research papers

CrossMark

A covalent adduct of MbtN, an acyl-ACP dehydrogenase from *Mycobacterium tuberculosis*, reveals an unusual acyl-binding pocket

Ai-Fen Chai, Esther M. M. Bulloch, Genevieve L. Evans, J. Shaun Lott, Edward N. Baker* and Jodie M. Johnston*

Laboratory of Structural Biology, School of Biological Sciences and Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. *Correspondence e-mail: en.baker@auckland.ac.nz, jm.johnston@auckland.ac.nz

Mycobacterium tuberculosis (*Mtb*) is the causative agent of tuberculosis. Access to iron in host macrophages depends on iron-chelating siderophores called mycobactins and is strongly correlated with *Mtb* virulence. Here, the crystal structure of an *Mtb* enzyme involved in mycobactin biosynthesis, MbtN, in complex with its FAD cofactor is presented at 2.30 Å resolution. The polypeptide fold of MbtN conforms to that of the acyl-CoA dehydrogenase (ACAD) family, consistent with its predicted role of introducing a double bond into the acyl chain of mycobactin. Structural comparisons and the presence of an acyl carrier protein, MbtL, in the same gene locus suggest that MbtN acts on an acyl-(acyl carrier protein) rather than an acyl-CoA. A notable feature of the crystal structure is the tubular density projecting from N(5) of FAD. This was interpreted as a covalently bound polyethylene glycol (PEG) fragment and resides in a hydrophobic pocket where the substrate acyl group is likely to bind. The pocket could accommodate an acyl chain of 14–21 C atoms, consistent with the expected length of the mycobactin acyl chain. Supporting this, steady-state kinetics show that MbtN has ACAD activity, preferring acyl chains of at least 16 C atoms. The acyl-binding pocket adopts a different orientation (relative to the FAD) to other structurally characterized ACADs. This difference may be correlated with the apparent ability of MbtN to catalyse the formation of an unusual *cis* double bond in the mycobactin acyl chain.

1. Introduction

Mycobacterium tuberculosis (*Mtb*), the causative agent of tuberculosis (TB), is the world's most devastating pathogen. According to the World Health Organization (WHO), there were 8.6 million new TB cases (13% of which were co-infections with HIV) and 1.3 million deaths from TB in 2012 (Zumla *et al.*, 2013).

Although effective treatment exists, the emergence of *Mtb* strains that are resistant to all of the front-line drugs has prompted a search for new drug targets.

Iron is an indispensable nutrient for all organisms. In most habitats, however, iron is not freely available. *Mtb* resides primarily in the phagocytic vacuoles of the macrophages, where the maximum concentration of free iron(III) is only 1–10 ng ml^{-1} (Miethke & Marahiel, 2007). *Mtb* counters this problem by producing a specific chelator (*i.e.* siderophore) called mycobactin T (De Voss *et al.*, 2000). The genes encoding the enzymes involved in mycobactin biosynthesis are found in two gene clusters: *mbt* and *mbt-2* (Krithika *et al.*, 2006; Cole & Barrell, 1998). The *mbt* cluster contains ten genes designated *mbtA–J*, which encode enzymes involved in the biosynthesis and assembly of the mycobactin core (Cole *et al.*, 1998). The
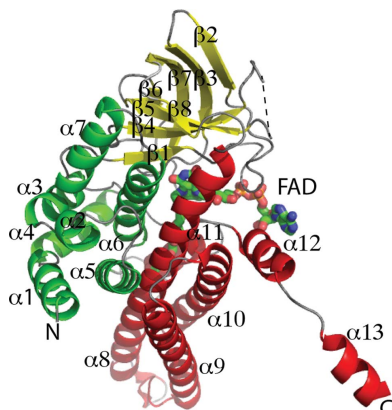
mbt-2 cluster contains enzymes that synthesize and add the acyl substituents to the mycobactin core and also genes that encode the iron-regulated transporters IrtA and IrtB. These gene clusters are essential for *Mtb* growth both *in vitro* (Sassetti & Rubin, 2003) and in macrophages (De Voss *et al.*, 2000), indicating that mycobactin-biosynthetic enzymes are promising targets for the development of new anti-TB drugs.

Most studies of mycobactin T biosynthesis have focused on the assembly of the core, with little being known about the addition of the acyl moiety to the core. There is structural variation in the acyl moiety of mycobactin T (Fig. 1*a*), but the main features are that the acyl moiety is a long unsaturated fatty-acid chain (C_{14} to C_{21}) with an unusual *cis* double bond at the α–β position adjacent to its connection to an *N*-hydroxy-lysine side chain of the mycobactin core (Snow, 1965, 1970; Kastrinsky *et al.*, 2010).

MbtN, encoded by Rv1346, is a putative acyl-ACP dehydrogenase (Cole & Barrell, 1998) predicted to catalyze the dehydrogenation of the acyl chain of mycobactin (Fig. 1*b*), and shares 23–28% sequence identity with previously characterized acyl-CoA dehydrogenases (ACADs). The ACAD family includes short-chain (SCADs), medium-chain (MCADs), long-chain (LCADs) and very long chain (VLCADs) enzymes, isovaleryl-CoA dehydrogenase (IVD) and isobutyryl-CoA dehydrogenase (iBD). Tertiary structures representing each subfamily have been reported and found to share a similar fold and a common substrate-binding mode (Tiffany *et al.*, 1997; Battaile *et al.*, 2002; Djordjevic *et al.*, 1994; Kim *et al.*, 1993; Lee *et al.*, 1996; McAndrew *et al.*, 2008; Satoh *et al.*, 2003). Although they vary in substrate specificity, these enzymes share a similar catalytic mechanism, catalyzing the α,β-dehydrogenation of fatty-acid acyl-CoA conjugates (Thorpe & Kim, 1995).

MbtN differs from previously characterized ACADs in several ways. Firstly, based on evidence detailed in this paper, the *in vivo* substrate for MbtN is likely to be an acyl-(acyl carrier protein) (acyl-ACP) rather than an acyl-CoA (Fig. 1*c*). Secondly, mycobactins commonly contain a *cis* double bond between the α and β C atoms of the acyl chain (Snow, 1965, 1970; Kastrinsky *et al.*, 2010), whereas all ACADs characterized to date catalyze the formation of products with *trans* double bonds. Here, we report the structure of MbtN in complex with its FAD cofactor as determined by X-ray crystallography and provide kinetic evidence for its preference for longer acyl chains. A polyethylene glycol (PEG) molecule is covalently bound to N(5) of FAD in the crystal structure, defining a long pocket that is predicted to accommodate the substrate acyl group. Importantly, the acyl-binding pocket has a different orientation from those of other ACADs, which may introduce steric constraints that are correlated with the formation of a *cis* double bond.
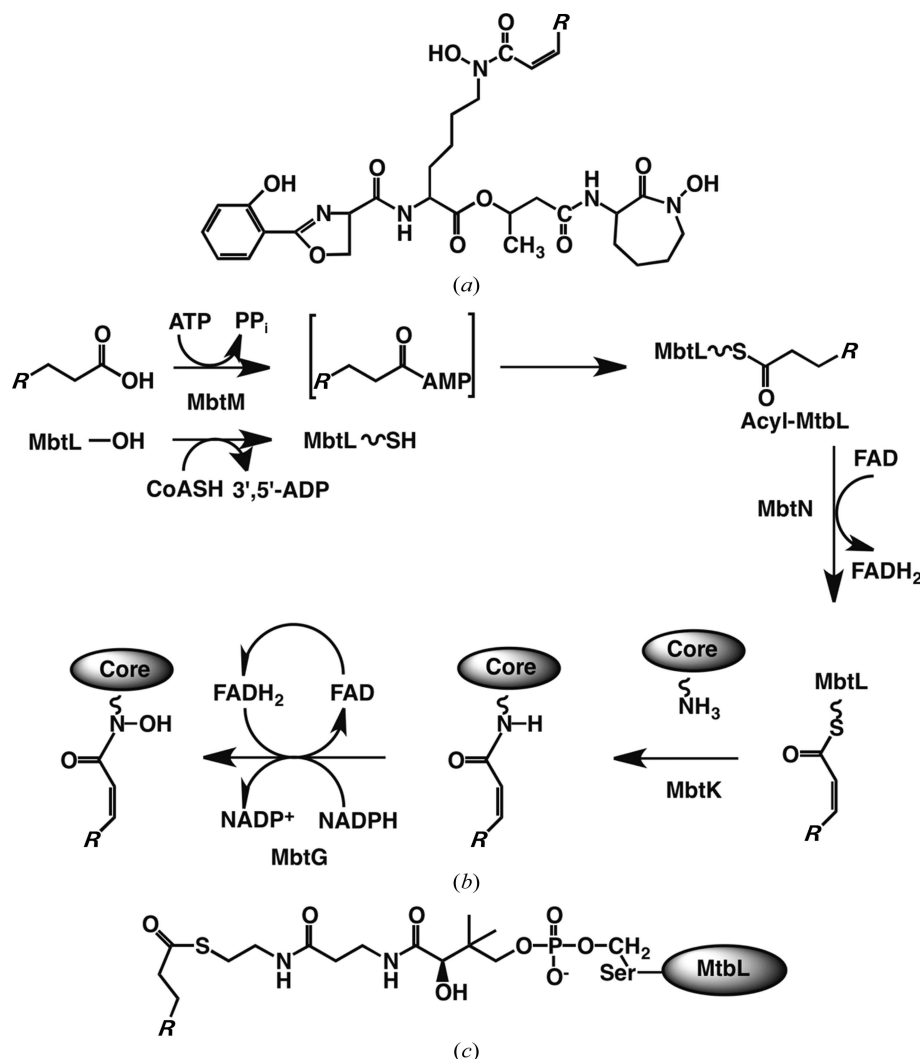

