Given the availability of structural and enzymatic data for *E. coli* PurK, we decided to examine differences between this protein and *S. aureus* PurK. *S. aureus* PurK and *E. coli* PurK share 30% sequence identity (see sequence alignment in Supplementary

Fig. 2) and have a common tertiary structure (r.m.s.d. = 1.8 Å over 355 residues). *S. aureus* PurK is slightly larger (374 *versus* 360 amino acids) as a result of a long loop in the N-terminus which is not present in the *E. coli* protein, as well as an additional α-helix–β-strand motif in the A domain inserted between strand β2 and helix α3 of the *E. coli* homolog. *E. coli* PurK has an additional 15-residue helix in the C-terminus. Interestingly, several differences exist between the active sites of *S. aureus* PurK and *E. coli* PurK. Fig. 3(*b*) shows a close-up detail of the ATP-binding sites of the two structures. Most notably, Tyr185 of *S. aureus* PurK, which forms π-stacking interactions with the adenine base of ATP, is replaced by a glycine in *E. coli* PurK. Ile188, which is located on the opposite side of the adenine base in *S. aureus* PurK, is replaced by a phenylalanine in *E. coli* PurK and does not π-stack with the adenine base.

Table 2
Kinetic parameters of PurK variants.

Reported errors are those obtained from curve-fitting the appropriate data set.

	K_m for AIR (μM)	K_m for ATP (μM)	K_m for HCO_3^- (mM)	k_{cat} (min^{-1})	k_{cat}/K_m† (μM^{-1} min^{-1})
PurK wild type	13.9 ± 1.7	43.2 ± 6.7	18.8 ± 3.9	1490 ± 66	107 ± 17
PurK F78I	51.3 ± 6.1	ND	ND	950 ± 75	18.5 ± 2.9
PurK F78W	103 ± 34	ND	ND	1080 ± 270	10.5 ± 6.1

† K_m for AIR.

