1 # Inhibiting *Mycobacterium tuberculosis* CoaBC by targeting a new allosteric

2 site.

3

4 Vitor Mendes[1*], Simon R. Green[2], Joanna C. Evans[3], Jeannine Hess[4], Michal Blaszczyk[1],

5 Christina Spry[4], Owain Bryant[1], James Cory-Wright[1], Daniel S-H. Chan[4], Pedro H.M.

6 Torres[1], Zhe Wang[5], Sandra O'Neill[2], Sebastian Damerow[2], John Post[2], Tracy Bayliss[2],

7 Sasha L. Lynch[3], Anthony G. Coyne[4], Peter C. Ray[2], Chris Abell[4], Kyu Y. Rhee[5], Helena

8 I. M. Boshoff[6], Clifton E. Barry III[3,6], Valerie Mizrahi[3], Paul G. Wyatt[2], Tom L.

9 Blundell[1*].

10

11 1 Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge, CB2 1GA,

12 UK

13 2 Drug Discovery Unit, College of Life Sciences, University of Dundee, Dow Street, Dundee, DD1 5EH,

14 Scotland, UK

15 3 MRC/NHLS/UCT Molecular Mycobacteriology Research Unit & DST/NRF Centre of Excellence for

16 Biomedical TB Research & Wellcome Centre for Infectious Diseases Research in Africa, Institute of

17 Infectious Disease and Molecular Medicine and Department of Pathology, Faculty of Health Sciences,

18 University of Cape Town, Anzio Road, Observatory 7925, South Africa

19 4 Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK

20 5 Division of Infectious Diseases, Weill Department of Medicine, Weill Cornell Medical College, New

21 York, NY 10065, USA

22 6 Tuberculosis Research Section, Laboratory of Clinical Immunology and Microbiology, National Institute

23 of Allergy and Infectious Disease, National Institutes of Health, 9000 Rockville Pike, Bethesda, Maryland

24 20892, USA

25

26 * To whom correspondence should be addressed

27   Vitor Mendes: vgm23@cam.ac.uk

28   Tom L. Blundell: tom@cryst.bioc.cam.ac.uk

Abstract

Coenzyme A (CoA) is a fundamental co-factor for all life, involved in numerous metabolic pathways and cellular processes, and its biosynthetic pathway has raised substantial interest as a drug target against multiple pathogens including *Mycobacterium tuberculosis*. The biosynthesis of CoA is performed in five steps, with the second and third steps being catalysed in the vast majority of prokaryotes, including *M. tuberculosis*, by a single bifunctional protein, CoaBC. Depletion of CoaBC was found to be bactericidal in *M. tuberculosis*. Here we report the first structure of a full length CoaBC, from the model organism *Mycobacterium smegmatis*, describe how it is organised as a dodecamer and regulated by CoA thioesters. A high-throughput biochemical screen focusing on CoaB identified two inhibitors with different chemical scaffolds. Hit expansion led to the discovery of potent inhibitors of *M. tuberculosis* CoaB. Crucially, we further show that these compounds bind to a novel cryptic allosteric site within CoaB.

Introduction

Tuberculosis (TB) is the most prevalent and deadly infectious disease worldwide and remains a global epidemic. Despite the availability of treatment, this disease, caused by *Mycobacterium tuberculosis*, still claims 1.5 million lives each year [1]. Current treatment regimens are long, which presents an obstacle for patient adherence and imposes a heavy social and economic burden on countries with a high incidence of TB. It is therefore critical to explore novel targets and find new and more effective drugs to combat this disease.

Coenzyme A (CoA) is an essential and ubiquitous cofactor, involved in numerous metabolic pathways with a large number of different enzymes requiring it for activity[2]. CoA is essential for the synthesis of phospholipids, fatty acids, polyketides, and non-ribosomal peptides, for the operation of the tricarboxylic acid cycle and in the degradation of lipids [3]. The importance of CoA for essential post-translational modifications of proteins is also well established in both eukaryotes and prokaryotes with various proteins post-translationally modified by thioester derivatives of CoA (acylation) or CoA itself (phosphopantetheinylation and CoAlation), while several other post-translational modifications depend indirectly on CoA through the mevalonate pathway [4-7]. Furthermore, dephospho-CoA, an intermediate of the CoA pathway, is incorporated into some RNA transcripts during transcription initiation thereby serving as a non-canonical transcription initiating nucleotide [8]. These RNA modifications have functional consequences and occur in both eukaryotes and bacteria [8]. In *M. tuberculosis*, CoA plays a pivotal role in the biosynthesis of complex lipids that are crucial components of the cell wall and required for pathogenicity [9]. It is also needed for the degradation of lipids, including cholesterol, which are the primary source of energy for this organism during infection [10,11]. The CoA pathway is therefore an attractive pathway for drug discovery for

79   many different infectious diseases, including TB, given its ubiquitous nature, wide

80   metabolic and functional impact of its inhibition and lack of sequence conservation

81   between prokaryotes and humans.

82   The biosynthesis of CoA from pantothenic acid (vitamin B_5) is performed in five steps,

83   sequentially catalysed by the enzymes pantothenate kinase (CoaA, also known as PanK),

84   phosphopantothenoylcysteine   synthetase   (CoaB),   phosphopantothenoylcysteine

85   decarboxylase (CoaC), phosphopantetheine adenylyltransferase (CoaD) and dephospho-

86   CoA kinase (CoaE). However, in the vast majority of prokaryotes, including *M.*

87   *tuberculosis*, CoaB and CoaC are encoded by a single gene to produce a fused

88   bifunctional enzyme (CoaBC). Transcriptional silencing of individual genes of the CoA

89   biosynthesis pathway of this pathogen identified CoaBC as uniquely bactericidal within

90   the CoA pathway, highlighting it as a good candidate for drug discovery [12].

91   CoaBC comprises two enzymes in a single polypeptide chain and converts 4'-

92   phosphopantothenate   to   4'-phosphopantetheine   in   three   steps.   First,   4'-

93   phosphopantothenate (*P*PA) reacts with CTP to form 4'-phosphopantothenoyl-CMP with

94   the release of pyrophosphate. This intermediate subsequently reacts with cysteine to form

95   4'-phosphopantothenoylcysteine (*P*PC) with the release of CMP with these two steps

96   being catalysed by CoaB. The product of CoaB is then decarboxylated by CoaC to 4'-

97   phosphopantetheine an enzyme of the homo-oligomeric flavin-containing decarboxylase

98   (HFCD) protein family. X-ray crystal structures have been reported for the individual

99   CoaB and CoaC enzymes in several organisms, including a structure of CoaB from

100   *Mycobacterium smegmatis*, a close relative of *M .tuberculosis* However, a structure of

101   the full-length bifunctional CoaBC was missing.

102   Here we report the structure of the bifunctional CoaBC of *M. smegmatis* at 2.5 Å. We

103   identify a previously unknown allosteric site in CoaB and crucially, we report the

104   identification of the first *M. tuberculosis* CoaBC allosteric inhibitors. Using X-ray

105   crystallography and enzyme kinetic experiments, we define the mode of binding of one

106   of the inhibitors and show its impact on the protein structure and function. These results

107   further illustrate the potential of CoaBC as a novel drug target in *M. tuberculosis*.

108

109   **Results**

110   **Overall structure of CoaBC**

111   Because the HFCD protein family of flavin-binding proteins are known to form homo-

112   oligomers [13], we performed native electrospray-ionization mass spectrometry (ESI-MS)

113   to investigate the stoichiometry of CoaBC, previously proposed to form a dodecamer [13].

114   Both *M. tuberculosis* CoaBC (MtbCoaBC) (Figure S1A) and *M. smegmatis* CoaBC

115   (MsmCoaBC) (Figure S1B) exclusively exhibited a dodecameric assembly, with no other

116   oligomeric species observed in the spectra, consistent with a strong interaction between

117   the subunits of the complex. The dodecamer of MtbCoaBC was centred around the 56+

118   charge state, with an observed mass of 537 kDa, while the dodecamer of MsmCoaBC was

119   centred around the 52+ charge state, with an observed mass of 523 kDa. These masses

120   are 1-2% higher than the expected masses of 525 and 518 kDa for MtbCoaBC and

121   MsmCoaBC respectively, which can be attributed to the non-specific binding of solvent

122   molecules or ions to the protein complexes under the soft ionization conditions employed.