Figure 1
(*a*) Chemical structure of mycobactin T, the siderophore produced by *Mtb*. The length of the acyl chain in mycobactins is reported to vary from 14 to 21 C atoms (Snow, 1965, 1970). (*b*) The proposed pathway for mycobactin fatty-acid chain biosynthesis. An oval indicates the mycobactin core structure, while *R* indicates an acyl chain of various C-atom lengths. (*c*) Chemical structure of the phosphopantetheine linker between the acyl moiety (*R*) and the acyl carrier protein MbtL, the proposed substrate for MbtN.

2. Experimental procedures

2.1. Materials

Unless otherwise stated, all chemicals were obtained from Sigma–Aldrich, Scharlau or Pure Science. Protein concentrations were determined in triplicate using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific). The substrates used for enzyme assays,

myristoyl-CoA and palmitoyl-CoA, were obtained from Sigma–Aldrich.

2.2. Protein expression, purification and crystallization

The open reading frame encoding MbtN (Rv1346, FadE14) from *Mtb* strain H37Rv was cloned into the pProExHtb vector and transformed into *Escherichia coli* BL21(DE3) cells for heterologous expression as a His$_6$-tagged protein as described previously (Chai *et al.*, 2013). The protein was purified by immobilized metal-affinity chromatography, followed by the removal of the His$_6$ tag using *Tobacco etch virus* protease and size-exclusion chromatography (Chai *et al.*, 2013). The final purified protein comprised 389 amino-acid residues, representing residues 2–386 of the mature protein, together with a four-residue N-terminal extension (GAMG) that remained after tag removal. Crystals were grown in hanging drops by mixing equal volumes of protein solution (10 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 2 m*M* β-mercaptoethanol) and reservoir solution [23%(*w/v*) polyethylene glycol 3350, 0.2 *M* diammonium hydrogen citrate pH 5.5] at 291 K.

2.3. Data collection and structure solution

X-ray diffraction data to a resolution of 2.30 Å were collected at the Australian Synchrotron (beamline MX2, ADSC Quantum 315R CCD detector) at 110 K and processed as described previously (Chai *et al.*, 2013). Full details of the data quality are given in Chai *et al.* (2013). The *BALBES* server (Long *et al.*, 2008) identified *Thermus thermophilus* acyl-CoA dehydrogenase (PDB entry 1ws9; RIKEN Structural Genomics/Proteomics Initiative, unpublished work; 24% sequence identity to MbtN) as the best search template. The coordinates of the 1ws9 monomer, edited to conform with the MbtN sequence, were used as the model in the molecular-replacement program *Phaser_MR* (McCoy *et al.*, 2007). Two monomers were found per asymmetric unit, giving a Matthews coefficient (Matthews, 1968) V_M of 2.89 Å^3 Da^{-1} (57% solvent content). The structure was refined using *REFMAC*5 v.5.5.0109 (Murshudov *et al.*, 2011) and *BUSTER* v.2.8.0 (Bricogne *et al.*, 2011) alternating with manual rebuilding of the molecular structure using *Coot* v.0.6.2 (Emsley *et al.*, 2010). The two modified FAD molecules were generated using the *PRODRG* server (Schüttelkopf & van Aalten, 2004). Dihedral angle restraints were used to keep the pyrimidine and dimethylbenzene rings of the isoalloxazine moiety planar, while the central pyrazine ring was not restrained. This allowed the isoalloxazine to adopt a bent conformation along the N(5)—N(10) axis. The final R_{work} and R_{free} values were 0.174 and 0.210, respectively. The final model quality was assessed with *MolProbity* (Chen *et al.*, 2010). The refinement statistics are summarized in Table 1. Images of the structures were generated using *PyMOL* (DeLano, 2002).

2.4. Analysis of flavin planarity in published structures

The coordinates for crystal structures containing bent isoalloxazine moieties were obtained from the Protein Data Bank (http://www.rcsb.org). A total of 13 structures were

Table 1
Model-refinement statistics.

Values in parentheses are for the highest resolution shell.

PDB entry	4m6z
Resolution range (Å)	46.00–2.30 (2.42–2.30)
No. of reflections (working/test)	39900/2118 (2902/175)
$R_{\text{work}}/R_{\text{free}}$	0.174/0.210 (0.205/0.232)
No. of atoms (non-H)	
Protein (two monomers)	5603
Solvent	202
Ligand (two molecules)	349
Root-mean-square deviations from ideality	
Bonds (Å)	0.010
Angles (°)	1.04
Average *B* factors (Å^2)	
Monomer *A*	47.9
Monomer *B*	38.3
Ligand in monomer *A*	51.8
Ligand in monomer *B*	44.5
Solvent	42.8
Ramachandran plot	
Residues in most favoured regions (%)	97.15

retrieved, including structures that contained reduced flavins or flavin N(5) adducts. The dihedral angles between the dimethylbenzene and pyrimidine rings were measured using *PyMOL* (Table 2).

2.5. Steady-state kinetics

Assays for dehydrogenase activity were performed using phenazine methosulfate (PMS) as the intermediate electron carrier and dichloroindolphenol (DCPIP) as the terminal electron acceptor. All assays were measured on a Cary 4000 UV–Vis spectrophotometer using quartz cuvettes thermally equilibrated at 25°C and the reaction was initiated by the addition of enzyme. The following components were included in the assay: buffer (50 m*M* potassium phosphate, 0.3 m*M* EDTA, pH 7.6), PMS (1.4 m*M*), DCPIP (36 µ*M*), enzyme (0.5 µ*M*) and either myristoyl-CoA (20–140 µ*M*) or palmitoyl-CoA (5–60 µ*M*). Steady-state kinetic parameters for MbtN were determined from the rates of DCPIP reduction at 600 nm (ε = 21 000 *M*$^{-1}$ cm^{-1}). The units of enzyme activity were expressed in nmol min^{-1} per milligram of protein. In all experiments, enzymatic rates were corrected for the rate of uncatalyzed acyl-CoA hydrolysis recorded prior to the addition of enzyme. The initial rates for all substrate concentrations were determined in triplicate.

The kinetic constants k_{cat} and K_{m} were calculated by fitting the initial velocity data to the following equation using *GraphPad Prism* (v.5.03 for Windows, San Diego, California, USA; http://www.graphpad.com),

$$v = \frac{k_{\text{cat}}[\text{S}]}{K_{\text{m}} + [\text{S}]}, \qquad (1)$$

where k_{cat} is the turnover number, [S] is the concentration of the varied substrate and K_{m} is the Michaelis–Menten constant for the substrate.

3. Results

3.1. Molecular structure

The final model contained 744 of the 778 residues in the asymmetric unit, together with 203 water molecules. Residues 5–170 and 177–381 of MbtN are modelled in monomer *A*, whereas residues 5–172 and 177–381 are modelled in monomer *B*. The main-chain torsion angles conform well to standard values, with 718 residues falling in the most favoured regions of the Ramachandran plot and none in disallowed regions. The model has a *MolProbity* score of 1.43 and a clashscore of 1.85.

