BIOLOGICAL CRYSTALLOGRAPHY
ISSN 1399-0047

Structure and inhibition of subunit I of the anthranilate synthase complex of *Mycobacterium tuberculosis* and expression of the active complex

Ghader Bashiri,[a,b]‡ Jodie M. Johnston,[a,b]‡ Genevieve L. Evans,[a,b]‡§ Esther M. M. Bulloch,[a,b] David C. Goldstone,[a,b] Ehab N. M. Jirgis,[a] Silke Kleinboelting,[a,b]¶ Alina Castell,[a,b]‡‡ Rochelle J. Ramsay,[a,b] Alexandra Manos-Turvey,[c] Richard J. Payne,[c] J. Shaun Lott[a,b]* and Edward N. Baker[a,b]*

[a]School of Biological Sciences, University of Auckland, 3a Symonds Steet, Private Bag 90210, Auckland 1142, New Zealand, [b]Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland, 3a Symonds Steet, Private Bag 90210, Auckland 1142, New Zealand, and [c]School of Chemistry, The University of Sydney, Sydney, NSW 2006, Australia. *Correspondence e-mail: s.lott@auckland.ac.nz, ted.baker@auckland.ac.nz

Received 30 July 2015
Accepted 14 September 2015

Edited by Z. S. Derewenda, University of Virginia, USA

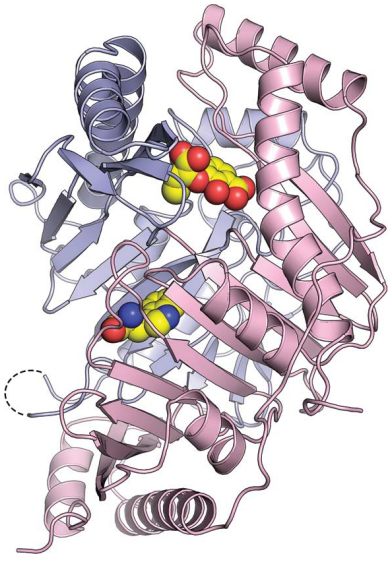
‡ GB, JMJ and GLE co-wrote this manuscript, contributed equally to this work and are co-first authors.
§ Current address: School of Chemistry and Biochemistry, The University of Western Australia, Crawley, WA 6009, Australia.
¶ Current address: Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305-5489, USA.
‡‡ Current address: Karolinska Institutet, Department of Microbiology, Tumor and Cell Biology, 171 77 Stockholm, Sweden.

Keywords: *Mycobacterium tuberculosis*; tryptophan biosynthesis; anthranilate synthase complex; inhibitor design.

PDB reference: anthranilate synthase subunit I, 5cwa

Supporting information: this article has supporting information at journals.iucr.org/d

The tryptophan-biosynthesis pathway is essential for *Mycobacterium tuberculosis* (*Mtb*) to cause disease, but not all of the enzymes that catalyse this pathway in this organism have been identified. The structure and function of the enzyme complex that catalyses the first committed step in the pathway, the anthranilate synthase (AS) complex, have been analysed. It is shown that the open reading frames Rv1609 (*trpE*) and Rv0013 (*trpG*) encode the chorismate-utilizing (AS-I) and glutamine amidotransferase (AS-II) subunits of the AS complex, respectively. Biochemical assays show that when these subunits are co-expressed a bifunctional AS complex is obtained. Crystallization trials on *Mtb*-AS unexpectedly gave crystals containing only AS-I, presumably owing to its selective crystallization from solutions containing a mixture of the AS complex and free AS-I. The three-dimensional structure reveals that *Mtb*-AS-I dimerizes *via* an interface that has not previously been seen in AS complexes. As is the case in other bacteria, it is demonstrated that *Mtb*-AS shows cooperative allosteric inhibition by tryptophan, which can be rationalized based on interactions at this interface. Comparative inhibition studies on *Mtb*-AS-I and related enzymes highlight the potential for single inhibitory compounds to target multiple chorismate-utilizing enzymes for TB drug discovery.



1. Introduction

Tuberculosis (TB) remains one of the world's deadliest bacterial diseases, killing more people annually than any other infectious disease (Zumla *et al.*, 2013). The emergence of multidrug-resistant (MDR) and extensively drug-resistant (XDR) strains of *Mycobacterium tuberculosis* (*Mtb*), the causative agent of TB, together with complications from co-infection with HIV/AIDS, demands the urgent development of new and effective anti-TB agents that have novel modes of action (Zumla *et al.*, 2013). Amino-acid biosynthesis pathways are attractive targets where pathogenic bacteria have significantly reduced access to amino acids, or their intermediary metabolites, from their human host, as is the case for pathogens with an intracellular lifestyle such as *Mtb* (Zhang & Rubin, 2013). While in the macrophage phageosome, *Mtb* has significantly restricted access to many nutrients, including key amino acids such as tryptophan and methionine (Zhang & Rubin, 2013; Berney *et al.*, 2015)

The tryptophan (Trp)-biosynthesis pathway is one such target in *Mtb*. Strains of *Mtb* that are auxotrophic for Trp show reduced intracellular survival in murine and human macrophages *in vitro* and are avirulent in immunocompetent and immunocompromised mice. This demonstrates that although Trp biosynthesis is not essential for *in vitro* growth, it is conditionally essential for the establishment and maintenance of infection (Smith *et al.*, 2001; Zhang *et al.*, 2013). This, combined with the lack of Trp biosynthesis in mammals, makes the pathway a tractable target for new anti-TB therapies. The Trp-biosynthesis pathway is conserved in mycobacterial species from the nonpathogenic *M. smegmatis* to the pathogens *Mtb* and *M. leprae*. Significantly, *M. leprae* has retained the Trp-biosynthetic pathway (Xie *et al.*, 2003) despite possessing a minimal mycobacterial genome (Cole *et al.*, 2001). These factors point to the importance and essentiality of the Trp-biosynthesis pathway in mycobacterial species and its potential as a target for new TB drugs. This is emphasized by the recent observation that 6-fluoroanthranilate, an orally effective antimycobacterial compound in a mouse model of TB, targets the Trp-biosynthetic pathway (Zhang *et al.*, 2013).

Trp is synthesized from chorismate, the final product of the shikimate pathway (Kerbarh, Bulloch *et al.*, 2005). Chorismate is also a common precursor for other aromatic amino acids and metabolites such as folic acid, salicylic acid, menaquinone (vitamin K_2) and ubiquinone (coenzyme Q). In *Escherichia coli*, the Trp-biosynthetic pathway requires seven enzymes, the genes for which are organized as a whole-pathway operon (*trpEGDFCBA*; Xie *et al.*, 2003). In contrast, in mycobacteria the Trp-biosynthesis genes are organized as a split-pathway operon (Xie *et al.*, 2003). In *Mtb* H37Rv, *trpE* (Rv1609), *trpC* (Rv1611), *trpB* (Rv1612) and *trpA* (Rv1613) form a partial-pathway operon, with the *trpF* gene (Rv1603) just upstream as part of the neighbouring *his* operon (Barona-Gómez & Hodgson, 2003; Due *et al.*, 2011). The *trpD* (Rv2192c) gene is located remotely (Lee *et al.*, 2006). The *trpG* gene, which in other organisms encodes the amidotransferase component of anthranilate synthase (AS), is the only gene from the Trp-biosynthetic pathway that has not been characterized in *Mtb* to date.

AS catalyses the first committed reaction in the Trp-biosynthetic pathway. It comprises two distinct functional components: AS-I (TrpE) and AS-II (TrpG). AS-I catalyses the production of anthranilate from chorismate and ammonia through a 2-amino-2-deoxyisochorismate (ADIC) intermediate (Morollo *et al.*, 1993), whereas the glutamine amidotransferase (GAT) activity of AS-II provides ammonia from glutamine (Mouilleron & Golinelli-Pimpaneau, 2007; Fig. 1*a*). Depending on the organism, AS-II either forms hetero-oligomers with AS-I (*i.e.* dimers or tetramers) or exists as a single fused polypeptide with AS-I (Ashenafi *et al.*, 2008; Romero *et al.*, 1995). The AS complex is key to the regulation of Trp biosynthesis, with most of the characterized AS enzymes being allosterically inhibited by the end product of the pathway, Trp (Romero *et al.*, 1995).

To date, all structurally characterized AS proteins are from microorganisms in which the genes for AS-I and AS-II are adjacent to each other in whole-pathway *trp* operons. Structures have been determined for AS from the thermophile *Sulfolobus solfataricus* (*Sso*-AS; Knöchel *et al.*, 1999) and from two mesophiles: *Serratia marcescens* (*Sma*-AS; Spraggon *et al.*, 2001) and *Salmonella typhimurium* (*Sty*-AS; Morollo & Eck, 2001). Although these AS complexes have the same 2:2 subunit stoichiometries, their oligomeric associations are different. The mesophilic enzymes (*Sma*-AS and *Sty*-AS) associate primarily through dimerization of AS-I (Spraggon *et al.*, 2001; Morollo & Eck, 2001), whereas in their thermophilic counterpart (*Sso*-AS) tetramer formation depends almost completely on association of the two AS-II subunits (Knöchel *et al.*, 1999). The reasons for the radically different quaternary structures in different bacteria are not known.

Here, we identify the *Mtb*-AS-II protein as being encoded by open reading frame (ORF) Rv0013 and show that it has GAT activity and that its complex with AS-I converts chorismate to anthranilate in the presence of glutamine. We have also examined the potency of inhibitors developed for another chorismate-utilizing enzyme, salicylate synthase (SS), including several with a 3-(1-carboxyethenyloxy)-2-hydroxybenzoic acid scaffold (AMT series; Fig. 1*b*; Manos-Turvey *et al.*, 2010) and have determined the crystal structure of *Mtb*-AS-I in the presence of the most potent inhibitor, methyl-AMT. Our results point to new opportunities for structure-based inhibitor design targeting this important metabolic pathway for TB drug lead discovery.