123   Structures of a few proteins of the HFCD family have been determined [14-18]. All of these

124   structures show either a homo-trimeric or homo-dodecameric arrangement of the flavin-

125   containing Rossmann-fold with trimers forming at each of the vertices of the tetrahedron

126   in the case of a dodecameric arrangement [15]. However, all of these proteins, unlike

127   CoaBC, contain only a single functional domain. We solved the structure of MsmCoaBC

128   at 2.5 Å resolution (Figure 1A), in the presence of CTP and FMN (Figure 1B, S2A and

129   S2B), using crystals belonging to the $H3_2$ space group with an asymmetric unit containing

130   four protomers forming two CoaBC dimers. Data collection and refinement statistics are

131   summarised in (Table S1). The final model (residues 2-412) covers both CoaC and CoaB,

132   but densities for several residues in three loop regions in CoaB are not observed (residues

133   290-298, 336-342, 363-376). However, all these residues except for 375 and 376 can be

134   seen the MsmCoaB X-ray crystal structure (PDB: 6TH2) that we also solved in this work

135   at 1.8 Å. The N-terminal CoaC of MsmCoaBC (residues 1-179) forms the same type of

136   dodecameric arrangement seen in other HFCD family proteins such as the peptidyl-

137   cysteine decarboxylase EpiD [15] and sits at the core of the dodecamer (Figure 1A and 1C)

138   with the two domains connected through a small loop region (residues 180-189) that

139   tightly interacts with both. The active site of CoaC sits at the interface between two

140   protomers of the CoaC trimer and a protomer of an adjacent CoaC trimer with the FMN

141   site facing inwards towards the hollow centre of the dodecamer (Figure 2A). A flexible

142   flap that was observed before covering a reaction intermediate bound to *Arabidopsis*

143   *thaliana* CoaC [19] is also observed in some of the protomers, but in an open conformation

144   (Figure 1B).

145   The C-terminal CoaB of MsmCoaBC also displays a Rossmann fold consistent with the

146   several other CoaB structures solved previously, including both the eukaryotic form, in

147   which CoaB exists as an individual polypeptide, and the bacterial form where CoaB is

148   typically fused with CoaC [20-22]. Each CoaB of MsmCoaBC (residues 190-414) dimerises

149   with a CoaB belonging to an adjacent trimer (Figure 1C). The full protein resembles a

150   tetrahedron with CoaB dimers positioned at the six edges and CoaC trimers at the four

151   vertices (Figure 1A).

152   The shortest distance between a pair of CoaB and CoaC active sites is ~30 Å (Figure 2B).

153   Nevertheless, a flexible loop (residues 362-377) that covers the 4'-phosphopantothenate

154   site when this substrate binds to the enzyme [20] can be seen in our MsmCoaB structure,

155   extending away from the active site. A superposition of our MsmCoaB dimer structure

156   with MsmCoaBC shows the loop extending towards the CoaC active site (Figure 2B).

157   This long loop (15-16 amino acids) is present in all CoaBCs (Figure S3) and it is possible

158   that it helps channel the substrate from the CoaB to the CoaC active site.

159   The small differences (RMSD = 1.147) in overall structure of CoaB dimers in the full

160   length MsmCoaBC and the MsmCoaB crystal structure solved at 1.8 Å (PDB: 6TH2) can

161   be attributed to artefacts of crystal packing (Figure S4). Similarly, the CoaC structure

162   does not seem to differ between full length MsmCoaBC and the available individual CoaC

163   structures. However, when MsmCoaB (residues 186-414) is expressed alone, the protein

164   does not dimerise in solution and is inactive (not shown). This contrasts with *E. coli*

165   CoaB, which still dimerises and is functional when expressed on its own without the N-

166   terminal CoaC [23,24]. The CoaB dimer interface is mostly conserved, but there are clear

167   differences in the dimerisation region between MsmCoaB and *E. coli* CoaB that could

168   help to explain the different observed behaviours (Figure S3). The absence of

169   dimerization for the MsmCoaB when expressed alone suggests that the interactions

170   between CoaC and CoaB in *M. smegmatis,* and likely all other *Mycobacteriaceae*, are

171   fundamental for CoaB dimerisation and activity. This idea is reinforced by the fact that

172   the residues located at the interface of the two enzymes (CoaB and CoaC) are well

173   conserved in all *Mycobacteriaceae* and somewhat conserved in the sub-order

174   *Corynebacterineae* but not outside of this group (Figure S3).

175   The CoaB dimerisation region forms a β-sandwich composed of eight anti-parallel β-

176   strands, related by 2-fold symmetry, that contacts with the active site (Figure 2C and 2D).

177   Comparison of the MsmCoaB with human CoaB reveals that the human and many other

178   eukaryotic CoaBs [21] possess two extra α-helices and β-strands involved in the

179   dimerisation interface that help stabilise the dimer in the absence of CoaC (Figure S5).

180   The CoaB active site is enclosed by a loop that extends from the opposing protomer and

181   is observed for the first time in this work. This loop contains a motif "K-X-K-K", which

182   is widely conserved in bacteria (Figure S3), with few exceptions, and each lysine either

183   interacts directly with the triphosphate group of CTP or through highly coordinated

184   waters (Figure 2C). Also observed for the first time is the coordination of a cation by the

185   triphosphate group and D281 (Figure 2C and S6). While magnesium or manganese are

186   the favoured cations for CoaB activity [25], calcium is observed in our structures instead,

187   due to the high concentration present in the crystallization condition.

188

189   **CoaBC is regulated by CoA thioesters**

190   It is known that CoA biosynthesis is regulated, in many organisms, at the first step of the

191   pathway, catalysed by the enzyme CoaA [3]. *M. tuberculosis* and many other mycobacteria

192   possess two isoforms of this enzyme (type I and III), however only the type I isoform

193   seems to be active based on studies in *M. tuberculosis* [26]. CoA and its thioesters

194   competitively inhibit this enzyme by binding to the ATP site with CoA being the strongest

195   regulator [27,28]. Nevertheless, at physiologically relevant levels of CoA there is only a low

196   level inhibition of CoaA [28]. It is also known that CoaD, the enzyme that catalyses the

197   fourth step of the pathway, is competitively inhibited by CoA and its product dephospho-

198   CoA [29,30]. However, nothing was known about the regulation of CoaBC. We therefore

199   examined the effect of CoA and several of its thioesters (acetyl-CoA, malonyl-CoA and

200   succinyl-CoA) on MtbCoaBC activity, using a coupled enzymatic assay that quantifies

201   the release of pyrophosphate (EnzCheck pyrophosphate assay). Controls performed to

202   assess the activity of these compounds against the two coupling enzymes and the

203   compounds showed an absence of inhibition at the tested range of concentrations.

204   Potent regulation of CoaB activity by CoA and acyl-CoAs was observed, with IC_{50} values

205   ranging from 38 to 148 µM, far below the predicted intracellular concentrations of acyl-

206   CoAs [31], with succinyl-CoA displaying the highest inhibition (Figure 3A and Table 1).

207   Competition assays with the three substrates and acetyl-CoA show a competitive mode

208   of inhibition relative to CTP and *P*PA, and non-competitive inhibition for L-cysteine

209   (Figure 3B). In the absence of a crystal structure to confirm the mode of binding, these

210   results suggest that acyl-CoAs most likely bind to the active site itself, competing directly

211   with CTP and *P*PA. Interestingly both acyl-CoAs, involved in fatty acid synthesis, as well

212   as those, involved in the TCA cycle, show inhibition of CoaB with larger fatty acyl chains

213   showing higher inhibition of CoaB.

214

215   **Identification of CoaB inhibitors using high-throughput screening**

216   Although the CoA biosynthetic pathway is considered an attractive target for drug

217   discovery, CoA pathway inhibitors displaying potent whole cell activity are rare and the

218   few CoaBC inhibitors that have been reported to date are in majority substrate mimicking

219   [24,32].

220   In order to identify novel MtbCoaBC inhibitors, we have conducted a high-throughput

221   screen of 215,000 small molecules targeting CoaB activity. To do this an end-point

222   pyrophosphate quantification assay was used (Biomol Green). The most potent hits

223   identified were compounds **1a** and **2a** with IC_{50} values of 9 and 3.1 µM respectively

224   (Table S2), originating from two different but related chemical scaffolds. A search was

225   then performed for commercially available analogues. Testing of analogues of the initial

226   hits resulted in the identification of more potent compounds with sub-micromolar IC_{50}

227   values (Table 1, Figure 4A, Table S2 and Figure S7). Of these, compounds **1b** and **2b**

228   (Figure 4A and Table 1), with IC$_{50}$ values of 0.28 and 0.08 μM respectively, were

229   identified as the most potent of the two chemical series and therefore were selected for

230   further work.

231

232   **Elucidation of the mode of inhibition**

233   Following the identification of potent MtbCoaB inhibitors we went on to determine their

234   mode of inhibition using kinetic assays. For this, the EnzCheck coupled enzyme assay

235   that measures the release of pyrophosphate was used. Control experiments were first

236   performed to assess compound activity against the two coupling enzymes and the

237   compounds were found to be inactive at 100 μM. The IC$_{50}$ values for the compounds

238   against CoaB were reassessed with this assay and the values obtained were in line with

239   the primary screening assay (Table S2).

240   Competition experiments were then performed using the three CoaB substrates for the

241   two most potent compounds of each chemical series (**1b** and **2b**). Compound **1b** was

242   found to preferentially bind to the enzyme-substrate complex with uncompetitive

243   inhibition observed for all substrates (Figure 4B), consistent with the compound binding

244   preferentially when the three substrates are bound. Compound **2b** shows mixed inhibition

245   relative to CTP and uncompetitive inhibition for L-cysteine and *P*PA (Figure 4C). It is

246   known that CoaB forms the phosphopantothenoyl-CMP intermediate in the absence of L-

247   cysteine [33] and it is likely, due to spatial constraints, that cysteine can only bind at the

248   active site after the release of pyrophosphate. The data is therefore consistent with

249   compound **1b** preferentially binding after L-cysteine enters the active site, for to the last

250   step of catalysis and the formation of 4'-phosphopantothenoylcysteine and CMP.