Table 2
Isoalloxazine bending angles of MbtN and flavoprotein crystal structures in the PDB.

The dihedral angles between the dimethylbenzene and pyrimidine rings are shown. The angles were measured using *PyMOL* (DeLano, 2002).

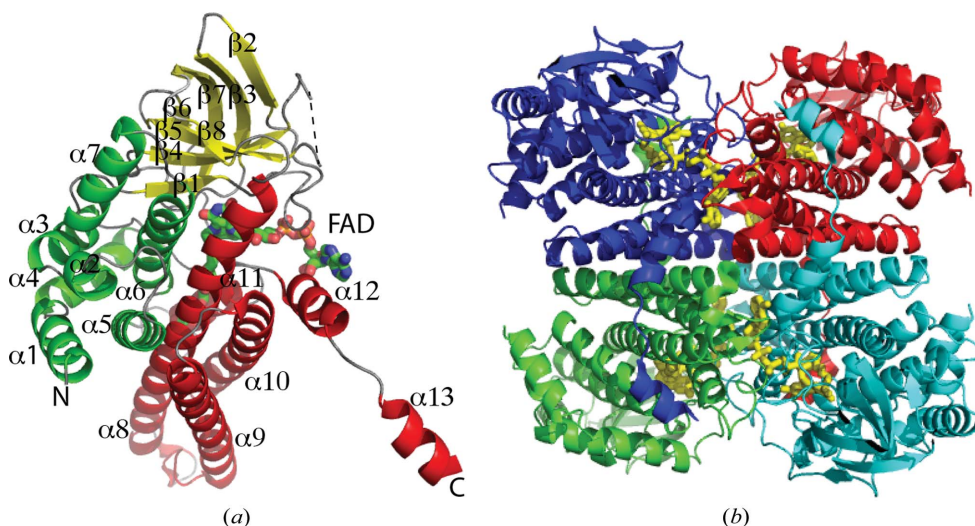| Protein	PDB code	C9—N10—N5—C4 (°)	Distortion (°)	Reference
Reduced FAD structures				
Acyl-CoA dehydrogenase	3r7k	157.7	22.3	Seattle Structural Genomics Center for Infectious Disease (unpublished work)
UDP galactopyranose	4gde	171.2	8.8	Dhatwalia *et al.* (2012)
Oxidoreductase	3rha	153.3	26.7	New York Structural Genomics Research Consortium (unpublished work)
Proline dehydrogenase	4h6r	160.4	19.6	Luo *et al.* (2012)
Choline oxidase	3ljp	182.7	7.3	Finnegan *et al.* (2010)
Glucose dehydrogenase	3tsj	155.1	24.9	D. Zafred, A. Nandy & W. Keller (unpublished work)
FAD N(5)-adduct structures				
D-Arginine dehydrogenase	3sm8	149.5	30.5	Fu *et al.* (2011)
Pyranose 2-oxidase	3lsm	155.3	24.7	Tan *et al.* (2010)
Sarcosine oxidase	2a89	162.7	17.3	Chen *et al.* (2005)
Adenosine-5-phosphosulfate reductase	2fjd	155.9	24.1	Schiffer *et al.* (2006)
Histone demethylase	2xah	152.9	27.1	Binda *et al.* (2010)
D-Amino-acid oxidase	1dao	151.9	28.1	Todone *et al.* (1997)
Nitroalkane oxidase	2c0u	161.2	18.8	Nagpal *et al.* (2006)
MbtN				
Modified FAD, monomer *A*		140.7	39.3	
Modified FAD, monomer *B*		148.6	31.4	

Figure 2
(*a*) The MbtN monomer. The N-terminal, middle and C-terminal domains are shown in green, yellow and red, respectively. The FAD adduct is shown as green sticks. (*b*) Tetrameric MbtN. Different monomers are shown in different colours and the FAD adduct is shown as yellow sticks.

MbtN exists as a tetramer in the crystal structure (Fig. 2*b*) and in solution based on size-exclusion chromatography multi-angle light scattering (SEC-MALS; data not shown). The crystallographic asymmetric unit contains a dimer, with the tetramer being formed by the application of twofold crystallographic symmetry. *PISA* (Krissinel & Henrick, 2007) analysis indicates that an average of 11% of the solvent-accessible surface per monomer is buried at the dimer inter-face and 12.8% at the major interface of the tetramer. The two monomers in the asymmetric unit are essentially the same, and each has a noncovalently bound FAD molecule defining the active site. The two monomers can be superimposed with a root-mean-square deviation (r.m.s.d.) of only 0.15 Å in the positions of all 371 common C^{α} atoms and of 0.28 Å in the FAD atomic positions. The C-terminal residues of each monomer extend into its dimeric partner and lie on the surface of the molecule.

As expected from the sequence similarity, the topology of the MbtN monomer conforms to that of other ACAD enzymes. It is folded into three domains: an N-terminal α-helical domain ($\alpha1$–$\alpha7$), a middle β-sheet domain ($\beta1$–$\beta8$) and a C-terminal α-helical domain ($\alpha8$–$\alpha13$) (Fig. 2*a*). The most variable part of the MbtN structure corresponds to the $\beta4$–$\beta5$ connecting loop, the electron density for which is too disordered to be accurately modelled. This may reflect structural flexibility in this region, which is located at the putative substrate-entrance site.

3.2. FAD is reduced and chemically modified

After manual model building of the protein structure, FAD was modelled into well defined electron density that was present in the active sites of both monomers. Its position corresponds very closely to that of the FAD cofactors in other ACAD enzymes and it is held in place by numerous hydrogen bonds and hydrophobic interactions. The pyrimidine ring of the isoalloxazine moiety is hydrogen-bonded to the main-chain atoms of Ile118, Ala120, Ser121 and Ser153 and the side-

chain hydroxyl of Ser121 (Fig. 3c), whereas the dimethylbenzene ring sits in a hydrophobic pocket. The ribitol and ADP portions are hydrogen-bonded to residues Ser127, Thr358 and Glu360 from the same monomer and residues His328, Gly332 and Arg263 from the other monomer of the dimer (Fig. 3c).

Surprisingly, the isoalloxazine ring of FAD is significantly bent about its N(5)—N(10) axis, with an angle of 140–149° between the planes of its pyrimidine and dimethylbenzene rings (Table 2). This suggests that FAD is present in its reduced form, a conclusion supported by the fact that the crystals were colourless. During initial refinement tests, FAD was modelled in both oxidized and reduced conformations, but the σ_A-weighted $2F_o - F_c$ Fourier map clearly showed that the reduced conformation fitted better.

After rounds of refinement with reduced FAD, it became clear that a long tubular piece of electron density extending from the FAD N(5) atom was actually continuous with it, suggesting that an unknown ligand was covalently bonded to N(5) (Fig. 3a). It was concluded that this adduct was most likely to be a PEG fragment derived from the crystallization mother liquor and that it had reacted with the FAD during the time taken for crystallization or during data collection under X-ray irradiation. The UV–visible spectrum showed that the original protein solution contained oxidized FAD, a conclusion supported by mass spectrometry, which gave a mass corresponding to oxidized FAD (data not shown). Mass-spectral analyses of dissolved crystals were unsatisfactory, however, owing to the presence of numerous PEG fragments.

The density was therefore modelled as a fragment of PEG 3350 covalently bonded to N(5) of the FAD through a PEG C atom. PEG fragments with chain lengths of 15, 18 and 21 atoms covalently linked to N(5) were modelled in both monomers, with PEG 21 found to be the best fit to the density in monomer A and PEG 18 in monomer B. However, in both MbtN monomers the pocket in which the PEG moiety is located is not entirely filled (Fig. 3b). The length of the pocket is ~23 Å, with 6–7 Å separating the end of the longest PEG molecule from the side chains of Arg251 and Ile255 at the end of the pocket. This suggests that a slightly longer chain (for example, an acyl group with two additional atoms) could be accommodated. The pocket is almost exclusively hydrophobic, with Leu85, Gly355 and Gly356 surrounding the PEG at the start of the pocket followed along its length by Val81, Ala352, Ile241, Ala77, Met349 and Ile245 in the middle and Trp345, Ile323, Ala252, Ser248 and Ala75 near the end (Fig. 3c).

