Data in Brief 15 (2017) 1019–1029

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Datasets, processing and refinement details for *Mtb*-AnPRT: inhibitor structures with various space groups

Genevieve L. Evans [a,*], Daniel P. Furkert [b], Nacim Abermil [b], Preeti Kundu [c], Katrina M. de Lange [a,1], Emily J. Parker [c], Margaret A. Brimble [b], Edward N. Baker [a], J. Shaun Lott [a,**]

[a] *Maurice Wilkins Centre for Molecular Biodiscovery and School of Biological Sciences, University of Auckland, 3 Symonds Street, Auckland 1142, New Zealand*
[b] *Maurice Wilkins Centre for Molecular Biodiscovery and School of Chemical Sciences, University of Auckland, 23 Symonds Street, Auckland 1142, New Zealand*
[c] *Maurice Wilkins Centre for Molecular Biodiscovery, Biomolecular Interaction Centre and Department of Chemistry, University of Canterbury, P.O. Box 4800, Christchurch 8140, New Zealand*

ARTICLE INFO

Article history:
Received 26 August 2017
Received in revised form
17 October 2017
Accepted 20 October 2017

Keywords:
Crystallography
Macromolecules
X-ray diffraction
Ligand binding
Structure-based inhibitor design

ABSTRACT

There are twenty-five published structures *of Mycobacterium tuberculosis* anthranilate phosphoribosyltransferase (*Mtb*-AnPRT) that use the same crystallization protocol. The structures include protein complexed with natural and alternative substrates, protein:inhibitor complexes, and variants with mutations of substrate-binding residues. Amongst these are varying space groups (*i.e.* $P2_1$, $C2$, $P2_12_12$, $P2_12_12_1$). This article outlines experimental details for 3 additional *Mtb*-AnPRT:inhibitor structures. For one protein:inhibitor complex, two datasets are presented – one generated by crystallization of protein in the presence of the inhibitor and another where a protein crystal was soaked with the inhibitor.

DOI of original article: http://dx.doi.org/10.1016/j.bbapap.2017.08.018
* Corresponding author. Present address: KU Leuven, Laboratory for Structural Neurobiology, Department of Cellular and Molecular Medicine, Herestraat 49, P.O. Box 601, B-3000 Leuven, Belgium
** Corresponding author.
E-mail addresses: g.evans@auckland.ac.nz (G.L. Evans), s.lott@auckland.ac.nz (J.S. Lott).
[1] Present address: Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1HH, United Kingdom

http://dx.doi.org/10.1016/j.dib.2017.10.051
2352-3409/© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Automatic and manual processing of these datasets indicated the same space group for both datasets and thus indicate that the space group differences between structures of *Mtb*-AnPRT:ligand complexes are not related to the method used to introduce the ligand.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Specifications Table

Subject area	*Proteomics and Biochemistry*
More specific subject area	*Structural biology*
Type of data	*Tables, figures and X-ray diffraction images*
How data was acquired	*X-ray macromolecular crystallography; Australian Synchrotron MX1 and MX2 beamlines*
Data format	*Unprocessed, processed, deposited with crystal packing analyzed.*
Experimental factors	*Protein crystals soaked and co-crystallized with ligands*
Experimental features	*Contrast of the different protein crystal packing associated with different inhibitors and/or introduced by soaking or co-crystallization. X-ray diffraction datasets for co-crystallization and soaked-in experiments with the same inhibitor indicate space group changes are independent of method used to introduce the inhibitor.*
Data source location	*Data collected in Melbourne, Australia.*
Data accessibility	*The atomic coordinates and observed structure factors are available from Protein Data Bank with accession numbers:* 5BO2, 5BO3 *and* 5BNE. *The X-ray diffraction image files corresponding to datasets from two experiments are hosted by Mendeley:* http://http://dx.doi.org/10.17632/2zrfgv34nb.1 http://http://dx.doi.org/10.17632/xgn5z8jnr7.1
Related research article	*Anthranilate phosphoribosyltransferase: Binding determinants for 5′-phospho-alpha-D-ribosyl-1′-pyrophosphate (PRPP) and the implications for inhibitor design "in press".* https://doi.org/10.1016/j.bbapap.2017.08.018

Value of the data

- *Mtb*-AnPRT is a target of interest in developing novel anti-tuberculosis agents. This protein's capacity to crystallize thus yield new protein:ligand complexes makes it of interest for structure-based inhibitor design.
- Previously *Mtb*-AnPRT:ligand complexes have been found with different space groups (*e.g.* $P2_1$, $C2$, $P2_12_12$, and $P2_12_12_1$), generated by using the same crystallization protocol.
- *Mtb*-AnPRT:inhibitor complex structures described herein have $C2$ and $P2_1$ space groups. These structures were solved using X-ray diffraction datasets from protein:ligand crystals generated by streak-seeded with wild-type $P2_12_12_1$ crystal.
- For one inhibitor, X-ray diffraction datasets are presented for both co-crystallization and soaked crystal experiments. Space group $C2$ occurred in both datasets and indicates the space group

change between ligand-free and inhibitor-bound structures are independent of method used to introduce the inhibitor.
- X-ray diffraction datasets utilizing different methods of ligand introduction and yielding equivalent protein:ligand structures are typically not made available. These publically available datasets in the context of multiple space groups observed for *Mtb*-AnPRT:ligand structures, and the analysis presented herein, demonstrate that space group changes can be independent of co-crystallization and soaking methods of ligand introduction. This has relevance to academic and industrial researchers who are pursuing structure-based inhibitor design.