2. Materials and methods

2.1. PCR amplification and cloning

The ORF Rv1609 encoding AS-I (TrpE) was amplified from *M. tuberculosis* H37Rv genomic DNA and cloned into the pProEX-HTb vector (Invitrogen). The resulting construct, pProEX-TrpE, expresses TrpE protein with an N-terminal His_6 tag that is cleavable using *Tobacco etch virus* (TEV) protease. The ORFs Rv0013, Rv0788, Rv1602 and Rv2604c, encoding putative AS-II (TrpG) proteins, were each amplified similarly and cloned into the pYUBDuet vector (Bashiri *et al.*, 2010) for individual expression. These constructs express the corresponding TrpG proteins with a noncleavable N-terminal His_6 tag.

For AS-I/AS-II co-expression, the *trpE* gene was cloned into the pYUBDuet-AS-II constructs. The resulting pYUBDuet-AS constructs co-express AS-I and AS-II under the control of two separate T7 promoters. The AS-I protein has no tag, allowing the His_6-tagged AS-II protein to be used as bait for AS-I during purification. For Rv0013, this co-expression consistently failed to produce viable cells, however, so the co-expression was instead carried out with His_6-tagged AS-I and untagged AS-II. To this end, the ORF Rv0013 was cloned into the pYUBDuet vector (Bashiri *et al.*, 2010) and *trpE* was subsequently cloned into the pYUBDuet-Rv0013 construct. In this construct, the N-terminal His_6 tag on the TrpE protein is cleavable using TEV protease. All cloning steps for the pYUBDuet-AS constructs were carried out using *E. coli*

Top10 (Invitrogen) electrocompetent cells. Positive clones were selected on low-salt LB agar medium supplemented with 50 µg ml^{-1} hygromycin B and were then verified using colony PCR, restriction digestion and sequencing.

2.2. Protein expression and purification

For the expression of AS-I (TrpE) alone, *E. coli* BL21 (DE3) cells were transformed with the pProEX-TrpE construct and a resulting positive colony from LB agar plates was grown as a starter culture in 100 ml non-inducing MDG medium overnight before being used to inoculate 10 l of autoinduction ZYM-5052 medium (Studier, 2005) for expression. All media were supplemented with ampicillin at 100 µg ml^{-1}. Expression was carried out at 37°C for 3 h, followed by incubation at 18°C overnight. For expression of the individual AS-II protein candidates and for co-expression of *Mtb*-AS constructs, *M. smegmatis* mc^{2}4517 cells (Wang *et al.*, 2010) were transformed with the appropriate pYUBDuet constructs and the proteins were expressed using autoinduction protocols (Bashiri *et al.*, 2010, Studier, 2005) over 4 d at 37°C. The identities of all expressed proteins were confirmed by mass spectrometry.

All proteins were purified using similar procedures. The cells were harvested, resuspended in lysis buffer and lysed in a cell disrupter (Microfluidizer M-110P). After centrifugation at 20 000*g* to separate insoluble material, the recombinant proteins were purified using Ni–NTA immobilized metal-ion chromatography (IMAC; 5 ml, HiTrap column), followed by anion-exchange chromatography (5 ml, QFF column) and size-exclusion chromatography (SEC; Superdex S200 10/300 column) steps. The His_6 tag was removed with rTEV protease (Blommel & Fox, 2007). The final (storage) buffer for AS-I was 50 m*M* HEPES pH 7.8, 150 m*M* NaCl, 5 m*M* $MgCl_2$, 2.5%(*v*/*v*) glycerol, 1 m*M* TCEP and that for AS was 20 m*M* HEPES pH 8.0, 10 m*M* EDTA, 1 m*M* TCEP. For biochemical assays of *Mtb*-AS, the purification buffers consisted of 150 m*M* NaCl, 50 m*M* glutamine, 2.5 m*M* TCEP, 5 m*M* EDTA. Purification involved an IMAC step followed by SEC on the same day. *Mtb*-AS was stored in 20 m*M* HEPES pH 8.0, 150 m*M* NaCl, 50 m*M* glutamine, 2.5 m*M* TCEP, 5 m*M* EDTA, 5%(*v*/*v*) glycerol at −80°C. The inclusion of glutamine in all buffers resulted in improved enzyme activity.

2.3. Crystallization and data collection

Initial crystallization conditions for both *Mtb*-AS-I and *Mtb*-AS were obtained by sitting-drop vapour diffusion using a Cartesian nanolitre dispensing robot (Genome Solutions) and in-house crystallization screens (Moreland *et al.*, 2005). Optimization by hanging-drop vapour diffusion at 18°C gave reproducible crystals in both cases. For *Mtb*-AS-I the best crystals were obtained by mixing 1 µl protein solution (15–23.5 mg ml^{-1}, 1 m*M* methyl-AMT) with 1 µl precipitant solution (150 m*M* trisodium citrate pH 5.5, 0.75–1.5 *M* ammonium sulfate, 0.25–0.5 *M* lithium sulfate). Fine screens were also undertaken, without success, with ethyl-AMT and phenyl-AMT (at 1 m*M*; Manos-Turvey *et al.*, 2010). For *Mtb*-AS, crystals that diffracted to about 3 Å resolution were obtained in 2–3 d by mixing 1 µl protein solution (20 mg ml^{-1}, 1 m*M* methyl-AMT) with 1 µl precipitant solution (0.6–1.0 *M* ammonium sulfate, 0.1 *M* bis-tris propane pH 7.0). These

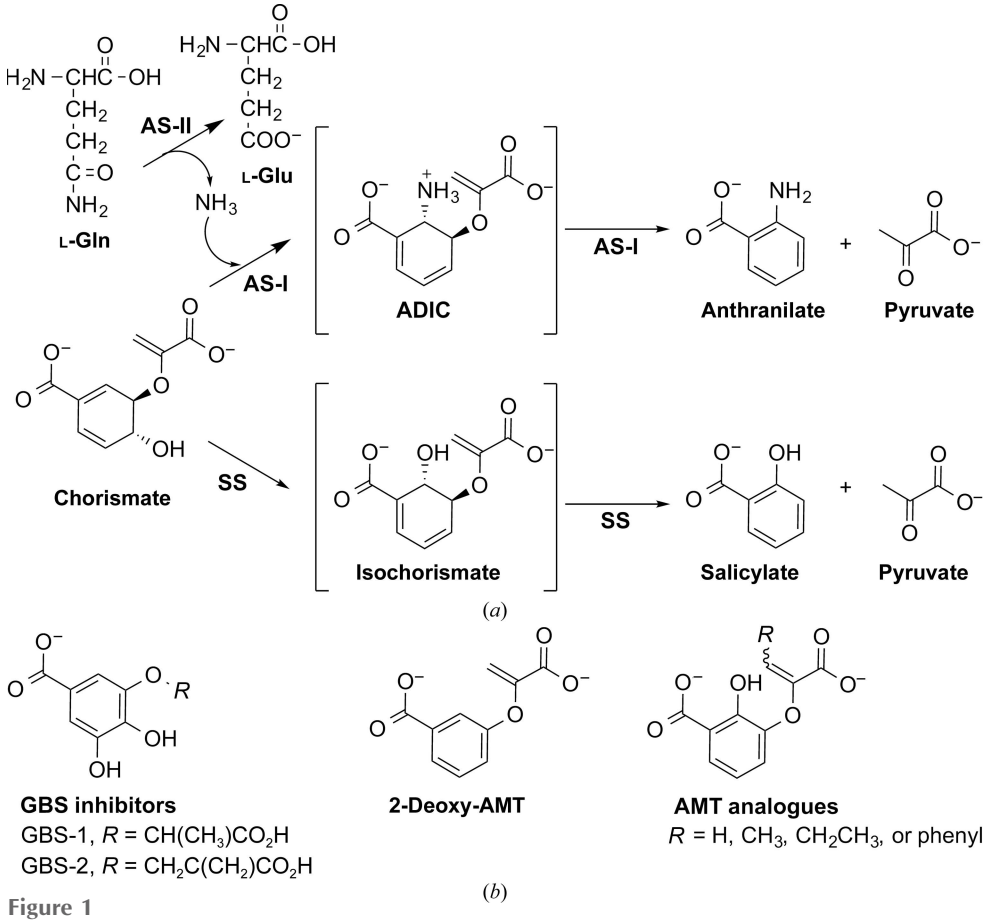


Figure 1
Molecular structures for the enzyme-catalysed reactions and inhibitor compounds. (*a*) The reaction catalysed by anthranilate synthase (AS) involves the production of ammonia from glutamine by AS component II (AS-II), followed by its utilization by AS component I (AS-I) in the formation of anthranilate from chorismate. The reactions catalysed by AS-I and salicylate synthase (SS) involve the same chorismate substrate, similar enzyme-bound intermediates (in parentheses) and products. (*b*) Potential inhibitor compounds tested against *Mtb*-AS in this study.

Table 1
Data collection and processing.

Values in parentheses are for the outermost shell.

Wavelength (Å)	0.8983
Space group	$P6_422$
Unit-cell parameters (Å, °)	$a = b = 156.41$, $c = 128.38$, $\alpha = \beta = 90$, $\gamma = 120$
Matthews coefficient (Å^3 Da^{-1})	4.06
Solvent content (%)	69.7
Resolution (Å)	135.45–2.10 (2.21–2.10)
$R_{\rm p.i.m.}$	0.034 (0.81)
Unique reflections	54312 (7768)
Observed reflections	1189598
Mean $I/\sigma(I)$	14.8 (1.1)
Multiplicity	21.9 (22.4)
Completeness (%)	100 (99.8)
CC_I mean (from *SCALA*)	0.99 (0.44)
Wilson *B* factor (Å^2)	48.3

Table 2
Refinement statistics.

Values in parentheses are for the outermost shell.

PDB code	5cwa
Resolution range (Å)	135.5–2.10 (2.16–2.10)
No. of reflections (working/test)	51515/2767
$R_{\rm work}/R_{\rm free}$ (%)	20.07/25.34 (36.1/36.6)
No. of atoms (non-H atoms)	
Protein	3861
Ligand	54
Water	205
R.m.s. deviations from ideality	
Bonds (Å)	0.012
Angles (°)	1.5
Average *B* factors (Å^2)	
Protein atoms	57.3
Methyl-AMT	52.3
Water molecules	57.6
Other solvent molecules (2 glycerol, 2 sulfate ions)	69.1
Ramachandran most favoured/outliers (%)	96.84/0.00
Cruickshank's DPI (Å)	0.152 [$R_{\rm work}$]
MolProbity score; percentile	1.76; 92nd

conditions were then optimized, with the best crystals being obtained using additives from the Hampton Research Silver Bullets Bio screen condition G6 (fumaric acid, *cis*-aconitic acid, DL-isocitric acid trisodium salt, oxalacetic acid, sodium pyruvate and HEPES pH 6.8).