3.3. Comparisons with other ACAD structures

Structural superpositions of MbtN onto other members of the ACAD family give r.m.s.d. values from 2.2 to 2.7 Å for all C^α atoms, indicating the similarity in the overall fold. However, there are variations. Like FkbI, an acyl-ACP dehydrogenase from *Streptomyces hygroscopicus* (Watanabe *et al.*, 2003), MbtN has fewer residues at the N-terminus prior to $\alpha 1$ than most ACADs, in which this region contributes to the tetramer interface (Watanabe *et al.*, 2003). In contrast, MbtN has a C-terminal extension, including an extra helix ($\alpha 13$), which contributes to the tetramer interface instead; the penultimate residue at the C-terminus (Val380) packs with its side chain ~3–4 Å from the indole ring of Trp345 from an adjacent subunit. The helices $\alpha 1$–$\alpha 7$ and $\alpha 9$ are also shifted relative to the equivalent helices in most ACADs, which results in a wider opening to the substrate-entrance site.

An intriguing feature of the ACAD family is that the catalytic Glu residue that extracts the α-proton from the acyl-CoA substrate during catalysis is conserved only in its approximate spatial location, and not in its position in the sequence, despite significant overall sequence conservation. Structural comparisons identify Glu237, located on helix $\alpha 8$, as this essential base in MbtN, equivalent to the catalytic Glu261 in human LCAD (Djordjevic *et al.*, 1994) and Glu254 in human IVD (Tiffany *et al.*, 1997) (Figs. 4a and 5). Glu237 is appropriately located on the rectus (*re*) side of the flavin ring. Interestingly, *PROPKA* (Rostkowski *et al.*, 2011) analysis predicts that the pK_a of Glu237 in MbtN is ~7.8, which is significantly higher than the theoretical pK_a of ~4.5 for a free Glu, further supporting its assignment as a catalytic residue. Gly356, which is located in the loop between $\alpha 11$ and $\alpha 12$ in MbtN, corresponds spatially to the alternative Glu position occupied by Glu376

(a)

(b)

(c)
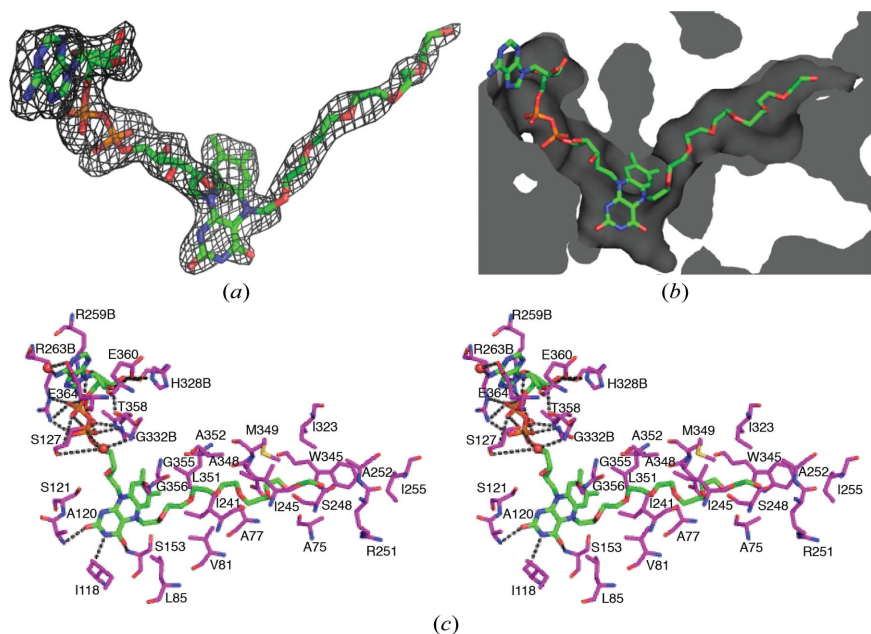
Figure 3
FAD adduct. (a) An $F_o - F_c$ OMIT map (black mesh) contoured at 2.0σ, showing the tubular density projecting from N(5) of FAD in monomer A. A covalently bound PEG fragment with a chain length of 21 C atoms (green sticks) was modelled here. (b) Cut-away surface diagram of the pocket in which the FAD adduct binds (the adduct is shown as green sticks), illustrating the length of the pocket. (c) Stereo diagram of the environment around the FAD–PEG adduct (green sticks) in monomer A. Protein residues are shown as magenta sticks, with dashed lines and red spheres indicating hydrogen bonds and water molecules, respectively.

in MCAD (Lee *et al.*, 1996), Glu368 in SCAD (Battaile *et al.*, 1996) and Glu422 in VLCAD (McAndrew *et al.*, 2008) (Figs. 4*a* and 5).

3.4. Acyl specificity and binding pocket

MbtN is predicted to catalyse the dehydrogenation of the acyl chain of mycobactin, which is linear and some 14–21 C atoms in length (Snow, 1965, 1970; Kastrinsky *et al.*, 2010).

Table 3
Summary of MbtN kinetics.

Acyl-CoA	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m (min^{-1} mM^{-1})
Myristoyl (C$_{14}$)	0.63 ± 0.09	0.062 ± 0.019	10.2 ± 4.7
Palmitoyl (C$_{16}$)	0.30 ± 0.02	0.010 ± 0.002	30 ± 10

(a)

(b)

Figure 4
Differently oriented acyl-binding pockets of MbtN and ACADs. (*a*) Overlay of FAD, catalytic Glu and various substrates (if present), all shown as sticks, for MbtN (green), human IVD (yellow; PDB entry 1ivh; Tiffany *et al.*, 1997), pig MCAD (red; PDB entry 1udy; Satoh *et al.*, 2003) and human VLCAD (cyan; PDB entry 3b9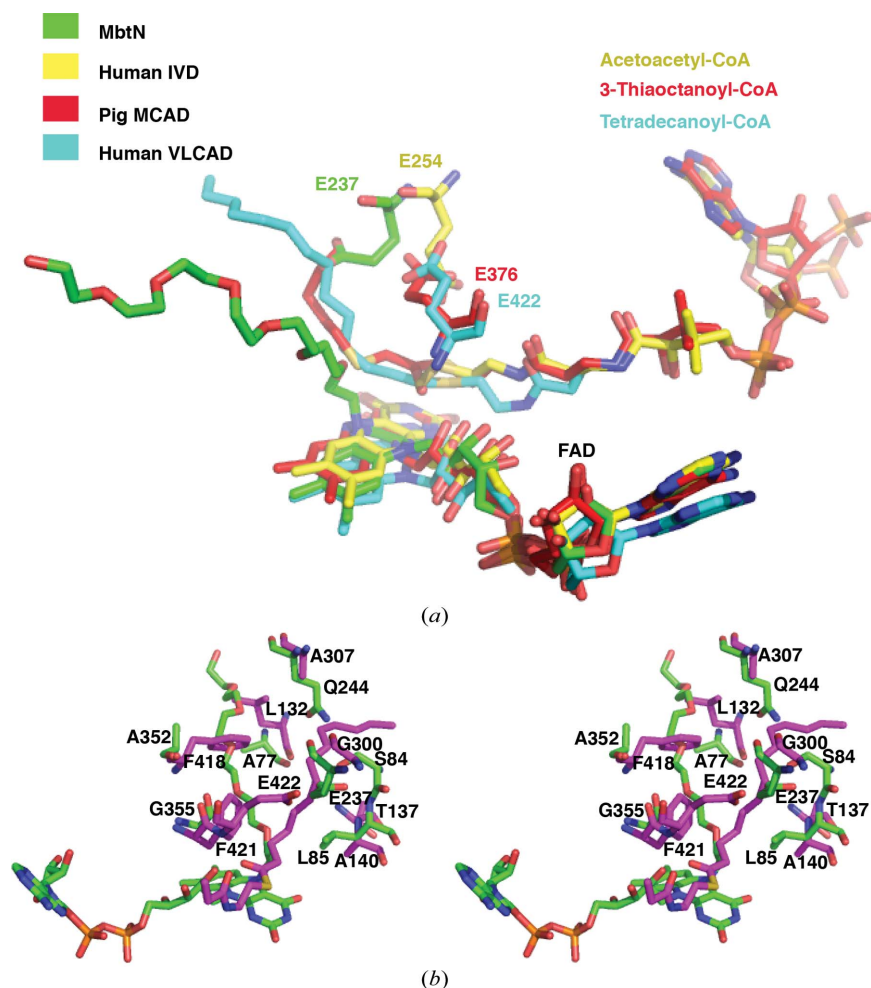6; McAndrew *et al.*, 2008). The FAD-binding site and catalytic Glu (in one of two positions) are conserved in the ACAD family. The orientation of the acyl-binding pocket (illustrated by the CoA substrates) is conserved for all but MbtN (shown by the PEG tail of the modified FAD). (*b*) Stereo diagram depicting the overlay of the acyl-binding pockets of MbtN (green) and VLCAD (PDB entry 3b96; magenta) and residues blocking the pockets. The putative acyl-binding pocket of MbtN has a different orientation from those of MCAD and VLCADs. Ligands and residues are shown as sticks.