1. Data

1.1. Overview

The experimental and data processing details for 3 new protein:ligand structures of *Mycobacterium tuberculosis* anthranilate phosphoribosyl transferase (*Mtb*-AnPRT) are described herein (Table 1). The protein structures are complexed with *Mtb*-AnPRT inhibitors characterized in [1] and annotated as:

- **8k** (2-(2-carboxyphenylamino)-5-(5-phosphonopentyloxy)benzoic acid)
- **8j** (2-(2-carboxyphenylamino)-5-(4-phosphonobutoxy)benzoic acid)
- **8i** (2-(2-carboxyphenylamino)-5-(3-phosphonopropoxy)benzoic acid)

These structures were determined with protein crystallized in presence of imidazole-malate and PEG4000. Crystallization drops seeded with crystal nuclei from a pre-existing *Mtb*-AnPRT crystal generate better crystal morphology for *Mtb*-AnPRT [2]. All new structures presented herein were generated by streak-seeding using wild-type $P2_12_12_1$ crystals. Also presented are two X-ray diffraction datasets corresponding to protein crystals either soaked and co-crystallized with the same inhibitor (Table 2).

Of the previously published structures of *Mtb*-AnPRT, 25 were determined from protein crystallized using this protocol (Table 3). Most of these structures have a ligand bound (*e.g.* inhibitor or substrate) and/or are protein variants with mutations in substrate-binding residues (Table 3). Amongst both cohorts (Tables 1 and 3) several different space groups have been observed, including $P2_1$, $C2$, $P2_12_12$, $P2_12_12_1$. Amongst the 28 structures referred to in this article (Tables 1 and 3), similar unit cells correspond to each space group.

$P2_12_12_1$ is the most common space group for macromolecular structures, and it has been proposed that this is due to its capacity to accommodate repositioning, *i.e.* rotations or translations, within the asymmetric unit, without loss of crystal contacts [3]. The structure of wild-type *Mtb*-AnPRT without ligands (PDB ID: 3QR9) has previously been solved from protein crystallized in imidazole-malate and PEG4000 in the space group $P2_12_12_1$, with two monomers (A, B) in the asymmetric unit, a unit cell of 79×92×120 Å, and 57% solvent content [2].

Mtb-AnPRT is a homodimeric protein with an extended "S"-shape, with each subunit containing two domains [4]. In the ligand-free structure, a single dimer (the biological assembly) is found in the

Table 1
Summary of *Mtb*-AnPRT structures described herein.

PDB ID	Ligand ID	Solvent content	Unit Cell				Space group	Resolution (Å)	Chains	PDB DOI
			β (°)	A (Å)	B (Å)	C (Å)				
5BO2	**8i**	46%	111°	94	78	103	$C2$	2.00	A,B	10.2210/pdb5bo2/pdb
5BO3	**8j**	46%	111°	95	78	103	$C2$	1.75	A,B	10.2210/pdb5bo3/pdb
5BNE	**8k**	46%	91°	77	78	117	$P2_1$	2.15	A,B,C D	10.2210/pdb5bne/pdb

Table 2
Crystallization of complexes, along with space group and unit cell from data processing.

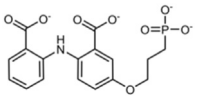
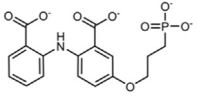
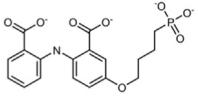
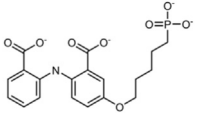
PDB ID	Ligand(s) bound	[Protein] mg mL^{-1}	Reservoir condition	Cryoprotectant/Soak	Notes	Space group	Unit cell
5BO2	**8i :** 	3.0	0.2 M imidazole. malate, pH 7.0, 9% PEG4000	0.2 M imidazole.malate, pH 7.0, 15% PEG4000, 1 mM **8i**	Streak seeded; **soaked crystal** for 4 h; 2 day old crystal **X-ray images available at:** http://http://dx.doi.org/10.17632/xgn5z8jnr7.1	*C*2	94×78×103 Å ßβ=111°
Not applicable	**8i :** 	3.0	0.2 M imidazole. malate, pH 7.0, 11% PEG4000, 1 mM **8i**	0.2 M imidazole.malate, pH 7.0, 15% PEG4000, 1 mM **8i**	Streak seeded; **protein co-crystallized with ligand**; 2 day old crystal **X-ray images available at:** http://http://dx.doi.org/10.17632/2zrfgv34nb.1	*C*2	95×78×103 Å β=111°
5BO3	**8j :** 	3.0	0.2 M imidazole. malate, pH 7.0, 11% PEG4000,	0.2 M imidazole.malate, pH 7.0, 15% PEG4000, 1 mM **8j**	Streak seeded; **soaked crystal** for 10 min; 2 day old crystal	*C*2	95×78×103 Å β=111°
5BNE	**8k :** 	3.1	0.2 M imidazole. malate, pH 7.0, 15% PEG4000, 1 mM **8k**	No cryo used, because crystallization condition contained 15% PEG4000	Streak seeded; **protein co-crystallized with ligand**; 2 day old crystal	$P2_1$	77×78×117 Å β=91°

Table 3
Space groups, unit cells and other information for previously published *Mtb*-AnPRT structures crystallized in imidazole-malate and PEG4000.