251   However, compound **2b** shows a mixed inhibition for CTP, reflecting a slightly different

252   mechanism of action. These results obtained for both compounds suggest the existence

253   of an allosteric site in the CoaB moiety of MtbCoaBC.

254

255   **Structural basis for inhibition of CoaB by allosteric inhibitors**

256   In order to elucidate the binding mode of compound **1b** we used a truncation of the

257   MsmCoaB (residues 187-414) that was previously crystallized before by others in the

258   presence of CTP (PDB code: 4QJI) at 2.65 Å resolution. The screening for new

259   crystallization conditions allowed us to find a new CTP containing condition that gave

260   excellent resolution (1.8 Å). Comparison of this structure with the full length MsmCoaBC

261   showed only minor differences that can be attributed to crystal packing (Figure S4).

262   Hence this crystallization system could be used to validate CoaB inhibitors binding

263   outside of the CTP site.

264   MsmCoaB was co-crystallized with CTP in the absence of compound **1b** and overnight

265   soaking of the crystals with this compound was performed. A Co-crystal structure of

266   MsmCoaB with compound **1b** was obtained and showed that the compound was bound

267   to a site at the dimer interface of CoaB, in a deep cavity that is occluded when the

268   compound is absent (Figure 5A-B and S8). Each CoaB dimer contains two of these sites,

269   which are formed by a sandwich of eight β-strands and a long loop that contains the

270   conserved "K-X-K-K" motif. This site opens to the active site and the inhibitor also

271   contacts with D281 that is involved in the coordination of the cation (Figure 5C). The

272   opening/closing of this cryptic allosteric site is mediated by the side chain of R207 of the

273   opposing protomer (Figure 5D) that moves 5.5 Å at the furthest point and to a smaller

274   extent by the side chain of F282 that moves 2 Å. R207 has previously been shown to be

275   critical for the second half of the reaction catalysed by CoaB, the conversion of the 4'-

276   phosphopantothenoyl-CMP intermediate to *P*PC, with almost no conversion of the

277   intermediate to *P*PC detected when this arginine is mutated to glutamine [33]. Given the

278   position of this arginine, it is likely that it is involved in the binding of cysteine. Despite

279   the absence of a crystal structure with cysteine, kinetic data showing uncompetitive

280   inhibition with cysteine is consistent with this.

281   This allosteric site is comprised of a large group of hydrophobic residues (I209, F282

282   L304 of protomer A and L203, I292, P299 and I302 of protomer B) many of which form

283   hydrophobic interactions with compound **1b** (Figure 5C). Several π-interactions between

284   the compound and the protein are also observed and involve D281 and F282 of protomer

285   A and R207 of protomer B (Figure 5D). Hydrogen-bond interactions are formed with

286   D281 and F282 of protomer A and R207 of protomer B. Water-mediated interactions are

287   also observed for a group of residues that sit at the outer edge of the site (L203, H286 and

288   D303) that is formed exclusively by protomer B (Figure 5C).

289   We propose that upon binding of L-cysteine, the R207 side chain moves towards the

290   active site, and is likely involved in stabilizing/orienting L-cysteine to attack the

291   phosphopantothenoyl-CMP intermediate. This movement opens the allosteric site, which

292   allows binding of allosteric inhibitors to the newly created cavity. The allosteric inhibitors

293   will then stabilize the enzyme in its substrate bound state with the position of R207

294   becoming locked by several hydrogen-bonds with the side chain of D281 of protomer A,

295   the backbone carbonyl group of I292 and the side chain of D204 of protomer B but also

296   by the π-interactions with the compound (Figure 5C). The residues around this site and

297   crucially R207 are conserved across many microorganisms, suggesting that this allosteric

298   site is present in most, if not all bacterial CoaBCs (Figure S3 and S9A). Interestingly,

299   even though overall sequence identity is very low between the human CoaB and

300   MsmCoaB (22%), the human enzyme also contains an arginine equivalent to R207 and a

301 roughly similar interface with several conserved residues, but there are stark differences

302 in the relative position of the residues at this site between the two enzymes (Figure S9B).

303 While we were not able to obtain co-crystal structures with other inhibitors, *in silico*

304 docking helped to provide a possible explanation for the structure-activity relationship

305 observed for series one and two. The highest-scoring docking pose of compound **1b** the

306 most potent inhibitor of series one, was almost identical to that observed in the co-crystal

307 structure (Figure S10A), and the analogues for which docking was performed adopted a

308 similar binding pose. The lower activity of compound **1a** relative to compound **1b** could

309 be explained by the loss of water-mediated hydrogen bonds (Figure 5C, S10B), while the

310 lower activity of compound **1c** could be explained by the loss of the carbonyl group which

311 faces a highly electropositive area of the protein (Figure S10C). Compound **2b** is

312 predicted to form direct hydrogen bonds at the bottom of the allosteric site, similar to

313 those formed by compound **1b** but also to interact directly with L203 and H286, forming

314 extra hydrogen bonds at the top of the allosteric site (Figure 6). For compound **1b** the

315 interactions at the top of the site are water mediated (Figure 5C). This could explain the

316 higher potency of compound **2b**. Compounds **2c** and **2d** are also predicted to form direct

317 hydrogen bond interactions at the top of the allosteric site, but the interactions at the

318 bottom of the site are not as favourable due to the presence of extra hydroxyl groups

319 (Figure S10D-F). The remaining compounds in series two, which have fewer hydroxyl

320 groups and/or with hydroxyl groups in different positions, lose the ability to form

321 hydrogen bonds, consistent with the weaker inhibitory effect observed (Table S2).

322

323 **Screening of CoaBC inhibitors against *M. tuberculosis***

324 The in vitro whole cell activity of the compounds was further evaluated by their ability

325 to inhibit *M. tuberculosis* growth on different carbon sources. None of the compounds

326   exhibited activity in media containing glycerol or cholesterol as the main carbon source.

327   We then tested whether the lack of inhibitory activity could be attributed to the presence

328   of BSA by determining the whole cell activity of the three most potent inhibitors against

329   *M. tuberculosis* in GAST/Fe minimal media. All the tested compounds exhibited

330   moderate to low activity in this media with compound **2b** displaying the best activity of

331   the three (Table 2). The observed differences in potency between the enzymatic assay and

332   whole cell activity are likely related to low compound permeation, high efflux or

333   metabolisation.

334

335   **Discussion**

336   CoA is an essential co-factor ubiquitous in all domains of live. For many years, this

337   pathway has been considered an attractive drug target to develop new antibiotics against

338   a wide range of pathogens including *M. tuberculosis*. Furthermore, the recent

339   identification of CoaBC as a key fragility point in the CoA pathway of this organism [12],

340   combined with the extremely low sequence identity with the human CoaB (25%) makes

341   this enzyme an highly attractive drug discovery target.

342   However, we were aware that the many questions remaining at the start of this work about

343   the organization and regulation of this bi-functional enzyme could have significant

344   implications for drug discovery. We therefore set out to obtain a full-length structure of

345   a CoaBC from mycobacteria and we were successful in solving the full length

346   MsmCoaBC but also MsmCoaB alone, which share very high sequence identity with the

347   *M. tuberculosis* orthologue (86% full-length, 84% CoaB enzyme) and hence valuable

348   tools for studying *M. tuberculosis* CoaBC. The CoaBC organization we revealed is

349   similar to other HFCD family proteins [15] but unique in the sense that it contains more than

350   one domain and highlights how the arrangement of the fused enzymes is essential for

351   mycobacterial CoaB dimerisation and function.

352   Regulation of the CoA biosynthesis pathway by its product (CoA) was known to occur

353   for other enzymes of the pathway but no information was available for CoaBC. We

354   demonstrate that both CoA as well as several CoA thioesters regulate CoaBC by

355   inhibiting CoaB activity with succinyl-CoA showing the greatest inhibition of the tested

356   acyl-CoAs, and that these molecules act in a competitive manner with CTP and PPA and

357   a non-competitive manner with L-cysteine. This is consistent with these molecules

358   binding to the CoaB active site but not to the L-cysteine sub-site.

359   CoA and acyl-CoAs inhibit both CoaA and CoaD enzymes to varying extents [28,30,34].

360   However, the inhibitory effect of CoA and its thioesters in their activity is low when

361   compared to what we observed in CoaBC and consequently the impact of the intracellular

362   level of these molecules will be higher in CoaBC . We therefore report a new and

363   important mechanism of regulation of de novo CoA biosynthesis, mediated by the action

364   of CoA thioesters on CoaBC. Since the reported intracellular levels of these molecules [31]

365   are normally above the observed IC_{50} the activity of CoaBC is highly inhibited. This

366   correlates well with previous work showing that "de novo" CoA biosynthesis closely

367   matches dilution due to cell division [35]. However, the data for intracellular concentrations

368   of CoA and CoA thioesters as well CoA half-life was not obtained for mycobacteria and

369   both interspecies differences as well as variations in growth conditions may affect these

370   conclusions.