Attempts to co-crystallize MbtN with lauroyl-CoA (C$_{12}$ acyl chain) and palmitoyl-CoA (C$_{16}$ acyl chain) were not successful. However, steady-state kinetic measurements were carried out for palmitoyl-CoA (C$_{16}$) and myristoyl-CoA (C$_{14}$). These showed a fivefold decrease in K_m and a threefold increase in catalytic efficiency (k_{cat}/K_m) for palmitoyl-CoA relative to myristol-CoA (Table 3), consistent with a preference for substrates with longer chain length.

A notable feature of MbtN is that the acyl-binding pocket, as defined by the bound PEG moiety, has a different orientation (relative to the FAD) from those in other currently structurally characterized ACADs (Fig. 4*b*). This results from sequence differences that open up alternative binding sites. In the MbtN acyl-binding pocket, residue 355 at the entrance appears to be a key 'gatekeeper'. Most ACADs characterized to date have a bulky hydrophobic residue, commonly Tyr or Phe, in this gatekeeper position (Phe421 in VLCAD; Figs. 4*b* and 5) which blocks the pocket; the Gly residue at this position in MbtN thus opens up a cavity that does not exist in other ACADs. The residues lining the pocket beyond the gatekeeper then determine its size and shape. One of these, the small residue Ala352 located in the middle of the binding pocket, is of note, as in other ACADs there is generally a bulky hydrophobic residue in the equivalent position (for example, Phe418 in VLCAD; Figs. 4*b* and 5).

Sequence differences in MbtN, notably involving Ser84, Leu85, Glu237 and Gln244, also occlude the acyl-binding pocket used by other ACADs (Fig. 4*b*). Leu85 is equivalent to smaller residues such as Gly100 in MCAD and Ala140 in VLCAD and affects the size of the opening to the pocket. The putative catalytic base in MbtN, Glu237, and its hydrogen-bond partner Ser84 block what in MCAD and VLCAD is the middle part of the acyl-binding pocket. In VLCAD the residue Ala307 helps to generate a deep pocket in which the long substrate can bind. In MbtN, which is expected to have a similar-sized substrate, the equivalent residue is the much larger Gln244. MCAD and IVD, which have shorter substrates that do not bind so deeply, also have larger residues (Gln95 and Met83, respectively) occupying a similar spatial position.