The *Mtb*-AS-I crystals were cryoprotected with increased concentrations of ammonium and lithium sulfate (1.0 and 1.5 *M*, respectively) and 2% ethylene glycol, and were flash-cooled in liquid nitrogen. These crystals were hexagonal, space group $P6_422$, with unit-cell parameters $a = b = 157.2$, $c = 127.6$ Å. Diffraction data were collected to 2.6 Å resolution on a Rigaku MicroMax-007 HF microfocus rotating-anode generator equipped with Osmic optics and a MAR345 image plate and were indexed and processed using *XDS* (Kabsch, 2010). The crystals grown from *Mtb*-AS solutions were cryoprotected in 30% glycerol and flash-cooled directly in liquid nitrogen. These proved to be essentially isomorphous to the *Mtb*-AS-I crystals, with space group $P6_422$ and unit-cell parameters $a = b = 156.4$, $c = 128.4$ Å compared with $a = b = 157.2$, $c = 127.6$ Å for the *Mtb*-AS-I crystals. Diffraction data were collected from these crystals on the Australian Synchrotron MX2 beamline using an ADSC Quantum 315r CCD detector. These data were indexed and processed with *XDS* (Kabsch, 2010), reindexed with *POINTLESS* (Evans, 2006) and scaled with *SCALA* (Evans, 2006) from the *CCP*4 program suite (Winn *et al.*, 2011). An appropriate diffraction limit of 2.1 Å was chosen using the $CC_{1/2}$ analysis of Karplus & Diederichs (2012). Data-collection statistics for this data set are given in Table 1.

2.4. Structure determination

An initial structural model for *Mtb*-AS-I was determined by molecular replacement with *Phaser* (McCoy *et al.*, 2007), using a search model derived from the *Sso*-AS structure (PDB entry 1qdl; Knöchel *et al.*, 1999). This structure was refined at 2.6 Å resolution using *REFMAC*5 (Murshudov *et al.*, 2011) to R = 24% and $R_{\rm free}$ = 28%. This model was in turn used as a search model to solve the structure of the *Mtb*-AS crystals by molecular replacement with *Phaser*. The AS-I component was readily placed, but no solution for the AS-II component could be found. The structure was refined at 2.1 Å resolution, starting with automated model building with *ARP*/*wARP* (Langer *et al.*, 2008) and followed by rounds of manual model building in *Coot* (Emsley, 2013) and refinement in *REFMAC*5 (Murshudov *et al.*, 2011) and *BUSTER* (v.2.11.5; Global Phasing Ltd, Cambridge, England). Additional electron density in the active site was modelled as a molecule of methyl-AMT, for which geometric restraints were generated in *Coot* from the inbuilt ligand builder *LIDIA* (Emsley, 2013). Solvent molecules were added based on appropriate shape and hydrogen-bond interactions. Full refinement statistics are given in Table 2, and the atomic coordinates and structure-factor amplitudes have been deposited in the Protein Data Bank as entry 5cwa. All structural figures were produced with *PyMOL* (Schrödinger, New York, USA).

2.5. Small-angle X-ray scattering (SAXS)

Protein samples for SAXS experiments were extensively dialyzed against 20 m*M* HEPES pH 8.0, 1 m*M* EDTA, 1 m*M* DTT. This buffer was also used as the buffer control and to dilute protein samples. Scattering curves were collected at concentrations of between 0.045 and 3 mg ml^{-1}. The SAXS experiments were performed on the SAXS beamline at the Australian Synchrotron using a PILATUS 1M detector at 1.6 m camera length, resulting in a q-range of 0.0007–0.0341 Å^{-1}. Initial buffer subtraction and data averaging were performed using *ScatterBrain* (v.1.71; Australian Synchrotron). Scattering curves were normalized against the scattering of water and the molecular mass was estimated from the scattering contrast (Mylonas & Svergun, 2007). The processed data were then analyzed using the *ATSAS* package (v.2.5). The radius of gyration ($R_{\rm g}$) was calculated by the Guinier approximation in *PRIMUS* (Konarev *et al.*, 2003). Pair distribution functions were calculated using *GNOM* (Svergun, 1992). SAXS experimental scattering curves were compared with theoretical curves calculated from AS-I crystal structures using *CRYSOL* (Svergun *et al.*, 1995).

2.6. SEC with multi-angle laser light scattering (SEC-MALS)

Protein samples (AS-I or AS with His_6 tags removed) were loaded onto an S200 10/300 column using an Ultimate 3000 HPLC with inline PSS SLD7000 MALS detector and Shodex RI-101 differential refractive detector. The system was pre-equilibrated in 20 m*M* HEPES pH 8.0, 1 m*M* EDTA, 1 m*M* DTT. Samples of 100 µl were loaded at concentrations of between 0.5 and 6 mg ml^{-1} and were eluted at a flow rate of 0.5 ml min^{-1}. The data were processed using the *WinGPC UniChrom* software package and a d*n*/d*c* value of 0.186 ml g^{-1}.

2.7. Steady-state kinetic and enzyme-inhibition data

AS activity was followed fluorimetrically using a 2104 EnVision Multilabel Reader (PerkinElmer) to measure the rate of appearance of anthranilate (Tamir & Srinivasan, 1970) in black 96-well microplates (Greiner) with $\lambda_{excitation}$ of 320 nm and $\lambda_{emission}$ of 460 nm. Plots of the fluorescence *versus* the anthranilate concentration were linear from 0.5 to 30 µ*M* (standard deviation of 1–5% for four replicates). Anthranilate was quantified using a coupled assay (Evans *et al.*, 2014). Chorismate concentrations were determined from the initial and final fluorescence in assays (in quadruplicate) where chorismate was the limiting reagent. The protein concentration was determined in triplicate using a Cary 100 UV–Vis spectrophotometer and an ε_{280} of 1.0195 ml mg^{-1} cm^{-1}.

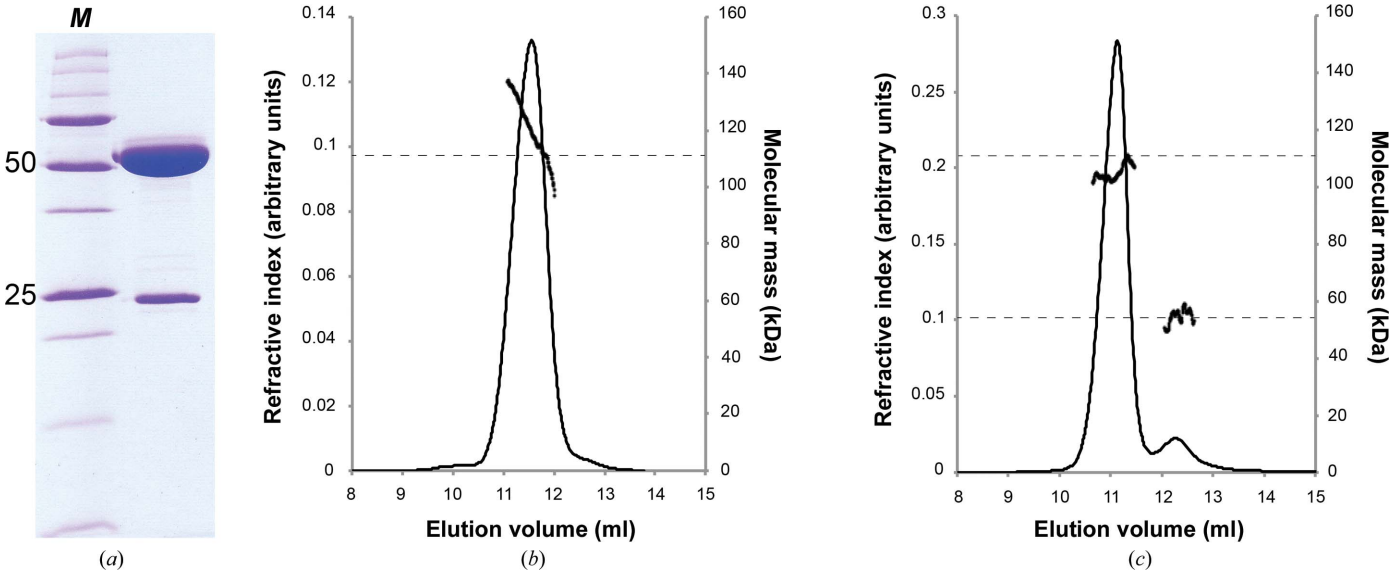
Glutamine-dependent assays for determination of the apparent Michaelis–Menten constant (K_m^{app}) consisted of 112 m*M* HEPES pH 8.0, 10 m*M* $MgCl_2$, 0.056 mg ml^{-1} *Mtb*-AS and varying concentrations of chorismate (3.7–118 µ*M*) in 375 µl with a constant glutamine concentration (12 m*M*). The K_m^{app} value for chorismate was obtained by nonlinear fitting to the Michaelis–Menten equation using *GraphPad Prism* (v.5.02; GraphPad Software, La Jolla, California, USA).

For determination of the IC_{50}, the glutamine-dependent assays were the same but with *Mtb*-AS and chorismate at concentrations of 0.075 mg ml^{-1} and 29 µ*M*, respectively, in 300 µl. All components except for chorismate and *Mtb*-AS were thermally equilibrated at 25°C for 5–15 min and the reaction was initiated by the addition of chorismate. Inhibitors were dissolved in DMSO and their concentrations were determined by ^{1}H-NMR. *In situ* twofold serial dilutions (with thorough mixing in each well) gave final inhibitor concentrations from 0.08 to 4600 µ*M*. For control reactions, DMSO was added without inhibitor over the same dilution range and had no measurable effect on enzyme activity. The IC_{50} values were determined as described in Evans *et al.* (2014), but with the max and min variables constrained to be 1 and 0, respectively. The Hill coefficient was unconstrained for methyl-AMT (Manos-Turvey *et al.*, 2010) and Trp data sets and was taken as 1 for the other inhibitors, assuming competitive inhibition. The assumption that the inhibitors are competitive is consistent with the structure of *Mtb*-AS-I, which shows the inhibitor bound in the active site. Furthermore, steady-state kinetic analysis of the AMT-series inhibitors against *Sma*-AS showed them to be competitive (Manos-Turvey *et al.*, 2010).