371   Although the CoA pathway and CoaBC have been the subject of many drug discovery

372   efforts, few non-substrate-mimicking inhibitors of CoaBC have been reported [24]. Our

373   work identifies two related chemical scaffolds that potently inhibit the activity of the

374   CoaB moiety of MtbCoaBC through a new cryptic allosteric site that sits in the dimer

375　interface region of the CoaB enzyme. This site is closed in the CTP-bound structure, by

376　the side chain of R207 a residue known to be involved in the second and final step of the

377　reaction catalysed by CoaB – the conversion of the 4'-phosphopantothenoyl-CMP

378　intermediate to *P*PC [33]. Considering the role of this in the final step of product formation

379　and that compound **1b** shows uncompetitive inhibition relative to all CoaB substrates, we

380　propose that the opening of this site occurs upon binding of the final substrate L-cysteine.

381　Currently it is not clear whether this new allosteric site is exploited by a natural ligand,

382　as we were unable to identify such a biomolecule. Nevertheless, the conservation of

383　residues at this site, across a variety of bacteria, indicates that this feature might be

384　common to many, if not all, bacterial CoaBs.

385　Drug discovery against *M. tuberculosis* is rich in examples of compounds with potent

386　activity against an essential enzyme but with a complete lack of whole cell activity due

387　to the impermeable cell wall of this organism, efflux pumps, target modification enzymes

388　and extensive capacity to metabolise compounds [36]. The lack of in vitro whole cell activity

389　displayed by the CoaB inhibitors reported in this work, may relate to any of these issues.

390　However, the activity observed in albumin free media (GAST media) also points to the

391　possibility of compound binding to albumin interfering with the assay. Nevertheless, the

392　promising biochemical and structural data described herein further validates CoaBC as a

393　promising novel anti-tubercular drug target by showing a new allosteric site that can be

394　targeted by potent inhibitors.

395

396　**Materials and methods**

397　**Cloning and protein purification**

398　*Mycobacterium tuberculosis* and *Mycobacterium smegmatis coaBC* genes were amplified

399　from genomic DNA of *M. tuberculosis* H37Rv strain, obtained from ATCC

400   (ATCC25618D-2) and genomic DNA of *M. smegmatis* mc^2 155 (graciously provided by

401   Dr. Nuno Empadinhas), and cloned into a pET28a vector (Novagen), modified to include

402   an N-terminal 6xHis-SUMO tag. *M. smegmatis coaB* construct was obtained from Seattle

403   Structure Genomics Center for Infectious Disease. The same protein purification protocol

404   was used for both *M. tuberculosis* and *M. smegmatis* CoaBC constructs.

405   *E. coli* BL21(DE3) containing pET28aSUMO-CoaBC was grown in 2XYT media at 37

406   ºC until an O.D.$_{600}$ = 0.6. IPTG was then added to a final concentration of 0.5 mM at the

407   temperature changed to 18 ºC for 18-20h. Cells were then harvested by centrifugation, re-

408   suspended in 50 mM TRIS pH 8.0, 250 mM NaCl, 20% (w/v) glycerol, 20 mM imidazole,

409   5 mM MgCl$_2$ with protease inhibitors tablets (Roche) and DNAseI (Sigma). Cells were

410   lysed with an Emulsiflex (Avestin) and the resultant cell lysate was centrifuged at 27000

411   x *g* for 30 min to remove cell debris. Recombinant CoaBCs were purified with a HiTrap

412   IMAC Sepharose FF column (GE-Healthcare), equilibrated with 50 mM TRIS pH 8.0,

413   250 mM NaCl, 20% glycerol (w/v) and 20mM Imidazole. Elution was performed in the

414   same buffer with 500mM Imidazole. Protein was dialysed in 25 mM TRIS pH 8 and 150

415   mM and SUMO tag was cleaved overnight at 4 ºC by adding Ulp1 Protease at 1:100

416   ratio. CoaBC was concentrated and loaded in a Superdex 200 column equilibrated with

417   25mM TRIS pH 8.0, 150 mM NaCl. Fraction purity was determined by SDS-page and

418   the purest fractions were pooled, concentrated to ~10 mg.ml^{-1} for *M. tuberculosis* CoaBC

419   and 30 mg.ml^{-1} for *M. smegmatis* CoaBC, flash frozen in liquid nitrogen and stored at -

420   80 ºC.

421   *E. coli* BL21(DE3) containing the *M. smegmatis* CoaB construct with a N-terminal non-

422   -cleavable 6xhis tag was grown and harvested as above and re-suspended in 20 mM

423   HEPES pH 7.0, 500 mM NaCl, 20 mM imidazole, 5 mM MgCl$_2$ with protease inhibitors

424   tablets (Roche) and DNAseI (Sigma). Cells were lysed with an Emulsiflex (Avestin) and

425   cell lysate was centrifuged at 27000 g for 30 mins to remove cell debris. Recombinat *M.*

426   *smegmatis* CoaB was purified with a HiTrap IMAC Sepharose FF column (GE-

427   Healthcare), equilibrated with 20 mM HEPES pH 7.0, 500 mM NaCl and 20 mM

428   imidazole. Elution was carried in the same buffer with 500 mM imidazole. Protein was

429   concentrated and loaded on a Superdex 200 column equilibrated with 20mM HEPES pH

430   7.0 and 500 mM NaCl. Fraction purity was assessed by SDS-page and the purest fractions

431   were pooled concentrated to 22 mg. ml^{-1}, flash frozen in liquid nitrogen and stored at -80

432   ºC.

433

434   **Native mass spectrometry**

435   Spectra were recorded on a Synapt HDMS mass spectrometer (Waters) modified for

436   studying high masses. MtCoaBC and MsCoaBC were exchanged into NH_4OAc (500 mM,

437   pH 7.0) solution using Micro Bio-Spin 6 chromatography columns (Bio-Rad). 2.5 μL of

438   sample solution was injected into a borosilicate emitter (Thermo Scientific) for sampling.

439   Instrument conditions were optimized to enhance ion desolvation while minimizing

440   dissociation of macromolecular complexes. Typical conditions were capillary voltage

441   1.8–2.0 kV, sample cone voltage 100 V, extractor cone voltage 1 V, trap collision voltage

442   60 V, transfer collision voltage 60 V, source temperature 20 °C, backing pressure 5 mbar,

443   trap pressure $3–4 \times 10^{-2}$ mbar, IMS (N_2) pressure $5–6 \times 10^{-1}$ mbar and TOF pressure 7–

444   8×10^{-7} mbar. Spectra were calibrated externally using cesium iodide. Data acquisition

445   and processing were performed using MassLynx 4.1 (Waters).

446

447   **Crystallization**

448   For both full length *M. smegmatis* CoaBC and CoaB alone, the crystallization screens and

449   optimization were performed at 18 ºC using the sitting-drop vapour diffusion method. For

450   CoaBC 300 nL of pure protein at 30 mg.ml^{-1}, pre-incubated with 3 mM CTP and 10 mM

451   MgCl$_2$, was mixed in 1:1 and 1:2 (protein to reservoir) ratio with well solution using a

452   mosquito robot (TTP labtech). Initial conditions were obtained in the Classics lite

453   crystallization screen (Qiagen), solution 1. Crystals obtained in this condition diffracted

454   poorly, therefore several rounds of optimization were performed. The final optimised

455   condition consisted of 0.1 M BisTris pH 6.5, 10 mM CoCl$_2$ 0.8 M 1,6-hexanediol.

456   Crystals appeared after three days in both conditions. A cryogenic solution was prepared

457   by adding ethylene glycol up to 30% (v/v) to mother liquor. Crystals were briefly

458   transferred to this solution, flash frozen in liquid nitrogen and stored for data collection.

459   For MsmCoaB 200 nL of pure protein at 22-24 mg.mL^{-1} with 10 mM CTP was mixed in

460   1:1 ratio with well solution using a Phoenix robot (Art Robbins). Crystals were obtained

461   in Wizards classics III&IV (Rigaku) solution G4 consisting of 20% (w/v) PEG 8000,

462   0.1M MES pH 6.0 and 0.2 M calcium acetate. Crystals appeared after 2 days.

463   To obtain ligand-bound structures, soaking was performed condition using the hanging-

464   drop vapour-diffusion method as follows: 2 uL of a solution containing 20% (w/v) PEG

465   8000, 0.1M MES pH 6.0, 0.2 M calcium acetate, 0.25 M NaCl 10% (v/v) DMSO and 1-

466   5 mM inhibitors was left to equilibrate against 500 µl of reservoir solution for 3 days.

467   Crystals were then transferred to the pre-equilibrated drops and incubated for 24h. A

468   cryogenic solution was prepared by adding 2-Methyl-2,4-pentanediol up to 25% (v/v) to

469   mother liquor. Crystals were briefly transferred to this solution, flash frozen in liquid

470   nitrogen and stored for data collection.

471   While we have also obtained *M. tuberculosis* CoaBC crystals they did not diffract and

472   our optimization efforts failed to improve them.