3.5. Preference for ACP over CoA as an acyl carrier

Acyl groups are typically carried either by acyl carrier proteins (ACPs) or CoA. Whereas most ACADs act on acyl-CoAs, MbtN is more likely to act on fatty acids
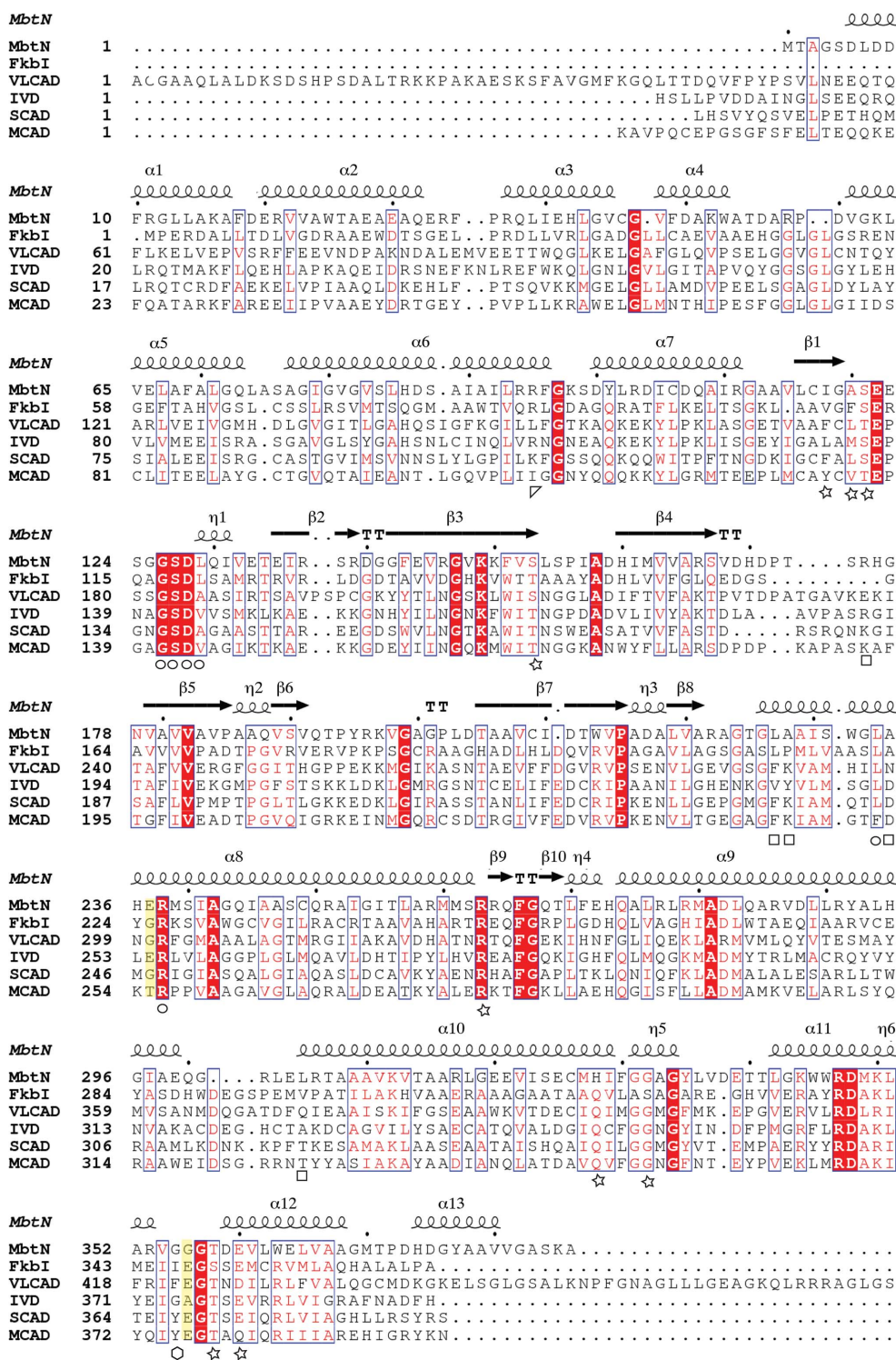
Figure 5 — sequence alignment (MbtN, FkbI, VLCAD, IVD, SCAD, MCAD)

```
MbtN                                                                    αααα
MbtN     1  ........................................................MTAGSDLDD
FkbI     1  .........................................................
VLCAD    1  ACGAAQLALDKSDSHPSDALTRKKPAKAESKSFAVGMFKGQLTTDQVFPYPSVLNEEQTQ
IVD      1  ..........................................HSLLPVDDAINGLSEEQRQ
SCAD     1  .....................................LHSVYQSVELPETHQM
MCAD     1  ................................KAVPQCEPGSGFSFELTEQQKE

              α1              α2                 α3            α4              αααα
MbtN         αααααααα   αααααααααααααα      αααααααααα     αααααααα       αααα
MbtN    10  FRGLLAKAFDERVVAWTAEAEAQERF..PRQLIEHLGVCG.VFDAKWATDARP..DVGKL
FkbI     1  .MPERDALLTDLVGDRAAEWDTSGEL..PRDLLVRLGADGLLCAEVAAEHGGLGLGSREN
VLCAD   61  FLKELVEPVSRFFEEVNDPAKNDALEMVEETTWQGLKELGAFLKELTSGKL..NCNTQY
IVD     20  LRQTMAKFLQEHLAPKAQEIDRSNEFKNLREFWKQLGNLGVLGITAPVQYGGSGLGYLEH
SCAD    17  LRQTCRDFAEKELVPIAAQLDKEHLF..PTSQVKKMGELGLLAMDVPEELSGAGLDYLAY
MCAD    23  FQATARKFAREEIIPVAAEYDRTGEY..PVPLLKRAWELGLMNTHIPESFGGLGLGIIDS

              α5                 α6               α7              β1
MbtN         αααααα          ααααααααααα.αααααα    ααααααααααα       ➝
MbtN    65  VELAFALGQLASAGTGVGVSLHDS.AIAILRRFGKSDYLRDICDQAIRGAAVLCIGASEE
FkbI    58  GEFTAHVGSLCSSLRSVMTSQGM.AAWTVQRLGDAGQRATFLKELTSGKL.AAVGFSER
VLCAD  121  ARLVEIVGMH.DLGVGITLGAHQSIGFKGILLFGTKAQKEKYLPKLASGETVAAFCLTEP
IVD     80  VLVMEEISRA.SGAVGLSYGAHSNLCINQLVRNGNEAQKEKYLPKLISGEYIGALMSEP
SCAD    75  SIALEEISRG.CASTGVIMSVNNSLYLGPILKFGSSQQKQQWITPFTNGDKIGCFALSEP
MCAD    81  CLITEELAYG.CTGVQTAIEANT.LGQVPLIIGGNYQQQKKYLGRMTEEPLMCAYCVTEP

              η1        β2        β3             β4
MbtN         ααα      ..➝ TT ➝➝➝➝➝        ➝➝ TT
MbtN   124  SGGSDLQIVETEIR..SRDGGFEVRGVKKFVSLSPIADHIMVVARSVDHDPT....SRHG
FkbI   115  QAGSDLSAMRTRVR..LDGDTAVVDGHKVWTTAAAYADHLVVFGLQEDGS.........G
VLCAD  180  SSGSDAASIRTSAVPSPCGKYYTLNGSKLWISNGGLADIFTVFAKTPVTDPATGAVKEKI
IVD    139  NAGSDVVSMKLKAE..KKGNHYILNGNKFWITNGPDADVLIVYAKTDLA...AVPASRGI
SCAD   134  GNGSDAGAASTTAR..EEGDSWVLNGTKAWITNSWEASATVVFASTD.....RSRQNKGI
MCAD   139  GAGSDVAGIKTKAE..KKGDEYIINGQKMWITGGGMANWYFLLARSDPDP..KAPASKAF
              oooo                              ☆                    □

              β5      η2  β6                 β7      η3  β8
MbtN         ➝➝➝➝    ααα ➝➝     TT         ➝➝➝➝  ααα ➝➝      ααααααα.αααα
MbtN   178  NVAVVAVPAAQVSVQTPYRKVGAGPLDTAAVCI.DTWVPADALVARAGTGLAAIS.WGLA
FkbI   164  AVVVVPADTPGVRVERVPKPSGCRAAGHADLHLDQVRVPAGAVLAGSGASLPMLVAASLA
VLCAD  240  TAFVVERGFGGITHGPPEKKMGIKASNTAEVFFDGVRVPSENVLGEVGSGFKVAM.HILN
IVD    194  TAFIVEKGMPGFSTSKKLDKLGMRGSNTCELIFEDCKIPAANILGHENKGVYVLM.SGLD
SCAD   187  SAFLVPMPTPGLTLGKKEDKLGIRASSTANLIFEDCRIPKENLLGEPGMGFKIAM.QTLD
MCAD   195  TGFIVEADTPGVQIGRKEINMGQRCSDTRGIVFEDVRVPKENVLTGEGAGFKIAM.GTFD
                                   ☆                                □□  □□  o□

              α8                                β9  β10 η4      α9
MbtN         αααααααααααααααααααααααααα           ➝ TT ααα  ααααααααααα
MbtN   236  HERMSIAGQIAASCQRATGITLARMMSRRQFGQTLFEHQALRLRMADIQARVDLLRYALH
FkbI   224  YGRKSVAWGCVGILRACRTAAVAHARTREQFGRPLGDHQLVAGHIADLWTAEQIAARVCE
VLCAD  299  NGRFGMAAALGTMRGIIAKAVDHATNRTQFGEKIHNFGLIQEKLARMVMLQYVTESMAY
IVD    253  LERLVLAGGPLGLMQAVLDHTIPYHVREAFGQKIFQLMQGKMADMYTRLMACRQYVY
SCAD   246  MGRIGIASQALGIAQASLDCAVKYAENRHAFGAPLTKLQNIQFKLADMALALESARLLTW
MCAD   254  KTRPPVAAGAVGLAQRALDEATKYALERKTFGKLLAEHQGISFLLADMAMKVELARLSYQ
              o                                      ☆                 □

                                 α10           η5           α11      η6
MbtN         αααα        ααααααααααααααααααα    ααα      ααααααααααα  αα
MbtN   296  GIAEQG.....RLELRTAAAVKVTAARLGEEVISECMHIFGGAGYLVDETTLGKWWRDMKL
FkbI   284  YASDHWDGSPEMVPATILAKHVAAERAAAGAAATAAQVLASAGAREGHVVERAYRDAKL
VLCAD  359  MVSANMDQGATDFQIEAAISKIFGSEAAWKVTDECIQIMGGMGFMK.EPGVERVLRDLRI
IVD    313  NVAKACDEG.HCTAKDCAGVILYSAECATQVALDGIQCFGGNGYIN.DFPMGRFLRDAKL
SCAD   306  RAAMLKDNK.KPFTKESAMAKLAASEAAATAISHQAIQILGGMGYVT.EMPAERYYRDARI
MCAD   314  RAAWEIDSG.RRNTYYASIAKAYAADIANQLATDAVQVFGGNGFNT.EYPVEKLMRDAKI
                               □                              ☆

              α12          α13
MbtN         αα   αααααααα   αααααααα
MbtN   352  ARVGGTDEVLWELVAAGMTPDHDGYAAVVGASKA...........
FkbI   343  MEIIEGSSEMCRVMLAQHALALPA..................
VLCAD  418  FRIFEGTNDILRLFVALQGCMDKGKELSGLGSALKNPFGNAGLLLGEAGKQLRRRAGLGS
IVD    371  YEIGAGTSEVRRLVIGRAFNADFH..................
SCAD   364  TEIYEGTSEIQRLVIAGHLLRSYRS.................
MCAD   372  YQIYEGTAQIQRIIIAREHIGRYKN.................
              ⬡  ☆  ☆
```

Figure 5
Sequence alignment of MbtN with pig MCAD (PDB entry 1udy; Satoh *et al.*, 2003), human MCAD (PDB entry 1egc; Lee *et al.*, 1996), rat SCAD (PDB entry 1jqi; Battaile *et al.*, 2002), human VLCAD (PDB entry 3b96; McAndrew *et al.*, 2008), human IVD (PDB entry 1ivh; Tiffany *et al.*, 1997) and FkbI (PDB entry 1r2j; Watanabe *et al.*, 2003). Residues important for ligand binding as reported previously (Watanabe *et al.*, 2003) or in this paper are indicated as follows: FAD binding, star; phosphopantetheine binding, circle; CoA binding, rectangle; ACP binding, triangle. The gatekeeper residue Gly355 is indicated by a hexagon. The two positions where the catalytic Glu can reside are highlighted in yellow. The secondary structure for MbtN is given above the sequences, as are dots marking every tenth MbtN residue. The sequence alignment was generated with *ESPript* (Robert & Gouet, 2014).

tethered to the *Mtb* ACP (Rv1344; MbtL), which is encoded in the same *mbt-2* gene cluster. The resulting unsaturated acyl-ACP would then be a substrate for the next enzyme in the biosynthetic pathway, MbtK (Rv1347c). MbtK has been demonstrated to have a 50-fold lower K_m for lauroyl-ACP compared with lauroyl-CoA (Krithika *et al.*, 2006), supporting an acyl-ACP substrate for MtbN. This preference is also strongly supported by the present MbtN structure, when compared with other ACADs.