PDB ID	Ref	Solvent content	Unit Cell				Space Group	Resolution (Å)	Chains	Notes	PDB DOI
			β (°)	A (Å)	B (Å)	C (Å)					
3QS8	[2]	46%	90°	78	81	111	$P2_1$	2.00	A,B,C,D	Co-crystal	10.2210/pdb3qs8/pdb
3UU1	[2]	45%	90°	78	111	80	$P2_1$	1.82	A,B,C,D	Co-crystal	10.2210/pdb3uu1/pdb
3R6C	[2]	44%	110°	94	78	100	$C2$	1.83	A,B	Co-crystal	10.2210/pdb3r6c/pdb
4IJ1	[5]	45%	111	95	78	102	$C2$	1.79	A,B	Co-crystal	10.2210/pdb4ij1/pdb
4X58	[6]	50%	111	95	78	101	$C2$	1.75	A,B	Mutant	10.2210/pdb4x58/pdb
4X59	[6]	50%	112	95	78	102	$C2$	1.80	A,B	Mutant	10.2210/pdb4x59/pdb
4X5A	[6]	49%	112	94	78	102	$C2$	1.93	A,B	Mutant	10.2210/pdb4x5a/pdb
4X5B	[6]	45%	111	94	78	100	$C2$	2.47	A,B	Mutant	10.2210/pdb4x5b/pdb
4X5C	[6]	49%	111	94	78	101	$C2$	2.33	A,B	Mutant	10.2210/pdb4x5c/pdb
4X5E	[6]	50%	110	95	79	101	$C2$	1.77	A,B	Mutant	10.2210/pdb4x5e/pdb
4GIU	[5]	46%	90	111	81	79	$P2_12_12$	1.67	A,B	Co-crystal	10.2210/pdb4giu/pdb
4GKM	[5]	46%	90	111	81	78	$P2_12_12$	1.67	A,B	Co-crystal	10.2210/pdb4gkm/pdb
3QR9	[2]	57%	90	79	92	120	$P2_12_12_1$	1.87	A,B	Ligand-free	10.2210/pdb3qr9/pdb
4M0R	[5]	56%	90	79	92	120	$P2_12_12_1$	1.96	A,B	Co-crystal	10.2210/pdb4m0r/pdb
4N5V	[7]	57%	90	80	93	121	$P2_12_12_1$	1.90	A,B	Soak	10.2210/pdb4n5v/pdb
4N8Q	[7]	56%	90	80	91	120	$P2_12_12_1$	2.08	A,B	Soak	10.2210/pdb4n8q/pdb
4N93	[7]	57%	90	80	92	121	$P2_12_12_1$	2.03	A,B	Soak	10.2210/pdb4n93/pdb
4OWM	[7]	60%	90	79	92	121	$P2_12_12_1$	1.99	A,B	Soak	10.2210/pdb4owm/pdb
4OWN	[7]	60%	90	80	92	121	$P2_12_12_1$	2.11	A,B	Soak	10.2210/pdb4own/pdb
4OWO	[7]	60%	90	79	92	121	$P2_12_12_1$	1.99	A,B	Soak	10.2210/pdb4owo/pdb
4OWQ	[7]	61%	90	79	92	122	$P2_12_12_1$	1.89	A,B	Soak	10.2210/pdb4owq/pdb
4OWS	[7]	60%	90	80	92	121	$P2_12_12_1$	2.43	A,B	Soak	10.2210/pdb4ows/pdb
4OWU	[7]	60%	90	79	92	121	$P2_12_12_1$	1.89	A,B	Soak	10.2210/pdb4owu/pdb
4OWV	[7]	60%	90	80	92	120	$P2_12_12_1$	1.90	A,B	Soak	10.2210/pdb4owv/pdb
4X5D	[6]	60%	90	80	92	121	$P2_12_12_1$	2.30	A,B	Mutant	10.2210/pdb4x5d/pdb

asymmetric unit [2]. The two subunits of the dimer are related by a non-crystallographic two-fold symmetry axis (NCS).

1.2. Data for protein complexes with inhibitors

The three protein:inhibitor complexes (PDB IDs: 5BO2, 5BO3, 5BNE) were solved in the absence of metals and substrate. For these structures, the solvent content has decreased by ~ 10%, the unit cell has changed (*i.e.* dimension(s) decreased by 10–20 Å), and the space group has changed to $P2_1$ or $C2$, compared to the ligand-free wild-type structure.