473

474   **Data collection and processing**

475 The data sets were collected at stations I02 and I03 at Diamond Light Source (Oxford,

476 UK). The diffraction images were processed with AutoPROC [37] using XDS [38] for

477 indexing and integration with AIMLESS [39] and TRUNCATE [40] from CCP4 Suite [41] for

478 data reduction, scaling and calculation of structure factor amplitudes and intensity

479 statistics.

480

481 **Structure solution and refinement**

482 MsmCoaB and MsmCoaBC structures were solved by molecular replacement using

483 PHASER [42] from the PHENIX software package [43]. For MsmCoaB the atomic

484 coordinates of MsmCoaB structure (PDB entry 4QJI) were used as a search model.

485 Ligand bound structures were solved using our highest resolution MsmCoaB apo form

486 structure (PDB entry 6TH2). For MsmCoaBC atomic coordinates of *Arabidopsis thaliana*

487 CoaC (PDB entry 1MVL) [19] and our highest resolution CoaB structure (PDB entry 6TH2)

488 were used as search models. Model building was done with Coot [44] and refinement was

489 performed in PHENIX [43]. Structure validation was performed using Coot and PHENIX

490 tools [43,44]. All figures were prepared using Pymol (The PyMOL Molecular Graphics

491 System, Version 2.0 Schrödinger, LLC.) and ligand interactions calculated with Arpeggio

492 [45].

493

494 **High-throughput screening**

495 Potential inhibitors of CoaBC were assessed at room temperature using a PHERAStar

496 microplate reader (BMG Labtech). Pyrophosphate produced by CoaB was converted to

497 two molecules of inorganic phosphate using a pyrophosphatase. Phosphate was then

498 detected using the BIOMOL® Green reagent (Enzo Life Sciences), which when bound

499 to phosphate absorbs light at 650 nm. An end-point assay was carried out in clear, flat-

500   bottom, polystyrene, 384-well plates (Greiner) in an 50 µl reaction volume containing

501   100 mM TRIS, pH 7.6; 1 mM $MgCl_2$; 1 mM TCEP; 0.03 U/mL pyrophosphatase; 2 µM

502   CTP; 40 µM L-cysteine; 30 µM *P*PA and 30 nM MtbCoaBC. Assays were performed by

503   adding 25 µL of a 2-times concentrated reaction mixture containing all components with

504   the exception of the enzymes to all wells, and the reactions started by adding 25 µL of a

505   2-times concentrated enzyme mixture. The reaction was carried out for 2 h at room

506   temperature, before 50 µL of BIOMOL® Green reagent was added and incubated for a

507   further 20 min prior to reading.

508

509   **Inorganic pyrophosphatase-purine nucleoside phosphorylase PNP-PPIase assay.**

510   The commercially available EnzCheck pyrophosphate assay kit (E-6645) (Life

511   Technologies) was used for this assay. The final reaction composition used was 0.03

512   U/mL inorganic pyrophosphatase, 1 U/mL purine nucleoside phosphorylase, 1 mM

513    $MgCl_2$, 200 µM MESG, 100 mM TRIS pH 7.5, 1 mM TCEP, 32 nM MtbCoaBC, 125

514   µM CTP, 125 µM *P*PA, 500 µM L-cysteine, and various concentrations of compounds

515   being tested for inhibition all prepared from DMSO stock solutions (compounds of series

516   one and two) or water (CoA and CoA thioesters). Assays were performed on either a

517   CLARIOStar or PHERAStar microplate reader (BMG Labtech) in 96-well plates

518   (Greiner). A substrate mixture containing the substrates and the inhibitor was pre-

519   incubated at 25 °C for 10 min. An enzyme solution was prepared and separately pre-

520   incubated at 25 °C for 10 min. The reaction was initiated by the addition of the substrates

521   to the solution containing the enzyme to a final volume of 75 µL. Enzymatic activity was

522   monitored by following the absorbance at 360 nm for 30 min (100 cycles/20 s each cycle).

523   Assays were performed in triplicates, including a negative control (lacking *P*PA) and a

524   positive control (lacking inhibitor).

525 Competition assays were performed using the same conditions but with variable substrate

526 concentrations (31.25 µM, 62.5 µM, 125 µM, 250 µM and 500 µM for CTP and *P*PA,

527 31.25 µM, 62.5 µM, 93.75 µM 125 µM and 250 µM for L-cysteine).

528

529 ***M. tuberculosis* strains and growth conditions**

530 MIC determination for *M. tuberculosis* H37RvMA was performed as previously

531 described [46] in the following media: 7H9/ADC/glycerol (4.7 g/L Difco Middlebrook 7H9

532 base, 100mL/L Middlebrook albumin (BSA)-dextrose-catalase (ADC) Difco

533 Middlebrook, 0.2% glycerol and 0.05% Tween-80) , 7H9/Cholesterol/Tyloxapol (4.7 g/L

534 7H9 base, 0.81 g/L NaCl, 24 mg/L cholesterol, 5 g/L BSA fraction V and 0.05%

535 Tyloxapol), GAST/Fe (0.3 g/L of Bacto Casitone (Difco), 4.0/L g of dibasic potassium

536 phosphate, 2.0 g/L of citric acid, 1.0 g/L of L-alanine, 1.2 g/L of magnesium chloride

537 hexahydrate, 0.6 g/L of potassium sulfate, 2.0 g/L of ammonium chloride, 1.80 ml/L of

538 10 N sodium hydroxide, 10.0 ml of glycerol 0.05% Tween 80 and 0.05 g of ferric

539 ammonium citrate adjusted to pH 6.6.

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

Figure Legends

Figure 1: X-ray crystal structure of FMN and CTP bound MsmCoaBC.

(A) Full aspect of the dodecameric CoaBC with CoaC represented in teal and CoaB in gold. (B) View of a CoaBC dimer with FMN and CTP shown. Each protomer is coloured differently. The CoaC active site flexible flap is highlighted in blue. (C) In the left panel, a CoaBC trimer is shown with the CoaC coloured in teal and CoaB in gold. On the right panel dimerization of two CoaBC trimers is shown with CoaC coloured in teal or grey for different trimers. Each CoaB forms a dimer with protomers from different trimers.

Figure 2: Detailed view of MsmCoaBC active sites and MsmCoaB dimerisation interface.

(A) View of CoaC active site with FMN bound. The active site sits between two protomers of one trimer (gold and red) and a third protomer from an adjacent trimer (green). Hydrogen bonds are depicted in yellow and π-interactions are in blue. (B)

575   Superposition of a CoaB crystal structure (PDB code: 6TH2) in green, with full length

576   CoaBC (teal) showing the active site flaps (brown) of the CoaB and CoaC enzymes. (C)

577   Detailed view of the CTP binding site. Cartoon and residues belonging to each protomer

578   are coloured differently. Hydrogen bonds and π-interactions are coloured as in B.

579   Important waters are represented as red spheres and calcium as a green sphere. Calcium

580   coordination is depicted in purple. (D) CoaB dimerization interface. Each protomer is

581   coloured as in C. (D)

582

583   **Figure 3: Regulation of MtbCoaBC by CoA and CoA thioesthers.**

584   (A) Inhibition of MtbCoaBC by CoA, acetyl-CoA, malonyl-CoA and succinyl-CoA. (B)

585   Lineweaver-Burk plots showing the effect of varying the concentration of each substrate

586   in the presence of different concentrations of acetyl-CoA.

587

588   **Figure 4: Inhibition of MtbCoaBC by compounds 1b and 2b.**

589   (A) Dose response profiles and chemical structure of compounds **1b** and **2b** is shown. (B)

590   Lineweaver-burke plots showing the effect of varying concentrations of compound **1b**

591   (left) and **2b** (right) in the presence of varying concentrations of CTP, *P*PA and L-

592   cysteine.

593

594   **Figure 5: MsmCoaB X-ray structure showing the cryptic allosteric site.**

595   CoaB with the cryptic allosteric site closed (A) and opened conformation (B) with

596   compound **1b** (pink) bound. (C) Detailed view of the allosteric site with compound **1b**

597   (yellow) bound. The individual protomers of the CoaB dimer are coloured in green or

598   pink.. Hydrogen bonds are depicted in red, π-interactions are in grey, and hydrophobic

599   interaction in green. Important waters are represented as red spheres and calcium as a

600   green sphere. Calcium coordination is depicted in purple. (D) Gating mechanism of the

601   cryptic allosteric site showing the movement of R207 with the closed conformation in

602   yellow and the open conformation in pink. An *E. coli* structure (PDB code: 1U7Z) with

603   the 4'-phosphopantothenoyl-CMP (purple) intermediate bound is superimposed.

604

605   **Figure 6: Docking of compound 2b into MsmCoaB showing the highest scoring pose.**

606   Hydrogen bonds are shown in red. The individual protomers of the CoaB dimer are either

607   coloured in green or pink.

608

609

610

611

612   **Tables**

613

614   **Table 1:** Inhibition of CoaB domain by CoA, CoA thioesters and the most potent

615   inhibitors from of series one and two. IC_{50} values determined using the EnzChek

616   pyrophosphate assay are shown.