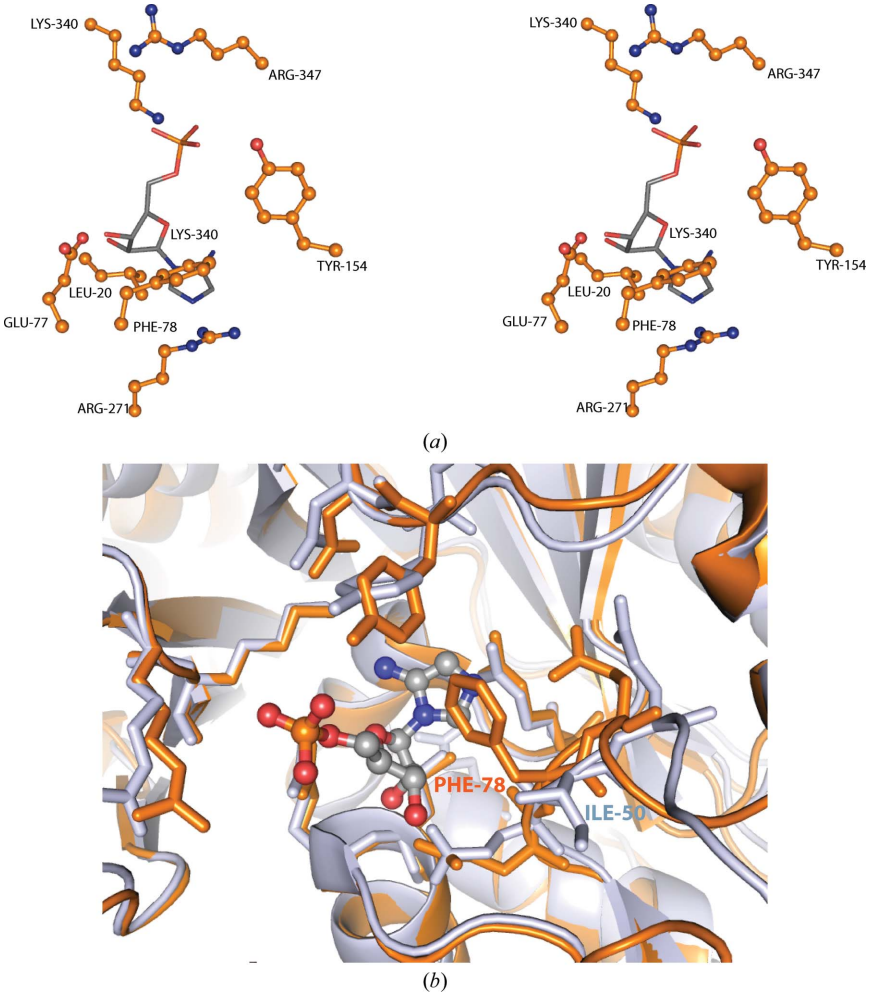


Figure 4
The active site of *S. aureus* PurK. (*a*) Stereoview of the *S. aureus* PurK AIR-binding site with AIR modeled in based on the *A. clavatus* complex structure (PDB entry 3k5i). (*b*) Overlay of the AIR-binding sites from *S. aureus* PurK (orange) and *E. coli* PurK (PDB entry 3eth; gray). AIR is modeled based on the superposition with *A. clavatus* PurK (protein not shown for clarity). Note the different orientations of Phe78 in *S. aureus* PurK and Ile50 in *E. coli* PurK. The r.m.s.d. was 0.94 Å.

Regarding the AIR-binding site, one remarkable difference exists between the structures of *S. aureus* and *E. coli* PurK. Fig. 4(*b*) shows a superposition of the AIR-binding sites of *E. coli* and *S. aureus* PurK with AIR modeled in based on superposition with the *A. clavatus* structure. Phe78 of *S. aureus* PurK is replaced by an isoleucine in *E. coli* PurK. This isoleucine, Ile50, points away from the active site in the *E. coli* structure, while Phe78 projects into the active site, taking up the space that AIR is likely to occupy when present. This different orientation of Ile50 compared with Phe78 suggests the possibility of Phe78 adopting a different conformation away from the active site to leave room for the imidazole base of AIR. In addition, the temperature factors for this residue are higher than those of the side chains of neighboring residues, indicating flexibility (average *B* values for side-chain atoms: Tyr76, 30 Å^2; Glu77, 39 Å^2; Phe78, 64 Å^2; Glu79, 56 Å^2; Asn80, 58 Å^2). We propose that *S. aureus* PurK uses Phe78 to fill the void when AIR is not present. This difference needs to be considered when searching for inhibitors based on the structure.

3.3. Enzymatic analysis of *S. aureus* PurK wild-type, F78I and F78W protein variants

Enzymatic analysis of *S. aureus* PurK was performed using a coupled assay as described in §2. In this assay the consumption of ATP by PurK is coupled to the generation of pyruvate by pyruvate kinase. The pyruvate produced is then coupled to NADH consumption by lactate dehydrogenase, which is monitored at 340 nm (Supplementary Scheme 2*a*). The steady-state kinetic parameters for this enzyme are shown in Table 2. As indicated by the Michaelis constants, the affinities of the enzyme

for AIR and ATP were 13.9 and 43.2 μ*M*, respectively. These K_m values are similar to those reported for the *E. coli* ortholog: 26 and 90 μ*M*, respectively (Firestine & Davisson, 1994). We determined the K_m for HCO_3^- to be 18.8 ± 3.9 m*M* (Supplementary Chart 1); to our knowledge, this value has never been measured before for other orthologs and is within the range of bicarbonate concentrations found in bacteria (Merlin *et al.*, 2003). Finally, this enzyme displays a catalytic efficiency (k_{cat}/K_m) of 107 μM^{-1} min^{-1}, which is similar to that reported for *E. coli* PurK.

We explored the role of Phe78 in enzyme function by site-directed mutagenesis. We speculated that Phe78 is used to fill a void in the enzyme active site when the substrate is not present and that it has to be flexible in order to leave room for AIR to bind. We tested this hypothesis by substituting this amino acid with isoleucine (to mimic other PurK enzymes) and with tryptophan (as a bulkier residue). Both of these substitutions resulted in a notable decrease in K_m and k_{cat}, as can be seen in Table 2. These effects may be explained by the fact that a bulkier residue such as tryptophan would have reduced mobility, resulting in lower catalytic efficiency, and by the fact that a smaller residue such as isoleucine would leave an empty space, resulting in a more flexible and less active enzyme. Taken together, these results show the importance of Phe78 in *S. aureus* PurK.