3. Results and discussion

3.1. Identification of the *trpG* ORF

Nine ORFs with predicted GAT activity are annotated in the *Mtb* genome. Of these, five are fused genes that encode



(*a*) (*b*) (*c*)

Figure 2
Analyses of *Mtb*-AS and its subunits. (*a*) SDS–PAGE gel of the purified *Mtb*-AS complex showing bands at ∼54 kDa for the AS-I subunit and ∼23 kDa for AS-II. Lane *M* contains molecular-mass markers (labelled in kDa). (*b*) SEC trace of the *Mtb*-AS complex, showing refractive-index (continuous line) and weight-averaged molecular-mass (darker line) data. The calculated molecular mass is substantially different across the elution peak, indicated by the sharp slope of the molecular-mass trace across the peak. (*c*) SEC trace of *Mtb*-AS-I alone, shown for comparison and plotted as for (*b*). The protein runs predominantly as an AS-I–AS-I dimer. The estimated molecular mass for the small monomer population is also shown. In both (*b*) and (*c*) data are shown for one concentration of protein for clarity.

large multidomain proteins, which are unlikely to be the AS-II component. Sequence comparisons of the four remaining GAT gene candidates (Rv0013, Rv0788, Rv1602 and Rv2604c) with two *trpG* genes from unrelated species (*Sma*-AS-II and *Sso*-AS-II) showed the conservation of key regions, including residues in the catalytic triad and oxyanion hole, indicative of class I GATs (Mouilleron & Golinelli-Pimpaneau, 2007; Raushel *et al.*, 1999). Of these four candidates, Rv0013 shares 38–40% sequence identity with *Sma*-AS-II and *Sso*-AS-II, compared with ∼18% for the next closest candidate (Rv1602).

Rv0788, Rv1602 and Rv2604c were expressed as individual proteins in soluble form, but showed no evidence of complex formation with AS-I. Likewise, co-expression of these proteins with AS-I in *M. smegmatis* host cells gave no AS complex formation. Attempts to express and purify Rv0013 on its own were unsuccessful, but a soluble putative AS complex was obtained when it was co-expressed with *Mtb*-AS-I. This complex was then purified by IMAC using the His_6 tag on AS-I; the untagged Rv0013 co-purified with AS-I (Fig. 2*a*), suggesting stable complex formation. Further purification by SEC resulted in a single peak, and mass spectrometry confirmed the presence and identity of both AS-I and Rv0013 proteins. From these experiments and its strong sequence similarity to other AS-II proteins, we identify Rv0013 as the AS-II protein in *Mtb*.

SEC-MALS analysis of the complex (Fig. 2*b*) indicated a range of M_r values across the peak, covering M_r values from ∼108 to 140 kDa rather than a single defined species. In contrast, the same analysis for AS-I alone (Fig. 2*c*) showed a much sharper peak corresponding to a dimer M_r of 108 kDa. The trace for the AS complex could thus encompass species ranging from the AS-I dimer (108 kDa) to complexes with one or two AS-II subunits (25 kDa each) bound to an AS-I dimer. We conclude that the AS complex sample also contains some uncomplexed AS-I present in solution. The AS complex is evidently stable, however, since no dissociation of AS-II was observed with SEC, with added salt (up to 1.0 *M*) or glycerol (up to 5%) or when the sample injected contained the methyl-AMT inhibitor (1 m*M*).

3.2. Functional characterization of the AS complex

To further characterize the *Mtb*-AS complex, its ability to catalyse the conversion of chorismate to anthranilate was tested. The complex catalysed the conversion of chorismate to anthranilate using glutamine as a nitrogen source, demonstrating that Rv0013 has GAT activity and that its complex with *Mtb*-AS-I is the authentic *Mtb*-AS complex. In contrast, the purified *Mtb*-AS-I on its own has no GAT activity and catalysed this reaction only in the presence of added ammonia. The glutamine-dependent activity of the *Mtb*-AS complex was found to be proportional to the amount of the enzyme complex, indicating that the oligomeric state did not change, or affect enzyme activity, over the concentration range used: 0.0044–0.087 mg ml^{-1} (data not shown).

The apparent Michaelis–Menten constant (K_m^{app}) of *Mtb*-AS for chorismate was determined to be 15 ± 2 μ*M* (Fig. 3*a*). This value is larger than the previously determined value for *Mtb*-AS-I of 7 μ*M* (Lin *et al.*, 2009) and the values for *Sma*-AS and

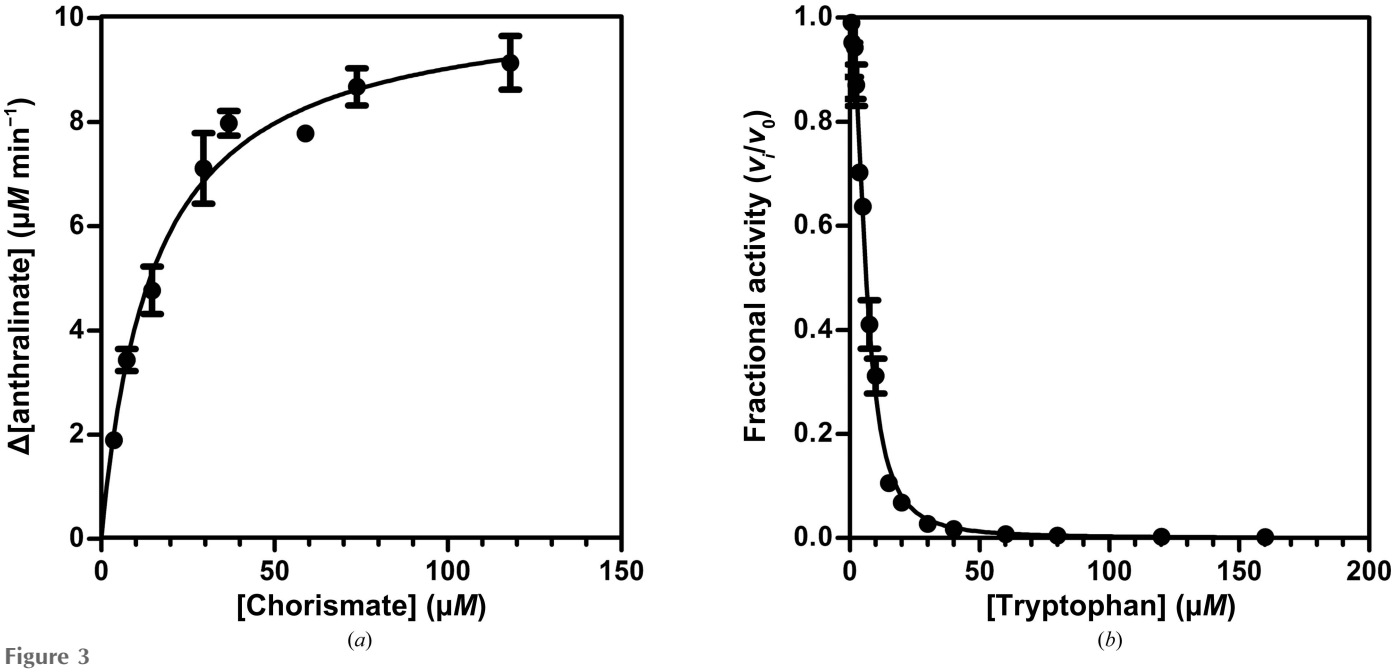


Figure 3
Enzymatic activity of *Mtb*-AS with varying concentrations of the substrate chorismate and in the presence of the allosteric inhibitor Trp. (*a*) Initial rates of anthranilate production were determined for different chorismate concentrations and were fitted to the Michaelis–Menten equation. K_m^{app} for chorismate was determined to be 15 ± 2 μ*M*, with a V_{max} of 10.4 ± 0.4 μ*M* min^{-1}. A k_{cat} of 0.26 ± 0.01 s^{-1} can also be determined; however, the calculation involves the assumption that AS-I and AS-II are in a 1:1 or a 2:2 stoichiometry; thus, this value may be inaccurate. (*b*) A dose-response curve comparing initial rates at various concentrations of Trp was fitted to determine a Hill coefficient of 2.1 ± 0.1 and an IC_{50} value of 6.3 ± 0.2 μ*M*. The error bars show the range of values for data points measured in duplicate.

Sty-AS (of 4.7 ± 0.8 and 3.9 ± 0.2 μ*M*, respectively) determined under similar conditions (*i.e.* saturating concentrations of Mg^{2+} and glutamine; Manos-Turvey *et al.*, 2010; Ziebart *et al.*, 2010). However, the K_m^{app} for *Mtb*-AS is lower than the values found for some plant AS enzymes (70–180 μ*M*; Bernasconi *et al.*, 1994; Poulsen *et al.*, 1993) and can be considered to be a reasonable value for a bacterial AS.

The pathway end product, Trp, was found to inhibit *Mtb*-AS activity (Fig. 3*b*), as previously shown for *Mtb*-AS-I (Lin *et al.*, 2009). A Hill coefficient of 2 was calculated at chorismate concentrations near the K_m^{app} (Fig. 3*b*) and at higher concentrations, indicating positive cooperativity. Similar Hill coefficients have been reported for a heterotetrameric plant AS (Bernasconi *et al.*, 1994; Poulsen *et al.*, 1993). Cooperative Trp inhibition is also seen for *Sma*-AS and *Sty*-AS, which oligomerize *via* AS-I dimers (Morollo & Eck, 2001; Spraggon *et al.*, 2001). Our result suggests that *Mtb*-AS also oligomerizes through interactions between its AS-I subunits.