Compound	IC_{50} EnzCheck (μM)
CoA	148 ± 11
AcCoA	121 ± 9
MlCoA	49 ± 3
SucCoA	38 ± 2
1b	0.28 ± 0.05
1c	4.6 ± 0.4
2b	0.08 ± 0.01
2c	0.41 ± 0.03
2d	0.54 ± 0.06
2e	3.0 ± 0.2

617

618 **Table 2:** Minimum inhibitory concentration (MIC) values of CoaB inhibitors against *M.*

619 *tuberculosis* H37Rv cultured in different media (μM).

Compound	7H9/ADC/ Glycerol	7H9/ Cholesterol/ Tyloxapol	GAST/ Fe
1b	>250	>250	125
1c	>250	ND	ND
2b	>250	>250	50
2c	>250	>250	125
2d	>250	>250	ND
2e	>250	ND	ND

620 *ND – Not determined.

621

622

623

624

625

626 **Acknowledgements**

638

References

1. World Health, O. *Global tuberculosis report 2019*, (World Health Organization, Geneva, 2019).

2. Strauss, E. Coenzyme A Biosynthesis and Enzymology. *Comprehensive Natural Products Ii: Chemistry and Biology, Vol 7: Cofactors*, 351-410 (2010).

3. Leonardi, R. & Jackowski, S. Biosynthesis of Pantothenic Acid and Coenzyme A. *EcoSal Plus* **2**(2007).

4. Tsuchiya, Y. et al. Protein CoAlation and antioxidant function of coenzyme A in prokaryotic cells. *Biochem J* **475**, 1909-1937 (2018).

5. Choudhary, C. et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-40 (2009).

6. Beld, J., Sonnenschein, E.C., Vickery, C.R., Noel, J.P. & Burkart, M.D. The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Nat Prod Rep* **31**, 61-108 (2014).

7. Wang, M. & Casey, P.J. Protein prenylation: unique fats make their mark on biology. *Nat Rev Mol Cell Biol* **17**, 110-22 (2016).

8. Bird, J.G. et al. The mechanism of RNA 5' capping with NAD+, NADH and desphospho-CoA. *Nature* **535**, 444-7 (2016).

9. Marrakchi, H., Laneelle, M.A. & Daffe, M. Mycolic acids: structures, biosynthesis, and beyond. *Chem Biol* **21**, 67-85 (2014).

10. Guerrini, V. et al. Storage lipid studies in tuberculosis reveal that foam cell biogenesis is disease-specific. *PLoS Pathog* **14**, e1007223 (2018).

11. Peyron, P. et al. Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence. *PLoS Pathog* **4**, e1000204 (2008).

12. Evans, J.C. et al. Validation of CoaBC as a Bactericidal Target in the Coenzyme A Pathway of Mycobacterium tuberculosis. *ACS Infect Dis* **2**, 958-968 (2016).

13. Kupke, T. et al. Molecular characterization of lantibiotic-synthesizing enzyme EpiD reveals a function for bacterial Dfp proteins in coenzyme A biosynthesis. *J Biol Chem* **275**, 31838-46 (2000).

14. White, M.D. et al. UbiX is a flavin prenyltransferase required for bacterial ubiquinone biosynthesis. *Nature* **522**, 502-6 (2015).

15. Blaesse, M., Kupke, T., Huber, R. & Steinbacher, S. Crystal structure of the peptidyl-cysteine decarboxylase EpiD complexed with a pentapeptide substrate. *EMBO J* **19**, 6299-310 (2000).

16. Blaesse, M., Kupke, T., Huber, R. & Steinbacher, S. Structure of MrsD, an FAD-binding protein of the HFCD family. *Acta Crystallogr D Biol Crystallogr* **59**, 1414-21 (2003).

17. Albert, A. et al. The X-ray structure of the FMN-binding protein AtHal3 provides the structural basis for the activity of a regulatory subunit involved in signal transduction. *Structure* **8**, 961-9 (2000).

18. Manoj, N. & Ealick, S.E. Unusual space-group pseudosymmetry in crystals of human phosphopantothenoylcysteine decarboxylase. *Acta Crystallogr D Biol Crystallogr* **59**, 1762-6 (2003).

19. Steinbacher, S. et al. Crystal structure of the plant PPC decarboxylase AtHAL3a complexed with an ene-thiol reaction intermediate. *J Mol Biol* **327**, 193-202 (2003).

20. Stanitzek, S., Augustin, M.A., Huber, R., Kupke, T. & Steinbacher, S. Structural basis of CTP-dependent peptide bond formation in coenzyme A biosynthesis catalyzed by Escherichia coli PPC synthetase. *Structure* **12**, 1977-88 (2004).

687 21. Manoj, N., Strauss, E., Begley, T.P. & Ealick, S.E. Structure of human
688   phosphopantothenoylcysteine synthetase at 2.3 A resolution. *Structure* **11**, 927-36
689   (2003).
690 22. Zheng, P. et al. Crystallographic Analysis of the Catalytic Mechanism of
691   Phosphopantothenoylcysteine Synthetase from Saccharomyces cerevisiae. *J Mol Biol*
692   (2019).
693 23. Kupke, T. Molecular characterization of the 4'-phosphopantothenoylcysteine
694   synthetase domain of bacterial dfp flavoproteins. *J Biol Chem* **277**, 36137-45 (2002).
695 24. Chan, D.S.-H. et al. Structural insights into Escherichia coli
696   phosphopantothenoylcysteine synthetase by native ion mobility–mass spectrometry.
697   *Biochemical Journal* **476**, 3125-3139 (2019).
698 25. Strauss, E., Kinsland, C., Ge, Y., McLafferty, F.W. & Begley, T.P.
699   Phosphopantothenoylcysteine synthetase from Escherichia coli. Identification and
700   characterization of the last unidentified coenzyme A biosynthetic enzyme in bacteria. *J*
701   *Biol Chem* **276**, 13513-6 (2001).
702 26. Awasthy, D. et al. Essentiality and functional analysis of type I and type III
703   pantothenate kinases of Mycobacterium tuberculosis. *Microbiology* **156**, 2691-701
704   (2010).
705 27. Song, W.J. & Jackowski, S. Kinetics and regulation of pantothenate kinase from
706   Escherichia coli. *J Biol Chem* **269**, 27051-8 (1994).
707 28. Vallari, D.S., Jackowski, S. & Rock, C.O. Regulation of pantothenate kinase by coenzyme
708   A and its thioesters. *J Biol Chem* **262**, 2468-71 (1987).
709 29. Wubben, T.J. & Mesecar, A.D. Kinetic, thermodynamic, and structural insight into the
710   mechanism of phosphopantetheine adenylyltransferase from Mycobacterium
711   tuberculosis. *J Mol Biol* **404**, 202-19 (2010).
712 30. Miller, J.R. et al. Phosphopantetheine adenylyltransferase from Escherichia coli:
713   investigation of the kinetic mechanism and role in regulation of coenzyme A
714   biosynthesis. *J Bacteriol* **189**, 8196-205 (2007).
715 31. Bennett, B.D. et al. Absolute metabolite concentrations and implied enzyme active site
716   occupancy in Escherichia coli. *Nat Chem Biol* **5**, 593-9 (2009).
717 32. Moolman, W.J., de Villiers, M. & Strauss, E. Recent advances in targeting coenzyme A
718   biosynthesis and utilization for antimicrobial drug development. *Biochem Soc Trans* **42**,
719   1080-6 (2014).
720 33. Kupke, T. Active-site residues and amino acid specificity of the bacterial 4'-
721   phosphopantothenoylcysteine synthetase CoaB. *Eur J Biochem* **271**, 163-72 (2004).
722 34. Yun, M. et al. Structural basis for the feedback regulation of Escherichia coli
723   pantothenate kinase by coenzyme A. *J Biol Chem* **275**, 28093-9 (2000).
724 35. Hartl, J., Kiefer, P., Meyer, F. & Vorholt, J.A. Longevity of major coenzymes allows
725   minimal de novo synthesis in microorganisms. *Nat Microbiol* **2**, 17073 (2017).
726 36. Nguyen, L. & Pieters, J. Mycobacterial subversion of chemotherapeutic reagents and
727   host defense tactics: challenges in tuberculosis drug development. *Annu Rev*
728   *Pharmacol Toxicol* **49**, 427-53 (2009).
729 37. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta*
730   *Crystallogr D Biol Crystallogr* **67**, 293-302 (2011).
731 38. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).
732 39. Evans, P.R. & Murshudov, G.N. How good are my data and what is the resolution? *Acta*
733   *Crystallographica Section D-Biological Crystallography* **69**, 1204-1214 (2013).
734 40. French, S. & Wilson, K. Treatment of Negative Intensity Observations. *Acta*
735   *Crystallographica Section A* **34**, 517-525 (1978).
736 41. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta*
737   *Crystallogr D Biol Crystallogr* **67**, 235-42 (2011).