3.4. Structure of *S. aureus* PurE at 1.45 Å resolution

Similar to other PurE enzymes, *S. aureus* PurE forms a homooctamer in the crystalline lattice (Fig. 2*b*) and in solution as revealed by gel-filtration chromatography (Supplementary Fig. 1). The octamer has the shape of a box with approximate dimensions of 75 × 75 × 40 Å and 422 symmetry. Each monomer is composed of a main domain and a C-terminal α-helix that extends over the adjacent monomer. The main domain is composed of a five-stranded parallel β-sheet with two α-helices packing against one face and three α-helices on the opposite face. Each face of the octamer has four shallow binding sites located at the interface of three monomers. In our structure each active site is occupied by two sulfates. Fig. 5(*a*) shows an overlay of the active sites of *S. aureus* and *E. coli* PurE with CAIR modeled in based on the *E. coli* structure (Hoskins *et al.*, 2007; PDB entry 2nsj). Based on this model, Asp12 anchors CAIR through the formation of hydrogen bonds to the 2′-hydroxyl and 3′-hydroxyl groups of the ribose. The phosphate group interacts with the positively charged Arg39 as well as with Ser9 and Ser11. The imidazole base is located in a hydrophobic pocket flanked by Ala63, Ala67, Leu69 and Pro104 from the adjacent monomer. The N3 position of the imidazole ring forms a hydrogen bond to Ser36. The N5 amino group is held by hydrogen bonds to the carbonyl O atoms of Gly64 and Ala66. The carboxylic group of CAIR forms hydrogen bonds to the backbone amide N atoms of Ala37 and His68.

The organization of the active site is consistent with the proposed mechanism for this enzyme, in which His38 functions as a general acid/base catalyst that assists with the conversion of N^5-CAIR to CAIR without exchange of CO_2 with the medium (Meyer *et al.*, 1999; Hoskins *et al.*, 2007; Schaefer *et al.*, 2007).

3.5. Enzymatic analysis of *S. aureus* PurE and the role of His38 in catalysis

We performed steady-state kinetic measurements using the assay described in §2. Given the instability of N^5-CAIR ($t_{1/2}$ = 0.9 min at pH 7.8 and 303 K), the substrate of PurE for the forward reaction, we decided to use a coupled assay to enzymatically characterize *S. aureus* PurE (Supplementary Scheme 2*b*). In this assay, PurK is used to generate N^5-CAIR *in situ* and PurC is used to convert CAIR to SAICAR. The