In structures of AnPRT from other prokaryote species, domain movement is observed within subunits due to substrate binding and results in compression of the homodimer by 10 Å (*e.g.* D_{max} (maximum distance) changes from 110 to 100 Å [8]). However, superposition of the subunits in the new *Mtb*-AnPRT structures onto the subunits of the ligand-free structure indicates there are no large changes (Table 4; [9]). Additionally, the longest dimension of the *Mtb*-AnPRT dimer is relatively unchanged between the ligand-free protein structure and the protein:ligand complex structures (Table 4; [10]). Thus, the changes in space group are not driven by domain movements within each subunit.

The $P2_1$ structure (PDB ID: 5BNE) contains an inhibitor annotated as **8k**, and is the third structure of *Mtb*-AnPRT with this space group determined for protein crystallized in the imidazole-malate condition. The increased components in this structure's asymmetric unit (chains A-D, vs. chains A and B; Fig. 1**A**) means that the lower symmetry described by $P2_1$ can generate equivalent protein content in a similarly-sized unit cell as is observed with structures defined by space groups $P2_12_12_1$ or $P2_12_12$. The ß angle of 91° could be taken to suggest that the space group should be orthorhombic (*e.g.* $P2_12_12$ or $P2_12_12_1$). Both POINTLESS [11] and ZANUDA [12] indicated that $P2_1$ was the correct space group for this dataset, however.

(A) Superposition of the dimer (cartoon) from a *Mtb*-AnPRT:inhibitor structure defined by space group $P2_1$, PDB ID: 5BNE (chain A, B, C and D in green, cyan, pink and yellow, respectively), onto that of the ligand-free *Mtb*-AnPRT structure (PDB ID: 3QR9 [2]; $P2_12_12_1$; chain A and B in dark and light grey, respectively). The figure includes adjacent dimers (ribbons) in equivalent crystal layers (c-b plane in 3QR9 [2] and c-a plane in 5BNE). In (B) the superposition is re-colored with the *Mtb*-AnPRT: inhibitor structure in yellow and the ligand-free *Mtb*-AnPRT structure in dark grey. Arrows highlight the reorientation of dimers relative to each other. (C) Superposition of the dimer (cartoon) from a *Mtb*-AnPRT:inhibitor structure defined by space group $C2$, PDB ID: 5BO2 (marine blue) onto that of the ligand-free *Mtb*-AnPRT structure (PDB ID: 3QR9 [2]; $P2_12_12_1$; dark grey). In (D) superpositions in panels B-C are combined.

Structural superposition of the dimers from *Mtb*-AnPRT:inhibitor structures (PDB ID: 5BNE and 5BO2) onto the dimer present in the $P2_12_12_1$ ligand-free *Mtb*-AnPRT structure (PDB ID: 3QR9 [2]) revealed a reorientation of the protein dimers relative to each other (Fig. 1**B and C**). The combination

Table 4
Comparison *Mtb*-AnPRT subunits found in the 3 new structures to ligand-free structure.

	RMSD[a] to chain A-3QR9 (Å)	RMSD[a] to chain B-3QR9 (Å)	Longest dimension of dimer (Å)
3qr9.pdb:A	–	0.89	109
3qr9.pdb:B	0.89	–	
5bo2.pdb:A	0.67	0.68	110
5bo2.pdb:B	0.66	0.73	
5bo3.pdb:A	0.63	0.53	110
5bo3.pdb:B	0.62	0.73	
5bne.pdb:A	0.57	0.53	109
5bne.pdb:B	0.46	0.74	
5bne.pdb:C	0.46	0.92	109
5bne.pdb:D	0.55	0.60	