738 42. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674
739   (2007).
740 43. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular
741   structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
742 44. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot.
743   *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
744 45. Jubb, H.C. et al. Arpeggio: A Web Server for Calculating and Visualising Interatomic
745   Interactions in Protein Structures. *J Mol Biol* **429**, 365-371 (2017).
746 46. Singh, V. et al. The complex mechanism of antimycobacterial action of 5-fluorouracil.
747   *Chem Biol* **22**, 63-75 (2015).

748

749

750 **Author contributions**

751 VMendes wrote the manuscript. VM designed and performed all the crystallographic

752 experiments with the help of MB, OB and JCW. VMendes and JH designed and

753 performed the kinetic experiments. JH synthesised 4'-phosphopantothenate. PHMT

754 performed docking experiments. DSC performed the native mass spectrometry

755 experiments. SG, TB, SON, SD, JP and CS developed and performed the high-throughput

756 screening. JCE, SLL and HIMB performed the microbiology experiments on *M.*

757 *tuberculosis* H37Rv. VMendes, JCE, SG, AGC, PCR, KR, CA, HIMB, CEB, VMizrahi,

758 PGW and TLB managed the project. All authors approved the manuscript.

759

760 **Accession numbers**

761 Coordinates and structure factors related to this work have been deposited in the PDB

762 with accession numbers: **6TGV**, **6TH2** and **6THC**.