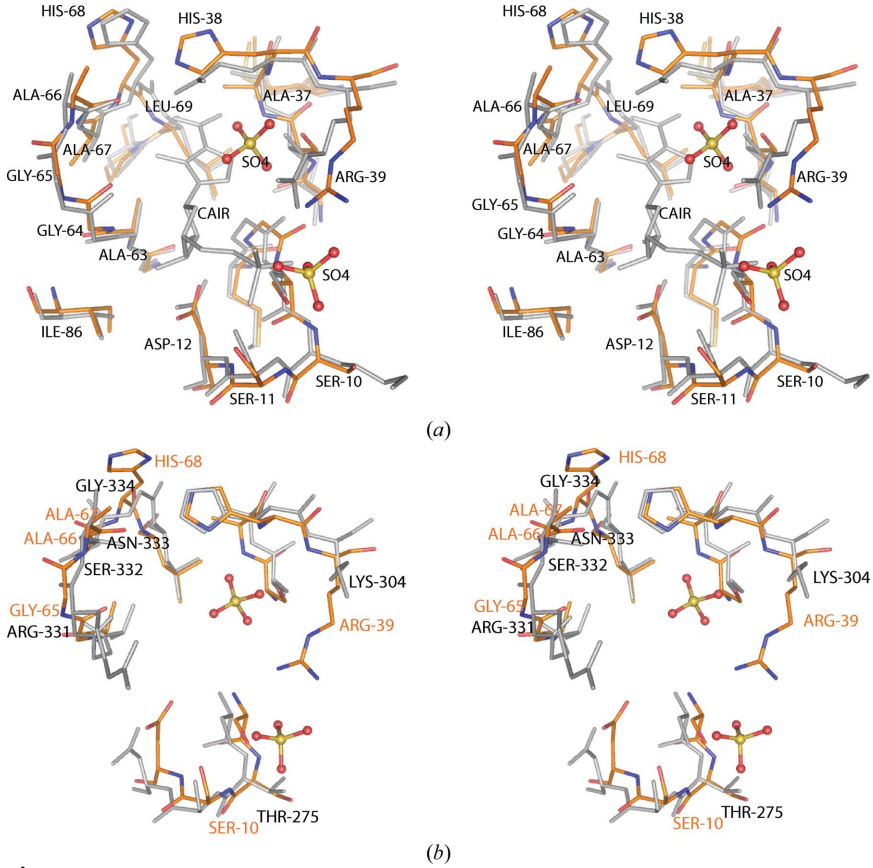


Figure 5
The active site of *S. aureus* PurE. (*a*) Overlay of the *S. aureus* PurE active site (in orange) with *E. coli* PurE containing CAIR (PDB entry 2nsj; gray; r.m.s.d. = 0.58 Å). Sulfates present in the *S. aureus* structure are shown as ball-and-stick models. Labels correspond to *S. aureus* residues. (*b*) Overlay of the *S. aureus* PurE active site (in orange) with the human PurE active site (human PAICS; PDB entry 2h31; gray; r.m.s.d. = 0.83 Å; Li *et al.*, 2007). Labels are only shown for the residues that differ between the two structures.

generation of SAICAR is then followed by monitoring the UV absorption at 282 nm ($\varepsilon_{282\,nm}$ = 8607 M^{-1} cm^{-1}). The Michaelis and catalytic constants for this protein are very similar to those reported for *E. coli* PurE (K_m = 33.1 μM for *S. aureus* PurE *versus* 22.2 μM for *E. coli* PurE and k_{cat} = 912 *versus* 954 min^{-1}).

We also confirmed the role of His38 as a general acid/base catalyst by mutating this histidine to asparagine, phenylalanine and tryptophan. All these substitutions rendered the enzyme inactive even in the presence of a hundredfold excess of enzyme compared with the wild type.

3.6. Comparison of *S. aureus* PurE and human PurE

Given the different substrate specificity of bacterial PurE and human PurE (Fig. 1), we decided to examine the differences in their active sites. Human PurE is part of a bifunctional enzyme composed of a SAICAR synthetase domain (PurC) and an AIR carboxylase domain (PurE). *S. aureus* PurE and human PurE share 28% sequence identity (51% identity between *S. aureus* and *E. coli* PurE) and an overall tertiary structure (r.m.s.d. = 1.1 Å; Supplementary Figs. 4 and 5). Fig. 5(*b*) depicts a superposition of the active sites of *S. aureus* PurE and the human bifunctional enzyme PurC–PurE. The most notable difference is that Arg39 and Gly65 in the bacterial structure are replaced by Gly304 and Arg331 in the human structure. These arginines interact with the negatively charged phosphate of the substrate from opposite sides. In addition, His68 is replaced by Gly334 in human PurE. These differences might explain why 4-nitro-5-aminoimidazole (NAIR) acts as a potent *in vitro* inhibitor of *Gallus gallus* PurE (the closest human homolog to be characterized) with a K_i of 0.34 nM, but only as a modest inhibitor of the *E. coli* enzyme (K_i = 0.5 μM) (Firestine *et al.*, 1994), which suggests the possibility that compounds with higher affinity for the bacterial protein might be developed.