[a] Root mean standard difference (RMSD) between the C_{alpha} atoms

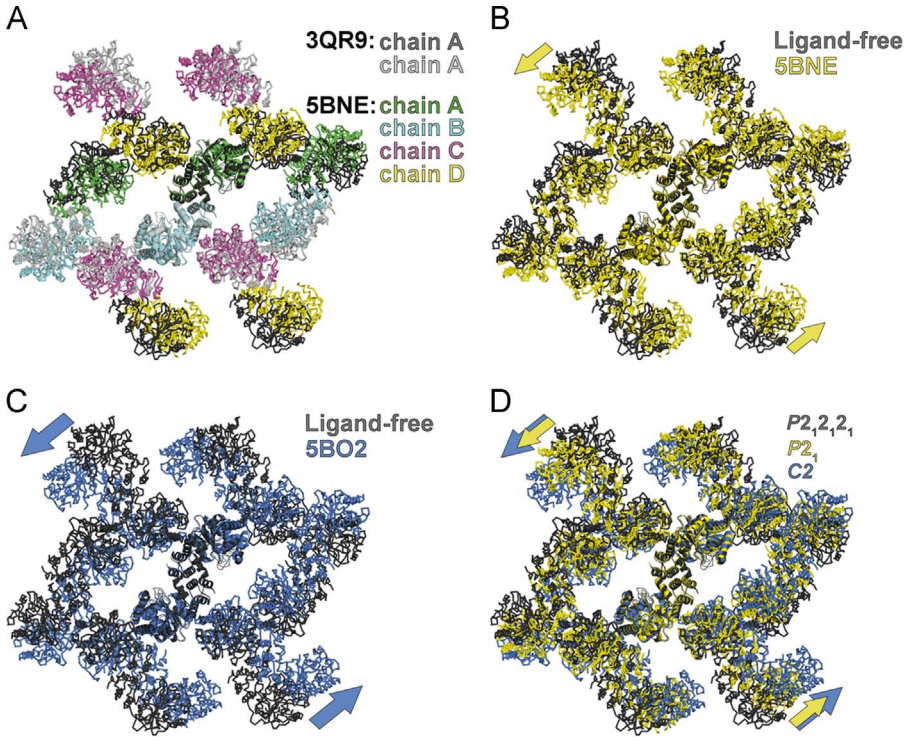


Fig. 1. Understanding changes in crystal packing in the *Mtb*-AnPRT structures with inhibitors. (A) Superposition of the dimer (cartoon) from a *Mtb*-AnPRT:inhibitor structure defined by space group $P2_1$, PDB ID: 5BNE (chain A, B, C and D in green, cyan, pink and yellow, respectively), onto that of the ligand-free *Mtb*-AnPRT structure (PDB ID: 3QR9 [2]; $P2_12_12_1$; chain A and B in dark and light grey, respectively). The figure includes adjacent dimers (ribbons) in equivalent crystal layers (c-b plane in 3QR9 [2] and c-a plane in 5BNE). In (B) the superposition is re-colored with the *Mtb*-AnPRT:inhibitor structure in yellow and the ligand-free *Mtb*-AnPRT structure in dark grey. Arrows highlight the reorientation of dimers relative to each other. (C) Superposition of the dimer (cartoon) from a *Mtb*-AnPRT:inhibitor structure defined by space group *C*2, PDB ID: 5BO2 (marine blue) onto that of the ligand-free *Mtb*-AnPRT structure (PDB ID: 3QR9 [2]; $P2_12_12_1$; dark grey). In (D) superpositions in panels B-C are combined.

of these superpositions (Fig. 1**D**) indicates that the lattice in the $P2_1$ structure corresponds to an intermediate position between the lattice observed for $P2_12_12_1$ and *C*2 structures. We propose the subunits that are related by crystallographic symmetry elements in the $P2_12_12_1$ structures are related by pseudosymmetry elements in the $P2_1$ structure. Pseudosymmetry occurs where a non-crystallographic symmetry element within the asymmetric unit is close to a crystallographic symmetry operators [13]. Thus, the $P2_1$ space group has been correctly assigned, even though the unit cell has a β of approximately 90 °.

The generation of the protein: inhibitor structures, involved experiments, with all three inhibitors, using both co-crystallization and soaking-in methods. The structures deposited on the PDB correspond to those where the clearest density was observed for the inhibitor (*i.e.* modelled with full occupancy; Fig. 2). In Table 5 there are data statistics corresponding to *Mtb*-AnPRT co-crystallized with inhibitor **8i**. In Table 6 are data statistics corresponding to inhibitor **8i** soaked into a wild-type ligand-free *Mtb*-APRT crystal. In both cases, with no manual intervention, the datasets processed in XDS [14] and AIMLESS [11] as space group *C*2. Thus, in this study, there no correlation was found between space group type and the method by which the ligand was introduced (*i.e.* soaking vs. co-crystallization).