3.3. Structure of the AS-I subunit

Diffracting crystals were obtained from solutions of both *Mtb*-AS-I and the *Mtb*-AS complex, in each case in the presence of the inhibitor methyl-AMT. Our initial structure, of *Mtb*-AS-I alone, was only determined at moderate resolution (2.6 Å), but could be used as a molecular-replacement model for the best diffracting *Mtb*-AS complex crystal. It was immediately apparent that the putative *Mtb*-AS crystals, which were isomorphous with the *Mtb*-AS-I crystals, contained no density corresponding to the *Mtb*-AS-II subunit, suggesting that it was either absent from the crystal or spatially disordered. Since the *Mtb*-AS-I structures obtained in the two cases are essentially identical, the higher resolution structure, refined at 2.1 Å resolution with a crystallographic R = 20.1% and R_{free} = 25.3% (Table 1), is presented here. One molecule of *Mtb*-AS-I is present in the asymmetric unit, with a solvent content of 69.7%. The final model includes residues 3–511, except for residues 116–119, for which no interpretable electron density was observed. These residues are located on the protein surface, remote from crystal contacts.

The AS-I monomer has an α/β fold, the core of which is a twisted β-sandwich formed by two orthogonal, antiparallel β-sheets built from 24 β-strands. This all-β core is surrounded by 14 helices (11 α-helices and three 3_{10}-helices; Fig. 4*a*), giving an overall fold that is topologically similar to other chorismate-utilizing enzymes (CUEs) of known structure. The structure is best described in terms of two subdomains, I and II (Knöchel *et al.*, 1999; Spraggon *et al.*, 2001), with each subdomain forming one side of the β-sandwich core (Fig. 4*a*). Subdomain I comprises residues 1–73, 163–306 and 459–511, and the rest of the residues belong to the other subdomain. The active site, which contains the methyl-AMT inhibitor, is located in a cleft between the subdomains, whereas the allosteric site is located 18 Å away within the β-sandwich core, as inferred from comparisons with homologous AS enzymes (Spraggon *et al.*, 2001; Morollo & Eck, 2001).

A search of the Protein Data Bank (PDB) using *SSM* (Krissinel & Henrick, 2004) shows that the best matches to

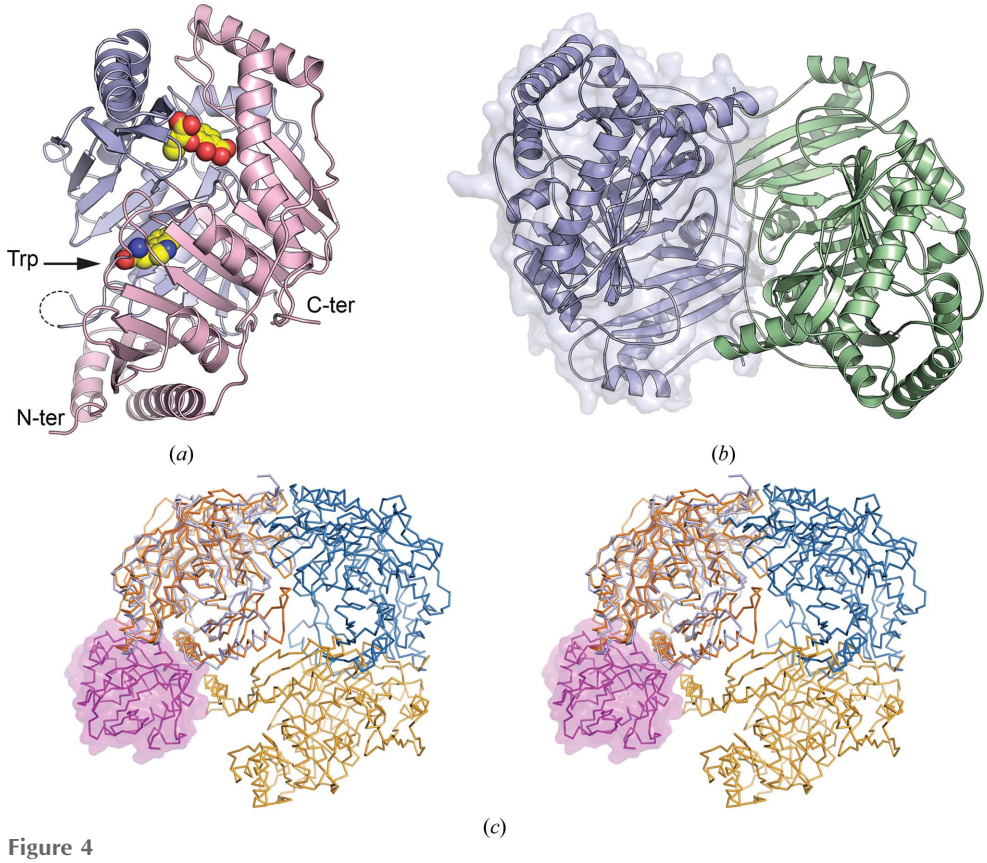


Figure 4
Three-dimensional structure of *Mtb*-AS-I. (*a*) The *Mtb*-AS-I protomer, with subdomain I coloured blue and subdomain II pink. The methyl-AMT inhibitor, shown in space-filling mode, is located in the active-site cleft between the two subdomains. Below it, some ∼20 Å away, is the allosteric Trp site (indicated with an arrow) identified by analogy with other AS enzyme structures. (*b*) Structure of *Mtb*-AS-I dimer with the two subunits coloured blue and green. The blue subunit is also shown in a transparent surface representation. (*c*) Stereoview of the *Mtb*-AS-I dimer (blue C^{α} trace) superimposed on the *Sma*-AS-I dimer (orange C^{α} trace) to highlight the different mode of dimerization. The two subunits in each AS-I dimer are coloured in different shades for clarity. Note that the viewing angle is different from that in (*b*) to show the different dimer interface more clearly. Also shown is the binding site for one of the two AS-II subunits in *Sma*-AS (magenta, transparent surface) showing that it is far removed from the new AS-I dimer interface in *Mtb*-AS. This demonstrates that the same AS-I–AS-II interaction is possible in *Mtb*-AS. (For clarity, the second AS-II subunit in *Sma*-AS is not shown, but by symmetry would be equally far removed from the new AS-I dimer interface.)

Mtb-AS-I are the three other structurally characterized AS-I subunits from AS complexes (Knöchel *et al.*, 1999; Morollo & Eck, 2001; Spraggon *et al.*, 2001), followed by CUEs from the folate (Parsons *et al.*, 2002; Bera *et al.*, 2012), phenazine (Li *et al.*, 2011), enterobactin (Kerbarh *et al.*, 2006) and mycobactin (Harrison *et al.*, 2006; Zwahlen *et al.*, 2007) biosynthetic pathways. Root-mean-square differences (r.m.s.d.s) in C^{α} positions range from 1.47 to 2.42 Å, highlighting the conservation of the core fold within this family of proteins. As has been noted before (Parsons *et al.*, 2002), most of the variations in different CUEs, in structure and in sequence, occur in the N-terminal portion of the protein where the Trp regulatory site is located in AS-I.

3.4. Oligomerization and identification of a novel dimer interface

Analysis of the crystal packing using *PISA* (Krissinel & Henrick, 2007) indicates that the biological unit of *Mtb*-AS-I is a dimer (Fig. 4*b*) that buries ∼1788 Å^2 or 8.5% of the solvent-accessible area of each subunit. This agrees with the SEC-MALS analysis for AS-I, which showed a 108 kDa dimer, similar to the predicted M_r of 112 kDa. In the crystal, the two subunits that comprise the AS-I dimer are related by a crystallographic twofold axis (Fig. 4*b*) and associate primarily through subdomain I (residues 3–6, 28–43, 57–65, 170–174, 206–222, 274–281 and 506–510); only one residue from subdomain II is involved.

Intriguingly, the AS-I–AS-I interface in *Mtb*-AS differs markedly from that observed in the other two mesophilic AS complexes, *Sma*-AS and *Sty*-AS (Fig. 4*c*), which involves more residues from subdomain II (equivalent to the region around α7, α8, β18 and β19 in our structure). In these two complexes the AS-I–AS-I interface is significantly smaller, burying only 725–806 Å^2 or 3.5–4% of the solvent-accessible surface area of each monomer, and is not given high significance by *PISA*. The fact that this AS-I–AS-I interface is essentially the same in both *Sma*-AS and *Sty*-AS argues that it is not a crystal-packing artefact, however. Cooperative interactions within the 2:2 heterotetramers must presumably enhance oligomer stability.

The different and much more extensive AS-I–AS-I interface in *Mtb*-AS results from changes on the surface of the *Mtb*-AS-I subunit, notably in the N-terminal portion of subdomain I. In particular, a 20-residue N-terminal extension that is not present in the other enzymes contributes significantly to each end of the interface. Interestingly, residues close to the allosteric Trp-binding site form part of the dimer interface, suggesting the possibility that Trp binding in one subunit could be signalled to the other; see §3.7.

Importantly, however, the novel AS-I–AS-I interface observed here should still allow the AS-II subunit to bind at the conserved site on AS-I that is utilized in all of the AS complexes characterized to date (Fig. 4*c*); there appear to be no steric clashes that would prevent this. Nevertheless, it is possible that stronger AS-I–AS-I association in the *Mtb*-AS complex may be linked to more facile dissociation of the AS-II subunits.

To confirm that the crystallographic dimer corresponds to that existing in solution, we performed SAXS analyses on the *Mtb*-AS-I protein (Fig. 5). Consistent with the SEC-MALS analysis (Fig. 2*b*), there was no concentration dependence evident in the scattering data, and the R_g was also consistent at 32.6 Å. Comparison of the experimental SAXS scattering data with the theoretical scattering curves from the crystal structures of the *Mtb*-AS-I dimer and the *Sma*-AS-I dimer shows that the *Mtb*-AS-I dimer fits much better, with χ = 0.47 for *Mtb*-AS-I but χ = 1.37 for *Sma*-AS-I (Fig. 5). This confirms that the *Mtb*-AS-I dimer seen in the crystal is the same as that in solution, and confirms that the dimer interface in *Mtb*-AS-I is different from those previously characterized.

The novel AS-I–AS-I dimer interface seen here fits with a pattern of variability within the wider CUE family among those enzymes that require a GAT subunit (Bera *et al.*, 2012; Parsons *et al.*, 2002). For example, the CUE from the folate-biosynthetic pathway in *E. coli* is a heterodimer involving a GAT subunit that can associate or dissociate under certain conditions (Roux & Walsh, 1992; Rayl *et al.*, 1996). In the AS subfamily, the AS-I–AS-II heterodimer interface is largely conserved between *Sso*-AS, *Sma*-AS and *Sty*-AS (Fig. 4*d*), but the quaternary structure differs (Morollo & Eck, 2001). In *Sso*-AS, dimerization occurs through AS-II and the AS-I subunits do not associate, whereas for *Sma*-AS and *Sty*-AS dimerization occurs through a relatively small AS-I–AS-I interface that differs from that seen here.