Residues involved in interactions with the adenyl group of CoA in ACADs (Watanabe *et al.*, 2003) are not conserved in MbtN (Fig. 5). Notably, a key hydrogen bond between Asp253 and the N(6) atom of CoA in human MCAD would be lost in MbtN, in which Asp253 is replaced by Ala235. The same substitution occurs in FkbI, an ACAD enzyme involved in polyketide biosynthesis, which also acts on acyl-ACPs (Watanabe *et al.*, 2003). In contrast, as also seen in FkbI (Watanabe *et al.*, 2003), residues involved in binding the phosphopantetheine arm, which is present in both acyl-ACP and acyl-CoA, are conserved or semi-conserved in MbtN (Fig. 5).

Examination of the structures of ACP-dependent enzymes has identified a basic patch near the active-site channel that is involved in electrostatic interaction with the acidic helix II of ACP (Zhang *et al.*, 2001). In MbtN, a similar basic patch, involving residues from helix α6 and the α6–α7 loop, is adjacent to the entrance of the proposed active-site channel. This includes six basic residues Arg58, Arg94, Arg95, Lys98, Arg103 and Arg302 (Fig. 6). Arg95 in MbtN is spatially equivalent to Arg87 of FkbI (Fig. 7), a key interacting residue identified during docking of FkbI onto its partner ACP

(FkbJ; Watanabe *et al.*, 2003). This further supports the prediction that MbtN binds an acyl-ACP.

4. Discussion

Siderophores are essential for the survival of *Mtb* in host macrophages, which are low in soluble iron. Mycobactin T biosynthetic enzymes are therefore potential anti-TB drug targets. In this study, we have carried out structural analysis of *Mtb* MbtN and confirmed its activity as an acyl-ACP dehydrogenase. MbtN has the typical ACAD fold, with an FAD cofactor bound to each monomer, and like most ACAD family members is tetrameric in the crystal and in solution. The deep acyl-binding pocket and the kinetic data are consistent with MbtN having a preference for long fatty-acyl chains, consistent with the structure of myco-bactins, which have a long unsaturated fatty-acid chain. MbtL is most likely to be the acyl carrier in mycobactin biosynthesis and the results here suggest that MbtL interacts with MbtN through a basic patch on the surface of MbtN adjacent to the entry to the active site.

Features of particular interest are the formation of the covalent PEG adduct, with the accompanying distortion of the isoalloxazine ring of FAD, and the unusual *cis* double-bond formation associated with MbtN. Distortion of flavin geometry is a common feature of flavoproteins. Oxidized and one-electron-reduced semiquinone flavins typically adopt a planar conformation, whereas the two-electron-reduced flavin has a bent conformation (Lennon *et al.*, 1999). In the crystal structure described here, the flavin was evidently in its reduced state, as shown by the lack of yellow colour, but reduction of FAD alone is not usually enough to produce such a significant bend (>30° distortion from planarity); typical bending owing to reduction is in the range 8–25°. Covalent bonding of the PEG at the N(5) position may have contributed to the larger bend in the isoalloxazine ring. Covalent modification at N(5) is not unprecedented, and has been observed in D-arginine dehydrogenase (Fu *et al.*, 2011), pyranose 2-oxidase (Tan *et al.*, 2010), sarcosine oxidase (Chen *et al.*,

2005), adenosine-5-phosphosulfate reductase (Schiffer *et al.*, 2006), histone demethylase LSD1 (Binda *et al.*, 2010), D-amino-acid oxidase (Todone *et al.*, 1997) and nitroalkane oxidase (Nagpal *et al.*, 2006) (see Table 2 for a summary). All flavins in these structures have a nonplanar isoalloxazine ring with the adduct facing the rectus (*re*) side of the flavin ring.
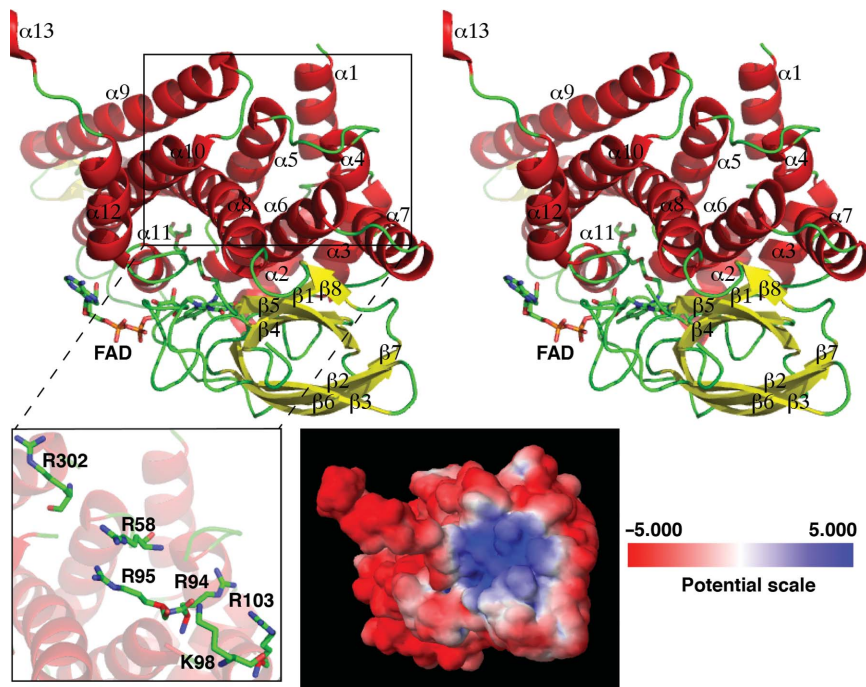

Figure 6
Proposed binding site for ACP. The top figure shows a cartoon monomer of MbtN in stereo, with secondary-structure elements labelled. The inset shows the basic residues in MbtN that are proposed to be important for electrostatic interaction with ACP. Below, a surface diagram showing the electrostatic potential. MbtN is in the same orientation as in the top figure.
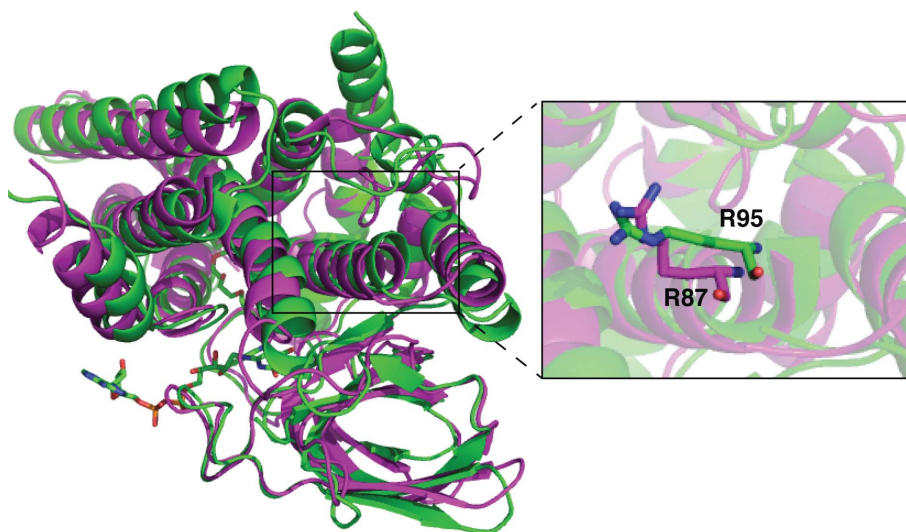

Figure 7
Overlay of the two acyl-ACP-binding proteins MbtN and FkbI in cartoon representation (FkbI in purple and MbtN in green) in a similar orientation to that in Fig. 6. Inset: a close-up view of the putative ACP-binding site, with Arg87 in FkbI, identified as likely to be important for ACP binding, and its MbtN equivalent Arg95 shown as sticks.