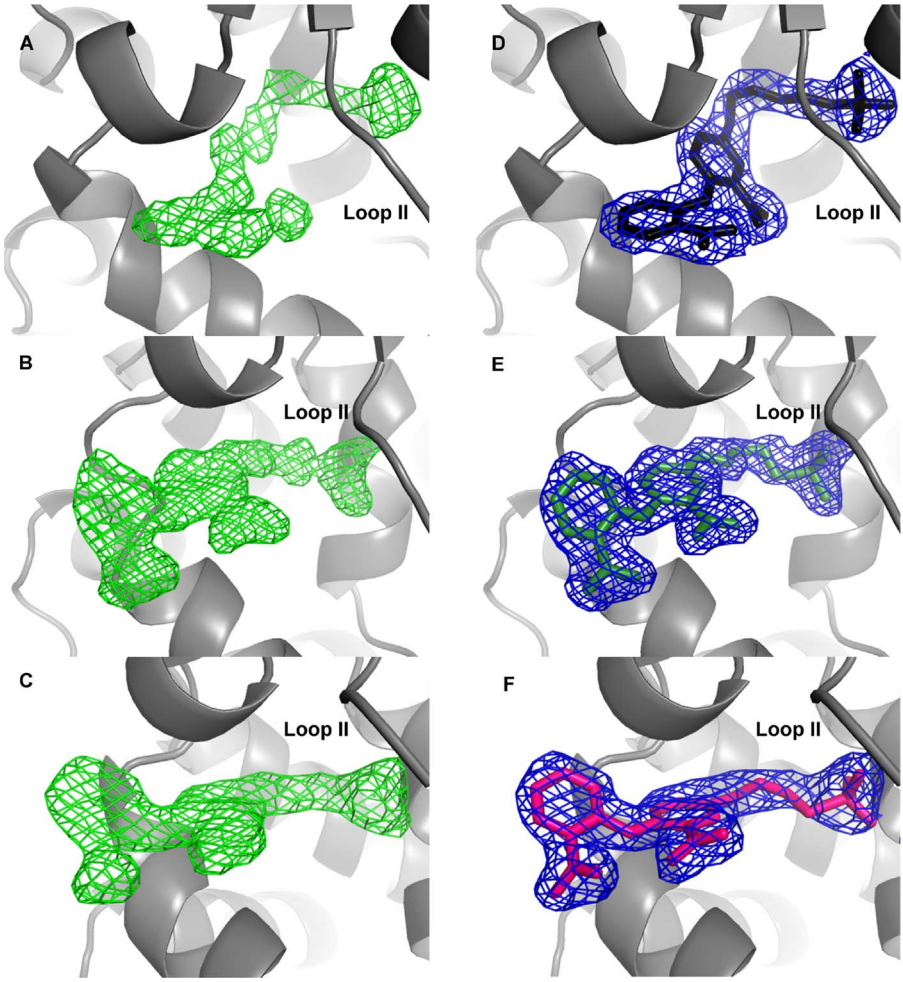


Fig. 2. Omit and fitted map for inhibitors bound in *Mtb*-AnPRT structures. The F_o-F_c map calculated (green, contoured at 3 σ) prior to the addition of ligands to the model for A) 8k, B) 8j and C) 8i (PDB entries 5BNE, 5BO3 and 5BO2, respectively). The $2F_o$-F_c map calculated (blue, contoured at 1 σ) after final refinement, with model including the ligands for D) 8k (black), E) 8j (green) and F) 8i (pink).

2. Experimental design, materials and methods

2.1. Materials

Unless otherwise stated, all chemicals were obtained from Sigma-Aldrich, Scharlau, or Pure Science. The purification of *Mtb*-AnPRT, as well as the synthesis and biochemical characterization of its inhibitors annotated as **8i**, **8j** and **8k** (Table 2), are outlined in the article entitled "*Anthranilate phosphoribosyltransferase: Binding determinants for 5′-phospho-alpha-D-ribosyl-1′-pyrophosphate (PRPP) and the implications for inhibitor design*" [1].

Table 5
Statistics for dataset of *Mtb*-AnPRT co-crystallized with inhibitor 8i.

Data collection	
AnPRT complexed with	**8i**
Space group	*C*2
Cell dimensions	
a, b, c (Å)	94.6, 78.1, 102.8
β (deg)	110.9
Unique reflections[a]	387946 (24179)
Resolution range (Å)[a]	48-1.95 (2.00-1.95)
R_{merge}[a]	0.138 (1.237)
$R_{p.i.m.}$[a]	0.054 (0.488)
Mean $I/\sigma(I)$[a]	9.9 (1.5)
$CC_{1/2}$[a]	0.997 (0.679)
Completeness (%)[a]	99.6 (94.0)
Redundancy[a]	7.6 (7.1)
Wilson B factor	20.5

[b] The average atomic temperature factor. [c] $R_{work}=(|F_{obs}| - |F_{calc}|)/|F_{obs}|$ and $R_{free}=\sum T (|F_{obs}| - |F_{calc}|)/\sum T |F_{obs}|$, where T is a test dataset of 5% of the total reflections randomly chosen and set aside before refinement. [d] RMSD from ideal geometry values from Engh and Huber (1991) [15].

[a] Outer resolution shell is shown in parentheses.

Table 6
Data and refinement statistics for AnPRT complexes with inhibitors.

Data collection			
AnPRT complexed with	**8i**	**8j**	**8k**
PDB code	5BO2	5BO3	5BNE
Space group	*C*2	*C*2	$P2_1$
Cell dimensions			
a, b, c (Å)	94.5, 78.0, 102.9	95.0, 78.1, 102.6	77.3, 78.4, 117.2
β (deg)	111.0	111.1	90.7
Unique reflections[a]	47170 (3415)	68993 (3669)	75079 (4387)
Resolution range (Å)[a]	47-2.00 (2.05-2.00)	48-1.75 (1.78-1.75)	47-2.15 (2.19-2.15)
R_{merge}[a]	0.121 (0.856)	0.121 (1.762)	0.111 (0.711)
$R_{p.i.m.}$[a]	0.076 (0.564)	0.046 (0.688)	0.073 (0.477)
Mean $I/\sigma(I)$[a]	8.5 (1.5)	13.3 (1.2)	6.9 (1.5)
$CC_{1/2}$[a]	0.993 (0.564)	0.998 (0.465)	0.993 (0.502)
Completeness (%)[a]	99.8 (97.7)	97.8 (93.7)	98.4 (77.3)
Redundancy[a]	3.4 (3.3)	7.8 (7.4)	3.0 (2.9)
Wilson B factor	15.1	16.3	21.4
Refinement			
Atoms, B factor (Å^2)[b]			
Protein	4788, 25.8	4746, 24.6	9280, 29.0
Solvent	352, 29.7	402, 29.2	353, 29.5
Ligands	79, 33.2	66, 26.6	107, 30.6
R_{work}/R_{free} (%/%)[a,c]	0.192/0.232 (0.276/0.324)	0.206/0.240 (0.319/0.336)	0.207/0.235 (0.284/0.332)
Ramachandran outliers (%)	0.31	0.16	0.31
R.m.s.d. of			
Bond lengths (Å)[d]	0.003	0.005	0.003
Bond angles (°)[d]	0.774	0.899	0.745