4. Conclusion

Here, we have presented a comprehensive structural and enzymatic characterization of the purine-biosynthesis enzymes PurK and PurE from *S. aureus*, which we believe are promising targets for the development of selective drugs against *S. aureus*. Purine-biosynthesis enzymes play major roles in bacterial virulence (Firestine *et al.*, 2009; Lan *et al.*, 2010; Jenkins *et al.*, 2011); among them, PurK and PurE are suitable targets as PurK is absent in humans and PurE uses different substrates. The structure of *S. aureus* PurK obtained at 2.23 Å resolution supports the previously proposed catalytic mechanism; however, several differences are seen in the structure compared with the archetype *E. coli* PurK. An additional α-helix–β-strand motif is present in the A domain of *S. aureus* PurK and the AIR-binding site of *S. aureus* PurK contains an unconserved phenylalanine (isoleucine in *E. coli*) which occupies the putative position of the imidazole ring of AIR. Enzymatically, *S. aureus* PurK has similar K_m and k_{cat} values to *E. coli* PurK. Although the differences in *S. aureus* PurK do not affect the enzyme function, such differences might be useful when trying to design inhibitors for *S. aureus* PurK.

With regard to PurE, several differences exist between the *S. aureus* and human proteins, such as an arginine coming from opposite sides of the substrate, that may account for the differences in substrate specificity (Fig. 1). In addition, these structural differences may account for the fact that NAIR inhibits the human enzyme but not the bacterial enzymes, suggesting that it may be possible to develop specific inhibitors for the bacterial enzyme.

This work was funded by the Chicago Biomedical Consortium with support from The Searle Funds at the Chicago Community Trust, NIH NIAID AI074658 and a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease Award (CH). Use of the Advanced Photon Source (beamlines 19-ID and 23-ID), Argonne National Laboratory was supported by the US Department of Energy. PB thanks Fundacion Caja Madrid for a scholarship. We thank Dr J. Stubbe for the gift of AIR and Dr D. Downs for a clone of the AIRs kinase protein.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Alekshun, M. N. (2005). *Expert Opin. Investig. Drugs*, **14**, 117–134.
Appelbaum, P. C. (2006). *Clin. Microbiol. Infect.* **12**, Suppl. s2, 3–10.
Boyle, M. P., Kalliomaa, A. K., Levdikov, V., Blagova, E., Fogg, M. J., Brannigan, J. A., Wilson, K. S. & Wilkinson, A. J. (2005). *Proteins*, **61**, 674–676.
Cersini, A., Martino, M. C., Martini, I., Rossi, G. & Bernardini, M. L. (2003). *Infect. Immun.* **71**, 7002–7013.
Chung, S.-O., Lee, J.-H., Lee, S. Y. & Lee, D.-S. (1996). *FEMS Microbiol. Lett.* **137**, 265–268.
Constantine, C. Z., Starks, C. M., Mill, C. P., Ransome, A. E., Karpowicz, S. J., Francois, J. A., Goodman, R. A. & Kappock, T. J. (2006). *Biochemistry*, **45**, 8193–8208.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Deleo, F. R., Otto, M., Kreiswirth, B. N. & Chambers, H. F. (2010). *Lancet*, **375**, 1557–1568.
Donnelly, M. I., Zhou, M., Millard, C. S., Clancy, S., Stols, L., Eschenfeldt, W. H., Collart, F. R. & Joachimiak, A. (2006). *Protein Expr. Purif.* **47**, 446–454.
Donovan, M., Schumuke, J. J., Fonzi, W. A., Bonar, S. L., Gheesling-Mullis, K., Jacob, G. S., Davisson, V. J. & Dotson, S. B. (2001). *Infect. Immun.* **69**, 2542–2548.
Drazek, E. S., Houng, H.-S., Crawford, R. M., Hadfield, T. L., Hoover, D. L. & Warren, R. L. (1995). *Infect. Immun.* **63**, 3297–3301.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Firestine, S. M. & Davisson, V. J. (1994). *Biochemistry*, **33**, 11917–11926.
Firestine, S. M., Paritala, H., McDonnell, J. E., Thoden, J. B. & Holden, H. M. (2009). *Bioorg. Med. Chem.* **17**, 3317–3323.
Firestine, S. M., Poon, S. W., Mueller, E. J., Stubbe, J. & Davisson, V. J. (1994). *Biochemistry*, **33**, 11927–11934.
Fischbach, M. A. & Walsh, C. T. (2009). *Science*, **325**, 1089–1093.
Hoskins, A. A., Morar, M., Kappock, T. J., Mathews, I. I., Zaugg, J. B., Barder, T. E., Peng, P., Okamoto, A., Ealick, S. E. & Stubbe, J. (2007). *Biochemistry*, **46**, 2842–2855.
Jenkins, A., Cote, C., Twenhafel, N., Merkel, T., Bozue, J. & Welkos, S. (2011). *Infect. Immun.* **79**, 153–166.