Analysis of the *Mtb*-AS-I crystal packing indicates that part of the expected binding face for *Mtb*-AS-II, based on homologous structures, is involved in crystal contacts. Thus, even

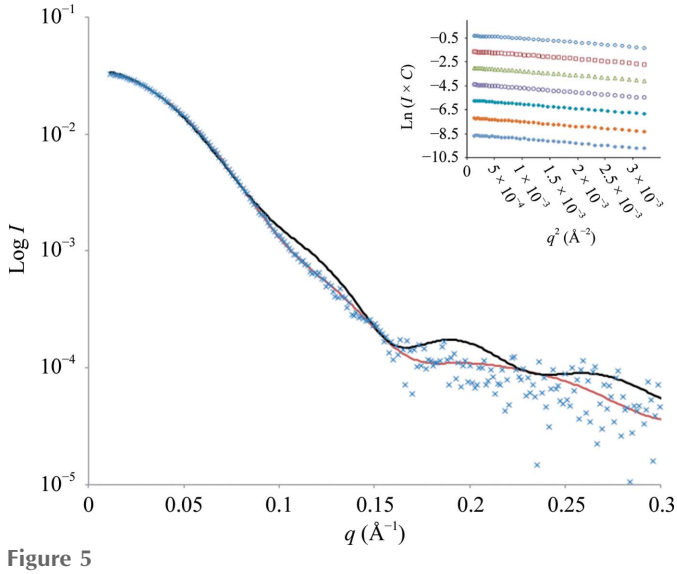


Figure 5
SAXS analyses of *Mtb*-AS-I protein. SAXS data are plotted against scattering angle and are compared with the *CRYSOL*-generated theoretical scattering (red line) obtained for our *Mtb*-AS-I crystal structure. The fitting has a χ value of 0.47. Also shown (black line) is the theoretical scattering curve for the different dimer seen for *Sma*-AS-I, which fits much less well (χ = 1.37). The inset shows the Guinier plot, in which the experimental SAXS intensity data are plotted against scattering angle over a range of concentrations [from 3 mg ml^{-1} (top line) with twofold dilutions to 0.045 mg ml^{-1} at the bottom]. The radius of gyration (R_g) is consistent at 32.6 Å over the concentration range.

though this crystal form had a high solvent content (69.7%) the AS-II subunits could not be accommodated in it. We conclude that the absence of AS-II from the crystal structure is owing to selective crystallization of AS-I arising either from the presence of some uncomplexed AS-I in the AS protein solution or to dissociation of the complex during crystallization.

3.5. Active site and methyl-AMT binding

The active site of *Mtb*-AS-I is located in a deep, solvent-filled funnel between the two subdomains (Fig. 4*a*) and is highly conserved when compared with the other AS-I (TrpE) structures (Spraggon *et al.*, 2001; Morollo & Eck, 2001; Knöchel *et al.*, 1999). AS-I active sites are inherently flexible, enabling closure over a bound substrate or the stabilization of a more open structure when Trp is bound as an allosteric inhibitor (Spraggon *et al.*, 2001; Morollo & Eck, 2001). Our present *Mtb*-AS-I structure corresponds to the closed form, consistent with the fact that the methyl-AMT inhibitor is bound in the active site. The active site is not fully formed, however, as the inhibitor displaces several side chains and no Mg^{2+} ion is bound. Mg^{2+} binding appears to be dependent on the presence of the carboxyl group of chorismate, or an analogue, and in *Mtb*-AS-I the residues expected to complete the Mg^{2+} site, Glu351 and Glu488, have moved away relative to the equivalent residues in *Sma*-AS (Spraggon *et al.*, 2001).

The methyl-AMT inhibitor is bound as its *Z*-isomer, with its aromatic ring sandwiched between the side chains of the conserved Lys492, with which it makes an N—H$\cdots\pi$ interaction, and Ile458 (Figs. 6*a* and 6*b*). As was also the case for its binding to *Mtb*-SS (Chi *et al.*, 2012), the inhibitor is flipped

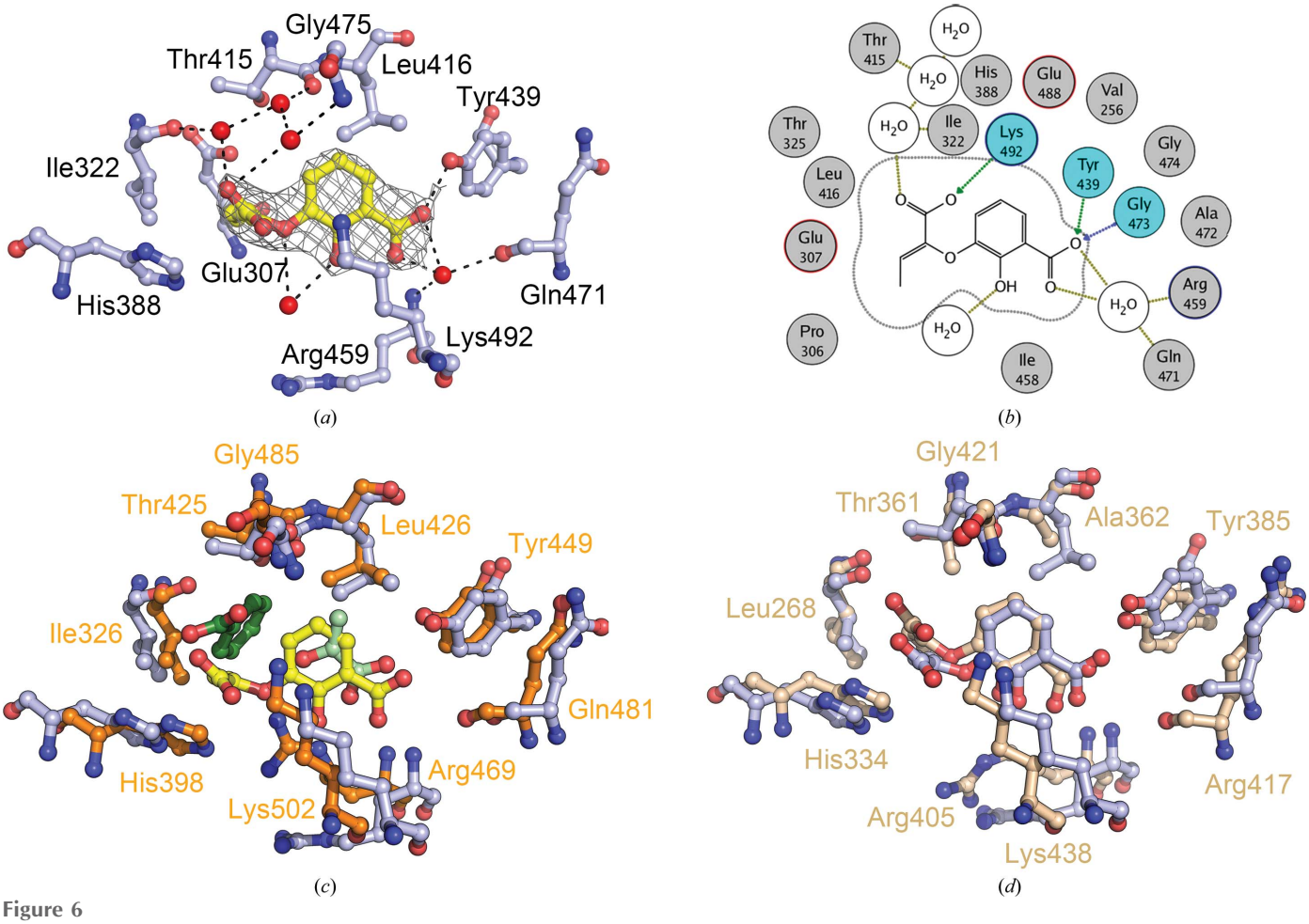


Figure 6
Mtb-AS-I active site and binding of the methyl-AMT inhibitor. (*a*) Methyl-AMT (yellow) binding in the *Mtb*-AS-I active site (blue) with $2F_o - F_c$ OMIT map (grey) contoured at 1.0σ. Water molecules making direct hydrogen bonds (black dashed lines) to methyl-AMT are shown as red spheres. The side chain of Ile458 packs directly below the aromatic ring of the inhibitor but is omitted for clarity. (*b*) Schematic representation of methyl-AMT binding in the active site with nearby residues shown as spheres. The residues that form hydrogen bonds directly to the inhibitor (cyan), *via* side-chain interactions (green lines) or *via* backbone (blue lines) are displayed. Acidic and basic amino acids are shown with red and blue outlines, respectively. (*c*) Superposition of the *Mtb*-AS-I active site (blue) on to that of *Sma*-AS-I (orange) indicates conservation of the residues in the active site of both proteins [residues in *Sma*-AS-I labelled in orange correspond to the residues in *Mtb*-AS-I shown in (*a*)]. The superposition also shows that the binding mode for methyl-AMT from *Mtb*-AS-I is rotated 180° with respect to benzoate (dark green) and pyruvate (light green) in the *Sma*-AS-I structure. (*d*) Superposition of active-site residues in *Mtb*-AS-I (blue) and *Mtb*-SS (wheat), indicating the conservation of active-site residues and the similar binding mode of methyl-AMT in both proteins. The methyl-AMT inhibitor is coloured the same as the protein side chains. Residues in *Mtb*-SS are labelled in this figure, superimposed on the *Mtb*-AS-I residues identified in (*a*).

Table 3
Inhibition of *Mtb*-AS by isochorismate and chorismate mimics compared with the results for other chorismate-utilizing enzymes.

ND, not determined.