[a] Outer resolution shell is shown in parentheses.
[b] The average atomic temperature factor.
[c] $R_{work}=(|F_{obs}| - |F_{calc}|)/|F_{obs}|$ and $R_{free}=\sum T (|F_{obs}| - |F_{calc}|)/\sum T |F_{obs}|$, where T is a test dataset of 5% of the total reflections randomly chosen and set aside before refinement.
[d] RMSD from ideal geometry values from Engh and Huber [15].

2.2. Crystallization

Crystals of *Mtb*-AnPRT were obtained in hanging drops of 1–2 μL protein solution (3.0–3.1 mg mL^{-1} in final storage buffer) with the equivalent amount of reservoir solution in a fine screen of 0.2 M imidazole-malate pH 7.0–8.5, 5–15% polyethylene glycol (PEG) 4000, as previously reported [2]. Complexes were prepared either by co-crystallization or by soaking with ligands (details in Table 2). The crystallization drops were seeded from previously formed crystals of *Mtb*-AnPRT, by streak seeding using a cat whisker. Data collection occurred within a few days after crystallization. In preparation for data collection, crystals were typically soaked in a reservoir solution containing cryoprotectant 15% PEG4000 and appropriate ligands (details in Table 2), before being flash-cooled in liquid nitrogen.

2.3. Data collection, structure solution and refinement

Data were collected at the Australian Synchrotron, Beamlines MX1 and MX2 at 110 K. The crystals of the *Mtb*-AnPRT complexes diffracted to a maximum resolution that varied between 1.75 and 2.15 Å. X-ray diffraction spots were indexed and integrated with XDS [14] and scaled with AIMLESS [11]. The high resolution cut-off was determined based on a correlation coefficient ($CC_{1/2}$) [16] exceeding 0.5, with mean I/σ between 1 and 1.5, and a $R_{p.i.m.}$ of 0.7 or less, as calculated by AIMLESS [11]. $R_{p.i.m.}$ is defined as $\sum hkl\ [1/(n-1)]\times R_{merge}$, with R_{merge} defined as $\sum hkl|I(hkl)- < I(hkl) > |/\sum hkl < I(hkl) >$, where $< I(hkl) >$ is the mean of symmetry-equivalent reflections of $I(hkl)$ [17]. Data and refinement statistics are given in Tables 5 and 6.

Mtb-AnPRT (chain A of 3QR9 [2]), without solvent or ligands, and with loops I and II removed, was used as the search model for structure determination by molecular replacement using Phaser [18]. Refinement and model building was performed with COOT [19], Refmac5 [20], and PHENIX [21]. After positioning waters, omit maps were examined (Fig. 2) and ligands placed, with subsequent refinement including restraints for the ligands generated by phenix.elbow [21]. Restraints on protein bond lengths and angles were based on the ideal values of Engh and Huber [15] and model quality was assessed using MolProbity [22]. Figures illustrating structural details were prepared using PyMOL.

The F_o-F_c map calculated (green, contoured at 3 σ) prior to the addition of ligands to the model for A) **8k**, B) **8j** and C) **8i** (PDB entries 5BNE, 5BO3 and 5BO2, respectively). The $2F_o$-F_c map calculated (blue, contoured at 1 σ) after final refinement, with model including the ligands for D) **8k** (black), E) **8j** (green) and F) **8i** (pink).

Acknowledgements

This work was funded by grants from FRST (Foundation for Research, Science and Technology grant number UOAX1005) and the Health Research Council of New Zealand (Programme award 12/1111). Coco the cat is acknowledged for her contribution of a whisker for protein crystallization. We also thank Dr Farah Javid-Majd and Prof. Jim Sacchettini from the Department of Biochemistry and Biophysics, Texas A&M University for the original gift of the *Mtb*-AnPRT expression plasmid. This work was performed as part of the *M. tuberculosis* Structural Genomics Consortium (http://www.webtb.org). The authors also wished to thank the anonymous reviewer who suggested examination/explanation of various space groups observed in *Mtb*-AnPRT structures.

Transparency document. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2017.10.051.

References

[1] G.L. Evans, D.P. Furkert, N. Abermil, P. Kundu, K.M. de Lange, E.J. Parker, M.A. Brimble, E.N. Baker, J. Shaun Lott, Anthranilate phosphoribosyltransferase: Binding determinants for 5'-phospho-alpha-D-ribosyl-1'-pyrophosphate (PRPP) and the implications for inhibitor design, Biochim. Biophys. Acta (2017), http://dx.doi.org/10.1016/j.bbapap.2017.08.018 (In press) (PII: S1570-9639(17)(30204-2)).