Figure 2
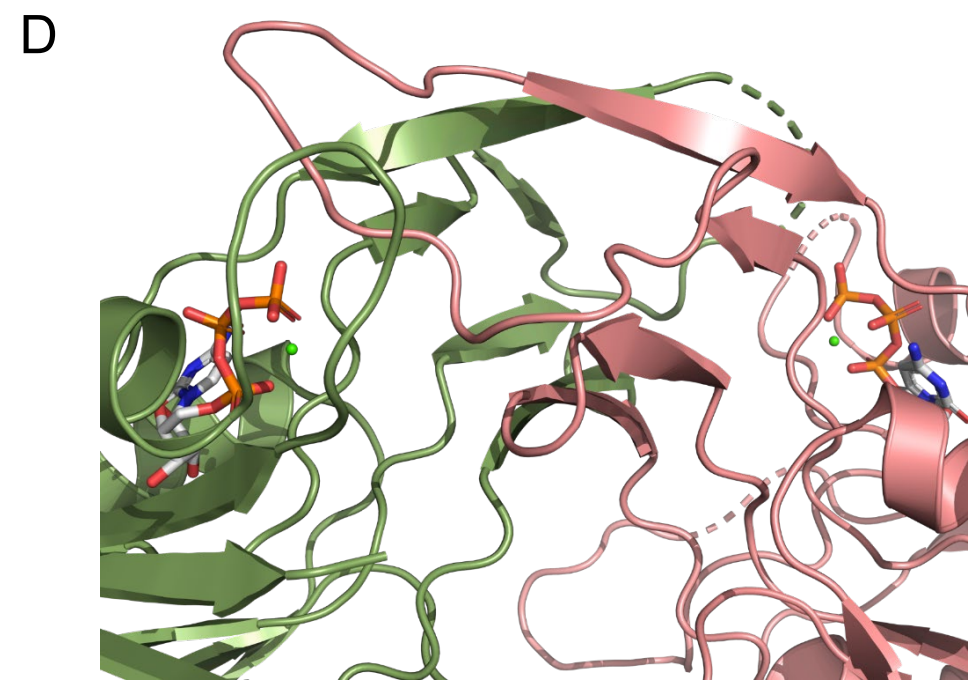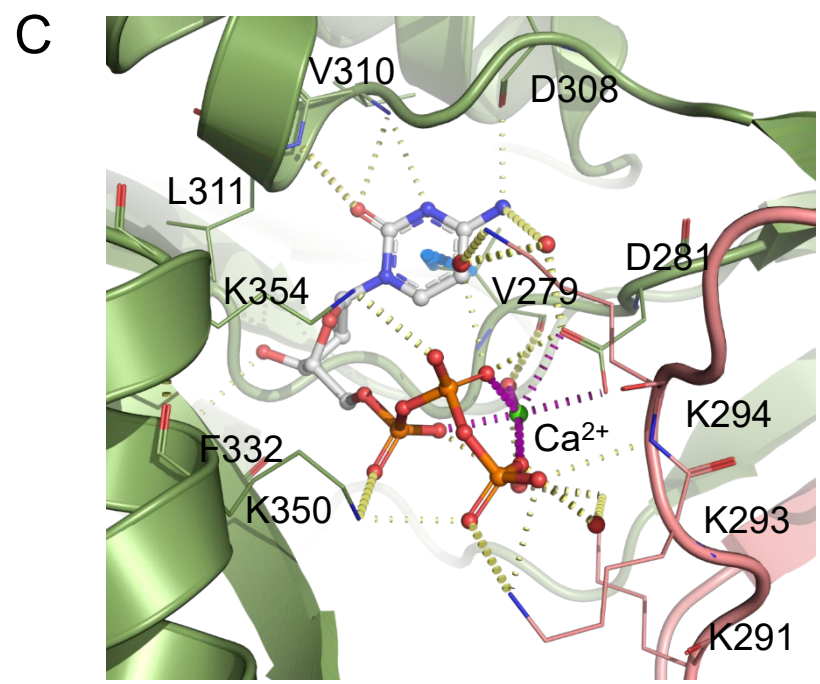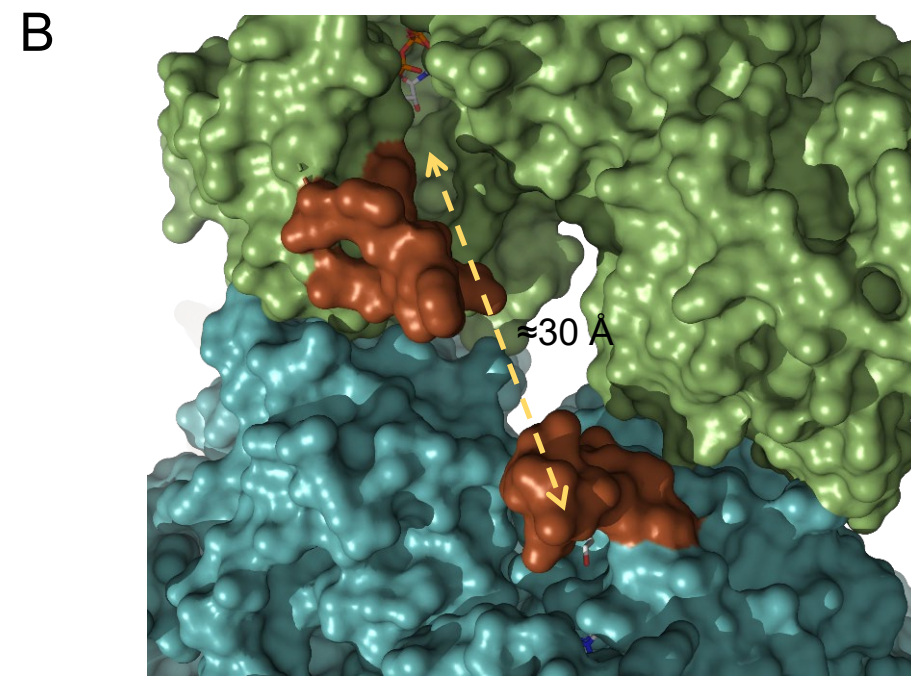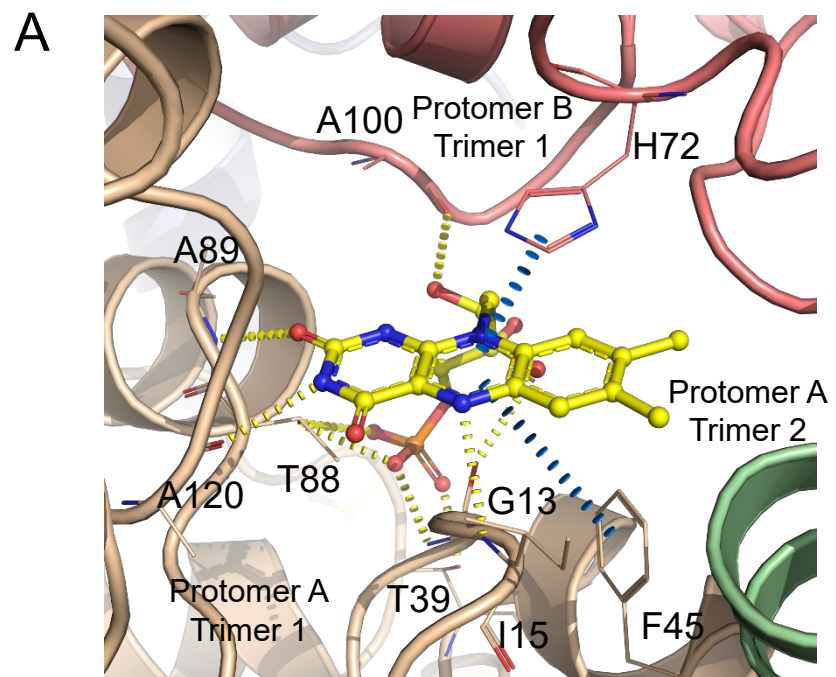
Figure 3
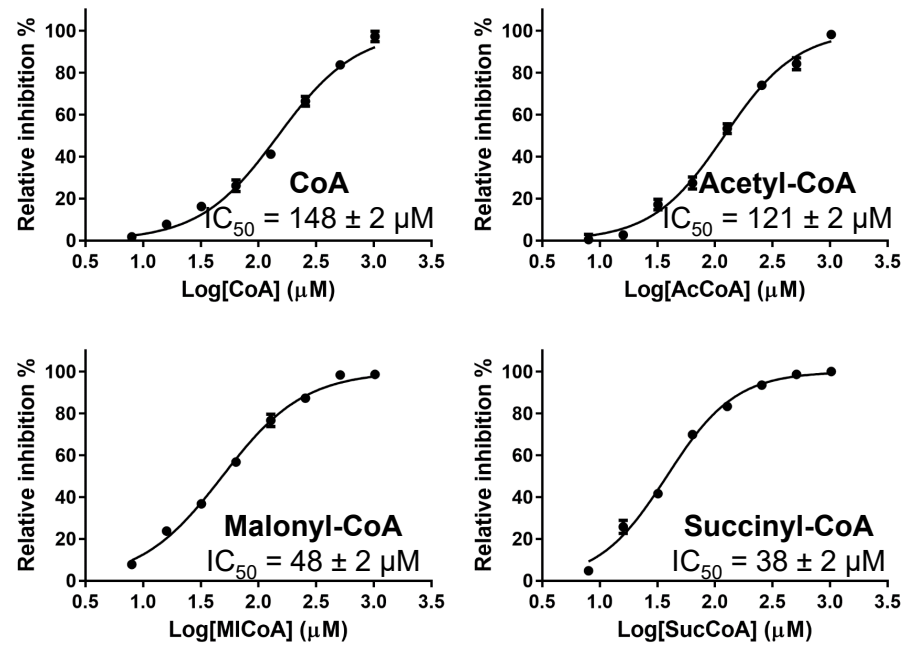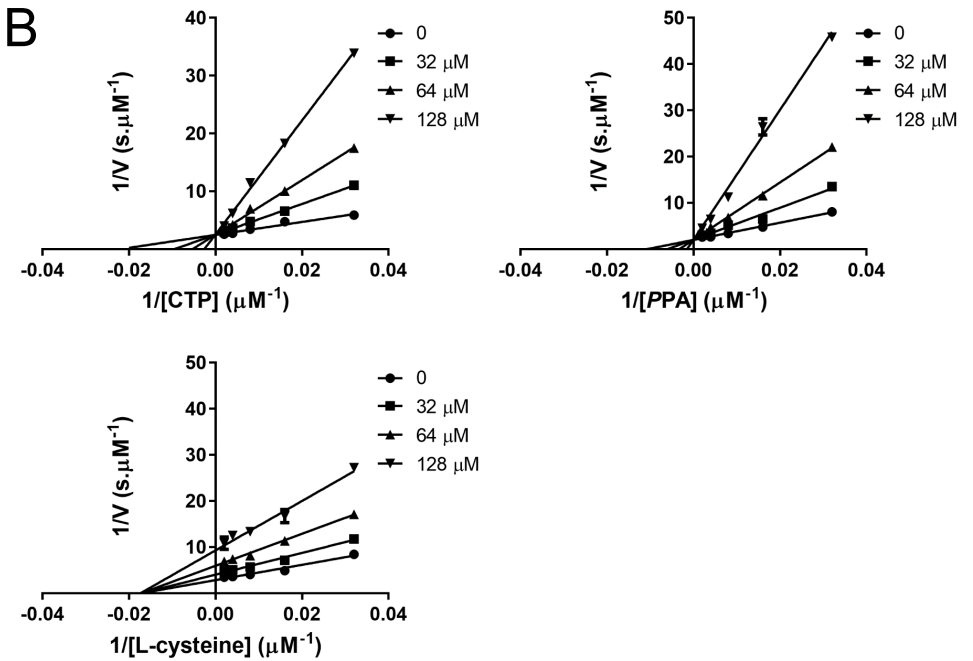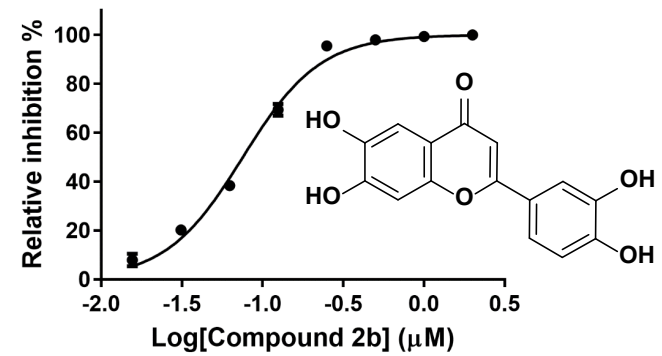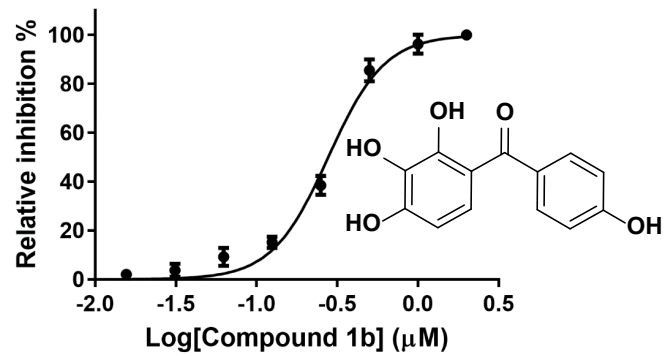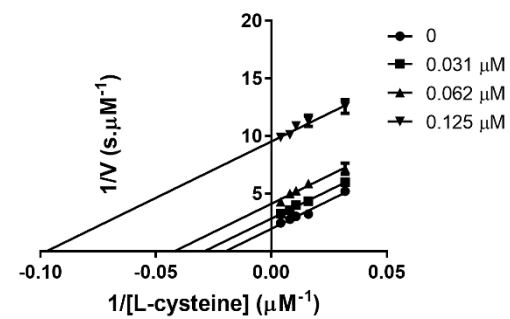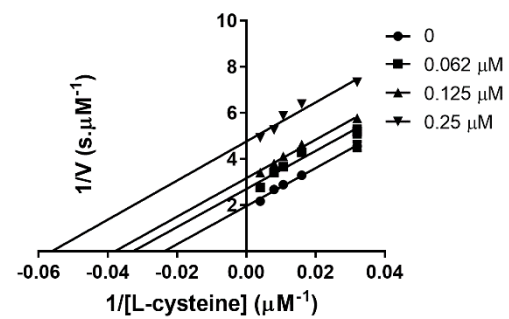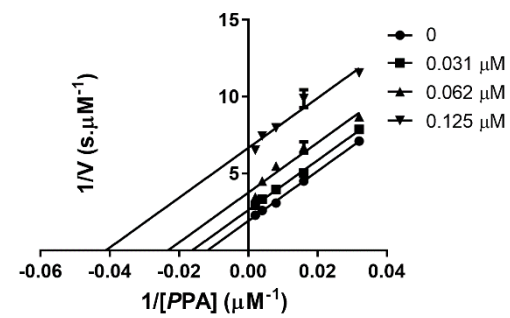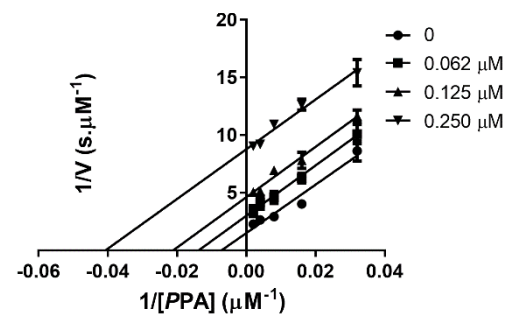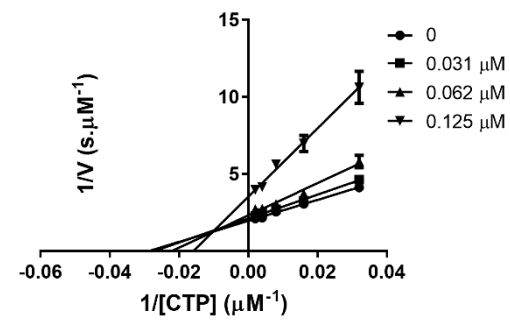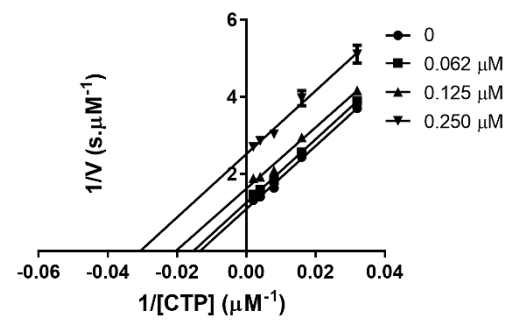
Figure 4

Figure 5

Figure 6

Figure 1