Inhibitor	IC_{50} for *Mtb*-AS (μM)	K_i for *Mtb*-SS† (μM)	K_i for *Sma*-AS† (μM)	Decreased potency relative to methyl-AMT		
				Mtb-AS	*Mtb*-SS	*Sma*-AS
Methyl-AMT (9:1 *Z*/*E*)	17 ± 1	11 ± 1	1.1 ± 0.1			
Ethyl-AMT (10:3 *Z*/*E*)	25 ± 1	12 ± 2	1.7 ± 0.3	1.5-fold	1.1-fold	1.5-fold
Phenyl-AMT (1:1 *Z*/*E*)	80 ± 5	21 ± 5	33 ± 4	5-fold	1.9-fold	30-fold
AMT	540 ± 30	240 ± 40	34 ± 6	32-fold	22-fold	30-fold
2-Deoxy-AMT	420 ± 50	500 ± 90	3.2 ± 0.3	25-fold	46-fold	2.9-fold
GBS-1	2200 ± 300	1400 ± 400	28 ± 7	140-fold	130-fold	26-fold
GBS-2	>2000	3000 ± 1000	90 ± 14	ND	270-fold	82-fold

† Data from Manos-Turvey *et al.* (2010).

almost 180° relative to the orientation predicted for the ADIC intermediate by analogy with the binding modes for benzoate and pyruvate in *Sma*-AS-I (Spraggon *et al.*, 2001; Fig. 6*c*). Thus, the salicylate-like portion of methyl-AMT sits between Tyr439 and Arg459, residues that bind the enol-pyruvate moiety in *Sma*-AS-I, and the enol-pyruvyl group of methyl-AMT is oriented towards the opening of the active site, in a position close to where the anthranilate-like component of ADIC is predicted to reside. This places it adjacent to the catalytically essential His388 (Morollo & Bauerle, 1993) and to the site occupied by the essential Mg^{2+} ion in *Sma*-AS-I. The side chain of Arg459, which is also important for catalysis, is displaced by the methyl-AMT and no longer hydrogen-bonds to His388 as in *Sma*-AS (Fig. 6*c*).

Diffracting crystals of *Mtb*-AS-I could only be obtained reproducibly in the presence of the most potent inhibitor of our AMT series, methyl-AMT, presumably because of the conformational movements that it induces. The structure shows that the methyl group contacts Pro306, Ile322 and Ile458 in a small hydrophobic depression. Larger substituents (*i.e.* those of ethyl-AMT and phenyl-AMT) could not occupy the same site without movements of the β-strands that form this part of the active-site wall, thus accounting for their twofold to fivefold lower inhibitor potency compared with methyl-AMT (Table 2) and very likely also for their failure to crystallize. Interestingly, methyl-AMT binds in exactly the same orientation in *Mtb*-SS, in which binding of the phenyl derivative is accommodated by movements in strands β15–β17 (Chi *et al.*, 2012). *Mtb*-AS and *Mtb*-SS thus not only share similar core structures, with an r.m.s.d of 2.1–2.2 Å for 387–393 common C^{α} atoms, but also share the same inhibitor-binding modes (Fig. 6*d*).

3.6. Comparative inhibition studies

AS-I belongs to a family of closely related CUEs that share a similar structure and catalytic mechanism (Kozlowski *et al.*, 1995; Ziebart *et al.*, 2010). Owing to the closely related active-site architecture, inhibitors of one member of this enzyme family can also inhibit other family members (Kozlowski *et al.*, 1995; Manos-Turvey *et al.*, 2010; Payne *et al.*, 2009; Walsh *et al.*, 1987; Ziebart *et al.*, 2010). Isochorismate is the intermediate in the reaction catalysed by salicylate synthase (SS; Kerbarh, Ciulli *et al.*, 2005) and is structurally similar to the intermediate of the AS-catalysed reaction, ADIC (Fig. 1*b*). We previously developed isochorismate mimetics (AMT series; Fig. 1) and chorismate mimetics [*e.g.* 2-deoxy-AMT, gallate-based synthesis (GBS) series; Fig. 1] as inhibitors of *Mtb*-SS (Manos-Turvey *et al.*, 2010) and found that these compounds also inhibit *Sma*-AS (Manos-Turvey *et al.*, 2010). To establish their inhibitory properties against *Mtb*-AS and draw comparisons with *Sma*-AS and *Mtb*-SS, we determined the IC_{50} values for a subset of these inhibitors (Table 3). Interestingly, comparison of the trends from values in our previous results obtained for *Mtb*-SS showed a strong correlation despite the differences in the catalysed reaction.

The compounds AMT, 2-deoxy-AMT, GBS-1 and GBS-2 are relatively potent inhibitors of *Sma*-AS (K_i values of 3–90 μM), but are only modest inhibitors of *Mtb*-SS (K_i values of 240–3000 μM; Manos-Turvey *et al.*, 2010). The IC_{50} values (Table 3) identify these as similarly modest inhibitors of *Mtb*-AS. The most potent *Mtb*-AS inhibitor of the AMT series is methyl-AMT, with an IC_{50} value of 17 ± 1 μM and a Hill coefficient of 1.0 ± 0.4 (indicating noncooperative inhibition). The potency of ethyl-AMT and phenyl-AMT against *Mtb*-AS is twofold to fivefold lower than that of methyl-AMT (Table 3). In the case of *Mtb*-SS, which is monomeric, the small changes in potency between methyl-AMT and phenyl-AMT (Table 3) are attributed to active-site flexibility (Chi *et al.*, 2012). In *Sma*-AS, however, the part of the enzyme that has to move to accommodate the phenyl moiety of phenyl-AMT is engaged in the AS-I dimer interface, explaining its much lower potency (Chi *et al.*, 2012). The similar behaviour for *Mtb*-AS and *Mtb*-SS, but the differences from *Sma*-AS, are consistent with the fact that in *Mtb*-AS the AS-I–AS-I dimer interface differs from that of *Sma*-AS.

3.7. Allosteric inhibition by Trp

The amino-acid sequence of *Mtb*-AS-I contains Trp-binding motifs in subdomain I (LLESX_{10}S at residues 53–67 and NPSPYMY at residues 280–286) that correspond to those in homologous enzymes, supporting our observation of allosteric control by Trp (Fig. 3*b*). In both *Sma*-AS and *Sty*-AS the cooperativity of allosteric inhibition by Trp arises because structural elements that interact across the AS-I–AS-I interface belong to secondary structures that also help form the active site (Spraggon *et al.*, 2001; Morollo & Eck, 2001). How, then, is cooperative inhibition by Trp achieved in *Mtb*-AS, given that it has a completely different AS-I–AS-I interface?

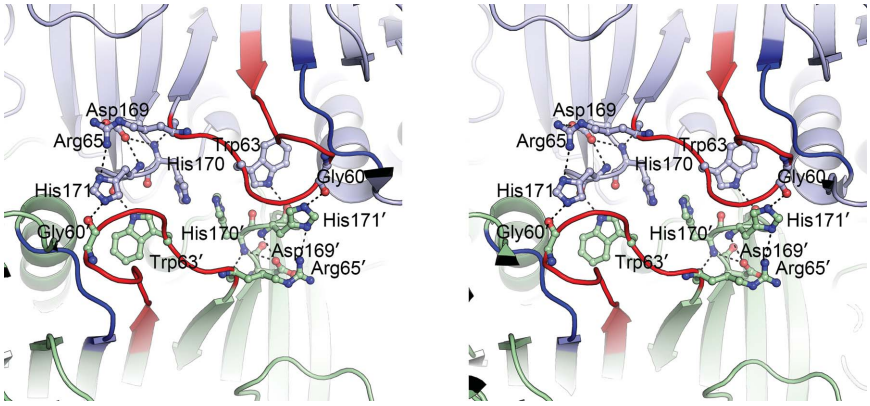


Figure 7
Mtb-AS-I allosteric site and proposed structural basis of cooperative allosteric inhibition by Trp. Stereoview of the *Mtb*-AS-I dimer interface around the allosteric binding site. The two monomers are shown as blue and green cartoon models, with the corresponding residues shown in ball-and-stick mode. The two allosteric motifs are shown in red (residues 53–67) and blue (residues 280–286) in the two monomers. Hydrogen bonds are shown as black dashed lines.

Our structure suggests that residues from the first allosteric loop region (residues 57–65) play a key part, as they approach each other at the interface (Fig. 7). Two histidine residues that are buried at the interface, His170 and His171, interact with allosteric loop residues (Ala60, Trp63 and Arg65) from either their own or the adjacent subunit, forming hydrogen bonds: His170 NH···Arg65 O (2.8 Å), His171 NH···Asp169 OD1 (2.9 Å), His171 O···Trp63 NE1 (3.0 Å), His171 NE2···Gly60 O (2.8 Å) and His171 ND1···Arg65 NH1 (3.1 Å). His170 also stacks against its symmetry mate, His170′, from the other monomer (Fig. 7). Beyond this residue, Asp169 stabilizes the turn containing His170 and His171 through multiple hydrogen bonds as well as by forming a salt bridge with Arg65 (2.8 Å). Upon productive Trp binding to one subunit, the conformation of this allosteric loop is likely to change, by analogy with the changes seen in homologous AS enzymes, and this would very likely alter the hydrogen-bonding patterns across the interface.

4. Conclusions

The Trp-biosynthesis pathway is of particular significance in *Mtb* as a target for new drugs that operate through a novel mechanism of action compared with current therapies. This is underscored by the report that fluorinated anthranilate analogues are effective at killing *Mtb* in a mouse model of disease, apparently through Trp starvation (Zhang *et al.*, 2013). This is likely to be associated with their turnover by the second enzyme in the pathway, *Mtb*-AnPRT (Cookson *et al.*, 2014). We are currently pursuing a rational drug-design approach for *Mtb*-AnPRT (Castell *et al.*, 2013; Evans *et al.*, 2014). In order to more fully characterize the Trp-biosynthesis pathway in *Mtb*, we have here used functional and structural studies to characterize the first committed step in Trp biosynthesis, which is catalysed by the AS complex. Our results indicate that the protein encoded by open reading frame Rv0013 is the amidotransferase (AS-II) component of anthranilate synthase in *Mtb*, forming a functional AS complex with AS-I and providing ammonia through glutamine hydrolysis. Attempts to crystallize the AS complex only gave crystals of its AS-I subunit, probably owing to its preferential crystallization from a solution containing both AS and AS-I.