[2] A. Castell, F.L. Short, G.L. Evans, T.V. Cookson, E.M. Bulloch, D.D. Joseph, C.E. Lee, E.J. Parker, E.N. Baker, J.S. Lott, The substrate capture mechanism of *Mycobacterium tuberculosis* anthranilate phosphoribosyltransferase provides a mode for inhibition, Biochemistry 52 (2013) 1776–1787. http://dx.doi.org/10.1021/bi301387m.

[3] S.W. Wukovitz, T.O. Yeates, Why protein crystals favour some space-groups over others, Nat. Struct. Biol. 2 (1995) 1062–1067.

[4] C.E. Lee, C. Goodfellow, F. Javid-Majd, E.N. Baker, J. Shaun Lott, The crystal structure of TrpD, a metabolic enzyme essential for lung colonization by *Mycobacterium tuberculosis*, in complex with its substrate phosphoribosylpyrophosphate, J. Mol. Biol. 355 (2006) 784–797. http://dx.doi.org/10.1016/j.jmb.2005.11.016.

[5] G.L. Evans, S.A. Gamage, E.M. Bulloch, E.N. Baker, W.A. Denny, J.S. Lott, Repurposing the chemical scaffold of the anti-arthritic drug lobenzarit to target tryptophan biosynthesis in *Mycobacterium tuberculosis*, Chembiochem 15 (2014) 852–864. http://dx.doi.org/10.1002/cbic.201300628.

[6] T.V. Cookson, G.L. Evans, A. Castell, E.N. Baker, J.S. Lott, E.J. Parker, Structures of *Mycobacterium tuberculosis* anthranilate phosphoribosyltransferase variants reveal the conformational changes that facilitate delivery of the substrate to the active site, Biochemistry 54 (2015) 6082–6092. http://dx.doi.org/10.1021/acs.biochem.5b00612.

[7] T.V. Cookson, A. Castell, E.M. Bulloch, G.L. Evans, F.L. Short, E.N. Baker, J.S. Lott, E.J. Parker, Alternative substrates reveal catalytic cycle and key binding events in the reaction catalysed by anthranilate phosphoribosyltransferase from *Mycobacterium tuberculosis*, Biochem. J. 461 (2014) 87–98. http://dx.doi.org/10.1042/bj20140209.

[8] M. Marino, M. Deuss, D.I. Svergun, P.V. Konarev, R. Sterner, O. Mayans, Structural and mutational analysis of substrate complexation by anthranilate phosphoribosyltransferase from *Sulfolobus solfataricus*, J. Biol. Chem. 281 (2006) 21410–21421. http://dx.doi.org/10.1074/jbc.M601403200.

[9] E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2256–2268. http://dx.doi.org/10.1107/S0907444904026460.

[10] A. Ortega, D. Amoros, J. Garcia de la Torre, Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models, Biophys. J. 101 (2011) 892–898. http://dx.doi.org/10.1016/j.bpj.2011.06.046.

[11] P.R. Evans, An introduction to data reduction: space-group determination, scaling and intensity statistics, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 282–292. http://dx.doi.org/10.1107/s090744491003982x.

[12] A.A. Lebedev, M.N. Isupov, Space-group and origin ambiguity in macromolecular structures with pseudo-symmetry and its treatment with the program Zanuda, Acta Crystallogr. D Biol. Crystallogr. 70 (2014) 2430–2443. http://dx.doi.org/10.1107/S1399004714014795.

[13] P.H. Zwart, R.W. Grosse-Kunstleve, A.A. Lebedev, G.N. Murshudov, P.D. Adams, Surprises and pitfalls arising from (pseudo) symmetry, Acta Crystallogr. D Biol. Crystallogr. 64 (2008) 99–107. http://dx.doi.org/10.1107/S090744490705531X.

[15] R.A. Engh, R. Huber, Accurate bond and angle parameters for X-ray protein-structure refinement, Acta Crystallogr. Sect. A 47 (1991) 392–400. http://dx.doi.org/10.1107/S0108767391001071.

[14] W. Kabsch, XDS, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132. http://dx.doi.org/10.1107/S0907444909047337.

[16] P.A. Karplus, K. Diederichs, Linking crystallographic model and data quality, Science 336 (2012) 1030–1033. http://dx.doi.org/10.1126/science.1218231.

[17] M.S. Weiss, Global indicators of X-ray data quality, J. Appl. Crystallogr. 34 (2001) 130–135. http://dx.doi.org/10.1107/S0021889800018227.

[18] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674. http://dx.doi.org/10.1107/S0021889807021206.

[19] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501. http://dx.doi.org/10.1107/S0907444910007493.

[20] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Crystallogr. D Biol. Crystallogr. 53 (1997) 240–255. http://dx.doi.org/10.1107/S0907444996012255.

[21] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213–221. http://dx.doi.org/10.1107/s0907444909052925.

[22] V.B. Chen, W.B. Arendall 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D. C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 12–21. http://dx.doi.org/10.1107/s0907444909042073.