Kirsch, D. R. & Whitney, R. R. (1991). *Infect. Immun.* **59**, 3297–3300.
Lan, L., Cheng, A., Dunman, P. M., Missiakas, D. & He, C. (2010). *J. Bacteriol.* **192**, 3068–3077.
Li, S.-X., Tong, Y.-P., Xie, X.-C., Wang, Q.-H., Zhou, H.-N., Han, Y., Zhang, Z.-Y., Gao, W., Li, S.-G., Zhang, X.-C. & Bi, R.-C. (2007). *J. Mol. Biol.* **366**, 1603–1614.
Lodise, T. P. & McKinnon, P. S. (2007). *Pharmacotherapy*, **27**, 1001–1012.
Lowy, F. D. (1998). *N. Engl. J. Med.* **339**, 520–532.
Mathews, I. I., Kappock, T. J., Stubbe, J. & Ealick, S. E. (1999). *Structure*, **7**, 1395–1406.
McFarland, W. C. & Stocker, B. A. (1987). *Microb. Pathog.* **3**, 129–141.
Merlin, C., Masters, M., McAteer, S. & Coulson, A. (2003). *J. Bacteriol.* **185**, 6415–6424.
Meyer, E., Kappock, T. J., Osuji, C. & Stubbe, J. (1999). *Biochemistry*, **38**, 3012–3018.
Meyer, E., Leonard, N. J., Bhat, B., Stubbe, J. & Smith, J. M. (1992). *Biochemistry*, **31**, 5022–5032.
Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J. & Stubbe, J. (1994). *Biochemistry*, **33**, 2269–2278.
Noskin, G. A., Rubin, R. J., Schentag, J. J., Kluytmans, J., Hedblom, E. C., Smulders, M., Lapetina, E. & Gemmen, E. (2005). *Arch. Intern. Med.* **165**, 1756–1761.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perfect, J. R., Toffaletti, D. L. & Rude, T. H. (1993). *Infect. Immun.* **61**, 4446–4451.
Polissi, A., Pontiggia, A., Feger, G., Altieri, M., Mottl, H., Ferrari, L. & Simon, D. (1998). *Infect. Immun.* **66**, 5620–5629.
Samant, S., Lee, H., Ghassemi, M., Chen, J., Cook, J. L., Mankin, A. S. & Neyfakh, A. A. (2008). *PLoS Pathog.* **4**, e37.
Schaefer, J., Jiang, H., Ransome, A. E. & Kappock, T. J. (2007). *Biochemistry*, **46**, 9507–9512.
Schmuke, J. J., Davisson, V. J., Bonar, S. L., Gheesling Mullis, K. & Dotson, S. B. (1997). *Yeast*, **13**, 769–776.
Settembre, E. C., Chittuluru, J. R., Mill, C. P., Kappock, T. J. & Ealick, S. E. (2004). *Acta Cryst.* D**60**, 1753–1760.
Sørensen, I. S. & Dandanell, G. (1997). *FEMS Microbiol. Lett.* **154**, 173–180.
Thoden, J. B., Holden, H. M. & Firestine, S. M. (2008). *Biochemistry*, **47**, 13346–13353.
Thoden, J. B., Holden, H. M., Paritala, H. & Firestine, S. M. (2010). *Biochemistry*, **49**, 752–760.
Thoden, J. B., Kappock, T. J., Stubbe, J. & Holden, H. M. (1999). *Biochemistry*, **38**, 15480–15492.
Tiedeman, A. A., Keyhani, J., Kamholz, J., Daum, H. A. III, Gots, J. S. & Smith, J. M. (1989). *J. Bacteriol.* **171**, 205–212.
Wang, J., Mushegian, A., Lory, S. & Jin, S. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 10434–10439.
Watanabe, W., Sampei, G., Aiba, A. & Mizobuchi, K. (1989). *J. Bacteriol.* **171**, 198–204.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.