Interestingly, the crystal structure of *Mtb*-AS-I revealed a dimer interface different from that observed for other AS complex enzymes, although this different dimer still supports cooperative allosteric AS inhibition by Trp and allows formation of the full *Mtb*-AS complex. Inhibitor potency trends against *Mtb*-AS did not correlate with those observed for *Sma*-AS, most likely because the flexibility in the active site of *Mtb*-AS is changed owing to its different AS-I dimer interface. We predict that inhibitors developed against a particular AS complex (*e.g.* *Sma*-AS or *Sty*-AS) would not necessarily be as potent against other AS enzymes (*e.g.* *Sso*-AS), owing to the difference in their oligomeric association and dimer interface.

The potential to design a single CUE inhibitor to target multiple essential biosynthetic pathways has been proposed by several research groups (Walsh *et al.*, 1987; Kozlowski *et al.*, 1995). Chorismate is the branch point for pathways leading to various essential compounds, including the aromatic amino acids, folate and siderophores (enterobactin and mycobactin). This 'magic shotgun' rather than 'magic bullet' approach has been growing in popularity and acceptance in recent years (Roth *et al.*, 2004; Morphy & Rankovic, 2005). Here, we demonstrate the potential for an inhibitor to target multiple CUEs in *Mtb*, since the compounds we tested had similar potency trends against *Mtb*-AS as they did against *Mtb*-SS (Manos-Turvey *et al.*, 2010). Furthermore, we observe that the most potent of these inhibitors, methyl-AMT, has an identical mode of binding to both enzymes (Chi *et al.*, 2012).

Acknowledgements

We thank Drs Mike Herbert and Chris Squire for help with SAXS and X-ray data collection, respectively, both of which were performed at the Australian Synchrotron. We thank Mr Martin Middleditch for mass spectrometry, which was performed at the Centre for Genomics, Proteomics and Metabolomics at the University of Auckland. This work was supported by research grants from the Health Research Council of New Zealand and Ministry of Business, Innovation and Employment of New Zealand. Access to the Australian Synchrotron was supported by the New Zealand Synchrotron Group.

References

Ashenafi, M., Carrington, R., Collins, A. C. & Byrnes, W. M. (2008). *Ethn. Dis.* **18**, S2-9–S2-13.
Barona-Gómez, F. & Hodgson, D. A. (2003). *EMBO Rep.* **4**, 296–300.

Bashiri, G., Rehan, A. M., Greenwood, D. R., Dickson, J. M. & Baker, E. N. (2010). *PLoS One*, **5**, e15803.

Bera, A. K., Atanasova, V., Dhanda, A., Ladner, J. E. & Parsons, J. F. (2012). *Biochemistry*, **51**, 10208–10217.

Bernasconi, P., Walters, E. W., Woodworth, A. R., Siehl, D. L., Stone, T. E. & Subramanian, M. V. (1994). *Plant Physiol.* **106**, 353–358.

Berney, M., Berney-Meyer, L., Wong, K.-W., Chen, B., Chen, M., Kim, J., Wang, J., Harris, D., Parkhill, J., Chan, J., Wang, F. & Jacobs, W. R. Jr (2015). *Proc. Natl Acad. Sci. USA*, **112**, 10008–10013.

Blommel, P. G. & Fox, B. G. (2007). *Protein Expr. Purif.* **55**, 53–68.

Castell, A., Short, F. L., Evans, G. L., Cookson, T. V., Bulloch, E. M., Joseph, D. D., Lee, C. E., Parker, E. J., Baker, E. N. & Lott, J. S. (2013). *Biochemistry*, **52**, 1776–1787.

Chi, G., Manos-Turvey, A., O'Connor, P. D., Johnston, J. M., Evans, G. L., Baker, E. N., Payne, R. J., Lott, J. S. & Bulloch, E. M. (2012). *Biochemistry*, **51**, 4868–4879.

Cole, S. T. *et al.* (2001). *Nature (London)*, **409**, 1007–1011.

Cookson, T. V. M., Castell, A., Bulloch, E. M. M., Evans, G. L., Short, F. L., Baker, E. N., Lott, J. S. & Parker, E. J. (2014). *Biochem. J.* **461**, 87–98.

Due, A. V., Kuper, J., Geerlof, A., von Kries, J. P. & Wilmanns, M. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 3554–3559.

Emsley, P. (2013). *Advancing Methods for Biomolecular Crystallography*, edited by R. J. Read, A. G. Urzhumtsev & V. Y. Lunin, pp. 205–209. Dordrecht: Springer.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Evans, G. L., Gamage, S. A., Bulloch, E. M. M., Baker, E. N., Denny, W. A. & Lott, J. S. (2014). *Chembiochem*, **15**, 852–864.

Harrison, A. J., Yu, M., Gårdenborg, T., Middleditch, M., Ramsay, R. J., Baker, E. N. & Lott, J. S. (2006). *J. Bacteriol.* **188**, 6081–6091.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.

Kerbarh, O., Bulloch, E. M., Payne, R. J., Sahr, T., Rébeillé, F. & Abell, C. (2005). *Biochem. Soc. Trans.* **33**, 763–766.

Kerbarh, O., Chirgadze, D. Y., Blundell, T. L. & Abell, C. (2006). *J. Mol. Biol.* **357**, 524–534.

Kerbarh, O., Ciulli, A., Howard, N. I. & Abell, C. (2005). *J. Bacteriol.* **187**, 5061–5066.

Knöchel, T., Ivens, A., Hester, G., Gonzalez, A., Bauerle, R., Wilmanns, M., Kirschner, K. & Jansonius, J. N. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 9479–9484.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.

Kozlowski, M. C., Tom, N. J., Seto, C. T., Sefler, A. M. & Bartlett, P. A. (1995). *J. Am. Chem. Soc.* **117**, 2128–2140.

Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.

Lee, C. E., Goodfellow, C., Javid-Majd, F., Baker, E. N. & Lott, J. S. (2006). *J. Mol. Biol.* **355**, 784–797.

Li, Q. A., Mavrodi, D. V., Thomashow, L. S., Roessle, M. & Blankenfeldt, W. (2011). *J. Biol. Chem.* **286**, 18213–18221.

Lin, X., Xu, S., Yang, Y., Wu, J., Wang, H., Shen, H. & Wang, H. (2009). *Protein Expr. Purif.* **64**, 8–15.

Manos-Turvey, A., Bulloch, E. M., Rutledge, P. J., Baker, E. N., Lott, J. S. & Payne, R. J. (2010). *ChemMedChem*, **5**, 1067–1079.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Moreland, N., Ashton, R., Baker, H. M., Ivanovic, I., Patterson, S., Arcus, V. L., Baker, E. N. & Lott, J. S. (2005). *Acta Cryst.* D**61**, 1378–1385.

Morollo, A. A. & Bauerle, R. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 9983–9987.

Morollo, A. A. & Eck, M. J. (2001). *Nature Struct. Biol.* **8**, 243–247.

Morollo, A. A., Finn, M. G. & Bauerle, R. (1993). *J. Am. Chem. Soc.* **115**, 816–817.

Morphy, R. & Rankovic, Z. (2005). *J. Med. Chem.* **48**, 6523–6543.

Mouilleron, S. & Golinelli-Pimpaneau, B. (2007). *Curr. Opin. Struct. Biol.* **17**, 653–664.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Mylonas, E. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s245–s249.

Parsons, J. F., Jensen, P. Y., Pachikara, A. S., Howard, A. J., Eisenstein, E. & Ladner, J. E. (2002). *Biochemistry*, **41**, 2198–2208.

Payne, R. J., Bulloch, E. M., Toscano, M. M., Jones, M. A., Kerbarh, O. & Abell, C. (2009). *Org. Biomol. Chem.* **7**, 2421–2429.

Poulsen, C., Bongaerts, R. J. & Verpoorte, R. (1993). *Eur. J. Biochem.* **212**, 431–440.

Raushel, F. M., Thoden, J. B. & Holden, H. M. (1999). *Biochemistry*, **38**, 7891–7899.

Rayl, E. A., Green, J. M. & Nichols, B. P. (1996). *Biochim. Biophys. Acta*, **1295**, 81–88.

Romero, R. M., Roberts, M. F. & Phillipson, J. D. (1995). *Phytochemistry*, **39**, 263–276.

Roth, B. L., Sheffler, D. J. & Kroeze, W. K. (2004). *Nature Rev. Drug Discov.* **3**, 353–359.

Roux, B. & Walsh, C. T. (1992). *Biochemistry*, **31**, 6904–6910.

Smith, D. A., Parish, T., Stoker, N. G. & Bancroft, G. J. (2001). *Infect. Immun.* **69**, 1142–1150.

Spraggon, G., Kim, C., Nguyen-Huu, X., Yee, M.-C., Yanofsky, C. & Mills, S. E. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 6021–6026.

Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.

Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.

Tamir, H. & Srinivasan, P. R. (1970). *Methods Enzymol.* **17**, 401–406.

Walsh, C. T., Erion, M. D., Walts, A. E., Delany, J. J. & Berchtold, G. A. (1987). *Biochemistry*, **26**, 4734–4745.

Wang, F., Jain, P., Gulten, G., Liu, Z., Feng, Y., Ganesula, K., Motiwala, A. S., Ioerger, T. R., Alland, D., Vilchèze, C., Jacobs, W. R. Jr & Sacchettini, J. C. (2010). *Antimicrob. Agents Chemother.* **54**, 3776–3782.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Xie, G., Keyhani, N. O., Bonner, C. A. & Jensen, R. A. (2003). *Microbiol. Mol. Biol. Rev.* **67**, 303–342.

Zhang, Y. J., Reddy, M. C., Ioerger, T. R., Rothchild, A. C., Dartois, V., Schuster, B. M., Trauner, A., Wallis, D., Galaviz, S., Huttenhower, C., Sacchettini, J. C., Behar, S. M. & Rubin, E. J. (2013). *Cell*, **155**, 1296–1308.

Zhang, Y. J. & Rubin, E. J. (2013). *Cell. Microbiol.* **15**, 1079–1087.

Ziebart, K. T., Dixon, S. M., Avila, B., El-Badri, M. H., Guggenheim, K. G., Kurth, M. J. & Toney, M. D. (2010). *J. Med. Chem.* **53**, 3718–3729.

Zumla, A., Nahid, P. & Cole, S. T. (2013). *Nature Rev. Drug Discov.* **12**, 388–404.

Zwahlen, J., Kolappan, S., Zhou, R., Kisker, C. & Tonge, P. J. (2007). *Biochemistry*, **46**, 954–964.