Covalent modification of the flavin moiety is a major form of inhibition of ACAD enzymes (Wenz *et al.*, 1981). Various mechanism-based inhibitors of ACADs have been reported, such as methylenecyclopropylacetyl-CoA (MCPA), spiropentylacetyl-CoA (SPA-CoA), 3-methyleneoctanoyl-CoA, 3-methyl-2-octenoyl-CoA and 3,4-alkadienoyl-CoA (Ghisla & Thorpe, 2004; Zeng *et al.*, 2006). Using rapid-reaction spectrophotometry, Wang and coworkers showed that 3,4-dienoyl-CoA derivatives could inhibit MCAD through the formation of a covalent flavin N(5) adduct (Wang *et al.*, 2001). These suicide inhibitors form an enolate intermediate that cannot efficiently discharge a hydride equivalent, meaning that the reduction of the flavin is no longer concerted as it is in the presence of normal substrates (Wang *et al.*, 2001). Wenz and coworkers reported the inhibition of pig MCAD by hypoglycin-A *via* the generation of a transient α-carbanion species that leads to covalent addition of the substrate to the flavin N(5) atom (Wenz *et al.*, 1981). Gomes and coworkers also proposed the carbanion mechanism as the inhibition pathway of butyryl-CoA dehydrogenase by 3-butynoylpantetheine (Gomes *et al.*, 1981).

Flavins can react with photoelectrons generated by X-radiation during data collection. Orru and coworkers have shown that crystalline flavoenzymes are redox-reactive and can be rapidly reduced by X-ray irradiation (Orru *et al.*, 2011). The covalent adducts in D-arginine dehydrogenase (DADH; Fu *et al.*, 2011) and cholesterol oxidase (Lyubimov *et al.*, 2007) are also thought to have arisen from a nucleophilic substitution in which the N(5) atom of the radiation-reduced flavin reacted with the C^{α} atom of the substrates. If the flavin N(5) adduct is formed *via* photoreduction, it represents a process separate from the normal catalytic mechanism. Alternatively, the PEG fragment in the structure here could act as a suicide or mechanism-based inhibitor and be representative of a type of complex formed during 'normal' catalysis.

The role of MbtN in mycobactin biosynthesis, and the observed *cis* double bond in the fatty-acyl moiety of mycobactin, raises intriguing questions about mechanism and stereospecificity. The stereochemistry of α,β-dehydrogenation in ACAD enzymes is always *pro-2R/pro-3R*, placing the α and β H atoms in the conformation required for an antiperiplanar reaction to generate a *trans*-2-enoyl-ACP product. The absence of an isomerase in the vicinity of the *mbt* and *mbt-2* gene clusters suggests that MbtN could be responsible for the production of the *cis* double bond. To generate a *cis*-2-enoyl-ACP product, MbtN could either utilize a two-step mechanism, in which a *trans* isomer is generated in the first step and is then converted to the *cis* form, or catalyze direct *cis* double-bond formation. Transition-state theory argues that it is less energetically expensive to utilize a substrate that resembles the final product conformation than to convert an intermediate to a product with a totally different conformation. This would favour direct *cis* double-bond formation by MbtN.

The different orientation of the acyl-binding pocket in MbtN, compared with those of other structurally characterized ACAD enzymes, may be a key element in this question. A substrate-binding site that is complementary to a *cis*-isomer could induce a substrate conformation suitable for *cis* double-bond formation, and lower the energy barrier for the formation of a *cis* product. In this case, α,β-dehydrogenation could still occur in the manner known for other ACADs, with the α and β H atoms still being removed in an antiperiplanar fashion, but with the pre-formed *cis*-like substrate conformation of the substrate ensuring *cis* double-bond formation. In an unrelated enzyme, Δ9-stearoyl ACP desaturase, a 'boomerang-shaped' binding cavity does indeed appear to affect stereoselectivity, enabling the formation of a product with a *cis* double bond (Guy *et al.*, 2011).

Our MbtN structure has underlined the importance of the gatekeeper residue 355 (MbtN numbering). In this case it is key to the different orientation of the acyl-binding pocket, but it can also be critical to acyl specificity, and this may aid in the identification of potential substrates for the as yet uncharacterized ACADs. Thus, Nishina and coworkers postulated that a tyrosine at this position favoured straight-chain specificity (Nishina *et al.*, 1995). In contrast, IVD (isovalerate dehydrogenase), which produces a short-chain branched product that is neither *cis* nor *trans*, is unusual amongst ACADs as it shares with MbtN the presence of a glycine at the gatekeeper position (Gly374; Tiffany *et al.*, 1997). In IVD this glycine allows the acyl-binding pocket to widen to accommodate the branched substrate (Tiffany *et al.*, 1997). In another branched short-chain ACAD, iBD, the gatekeeper residue is a Leu, which is proposed to affect the shape of the binding site, conferring its specificity for 2-methyl branched substrates (Battaile *et al.*, 2004). FkbI, which has an Ile in the gatekeeper position, also acts on a C^{α}-substituted substrate (Watanabe *et al.*, 2003).

As noted earlier, ACADs have two different spatial positions for the catalytic Glu. This may also be significant in helping to determine the stereochemistry of the final product. In IVD the Glu position is proposed to allow the correct approach and angle for abstraction of the *pro-R* C^{α} and C^{β} H atoms of the branched substrate (Tiffany *et al.*, 1997). Indeed, when IVD was mutated to shift the catalytic Glu to the location seen in MCAD and VLCAD the ability of the enzyme to act on branched-chain substrates was greatly reduced (to 4% of that of the wild type), but it retained an ability to accept straight-chain substrates (Tiffany *et al.*, 1997). MbtN shares with IVD the same spatial position for its catalytic Glu, so this observation is of potential significance. LCAD is another enzyme in which the catalytic Glu is in this position. Initially proposed to act on long-chain substrates, it has since been found this enzyme is likely to prefer branched-chain substrates such as 2,6-dimethylheptanoic-CoA (Wanders *et al.*, 1998).

Finally, the interaction of MbtN with the ACP portion of its acyl-ACP substrate may also help to direct the formation of the desired isomer, as suggested by Krithika *et al.* (2006) and as postulated for the family of acyl-ACP desaturases (Guy *et al.*, 2011). In MbtN, the residues predicted to be involved in ACP binding co-localize with the residues involved in active-site formation, catalytic activity and the acyl-binding pocket. In particular, Arg94, Arg95 and Lys98 (on the $\alpha6/\alpha7$ loop),

which are proposed to interact with ACP, are located just beyond Leu85, Ser84 and Ser88 (on helix α6), which influence the direction of the acyl-binding pocket in MbtN and/or hydrogen-bond to the catalytic Glu237. Other residues from helix α6 are also within 4 Å of the PEG adduct in the acyl-binding pocket. It is therefore plausible that ACP binding to MbtN helps to fashion the active site and acyl-binding pocket.

In summary, we have shown that the structure of MbtN is consistent with that of an acyl-ACP dehydrogenase rather than an acyl-CoA dehydrogenase. Furthermore, the capture of a covalently modified FAD identified the likely acyl-binding pocket of the enzyme and suggests that MbtN would be susceptible to inhibition by flavin modification. The pocket size is consistent with that expected for the mycobactin acyl-chain product. Importantly, the orientation of the pocket differs from those of the other ACADs structurally characterized to date, suggesting a mechanism to promote direct *cis* double-bond formation. While we have not been able to confirm a structural basis for *cis* double-bond formation, this study provides a framework on which to base future work on the catalytic activity of MbtN.

References

Battaile, K. P., Mohsen, A.-W. A. & Vockley, J. (1996). *Biochemistry*, **35**, 15356–15363.

Battaile, K. P., Molin-Case, J., Paschke, R., Wang, M., Bennett, D., Vockley, J. & Kim, J.-J. P. (2002). *J. Biol. Chem.* **277**, 12200–12207.

Battaile, K. P., Nguyen, T. V., Vockley, J. & Kim, J.-J. P. (2004). *J. Biol. Chem.* **279**, 16526–16534.

Binda, C., Valente, S., Romanenghi, M., Pilotto, S., Cirilli, R., Karytinos, A., Ciossani, G., Botrugno, O. A., Forneris, F., Tardugno, M., Edmondson, D. E., Minucci, S., Mattevi, A. & Mai, A. (2010). *J. Am. Chem. Soc.* **132**, 6827–6833.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C. & Womack, T. O. (2011). *BUSTER* v. 2.10.0. Cambridge: Global Phasing Ltd.

Chai, A.-F., Johnston, J. M., Bunker, R. D., Bulloch, E. M. M., Evans, G. L., Lott, J. S. & Baker, E. N. (2013). *Acta Cryst.* F**69**, 1354–1356.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Chen, Z.-W., Zhao, G., Martinovic, S., Jorns, M. S. & Mathews, F. S. (2005). *Biochemistry*, **44**, 15444–15450.

Cole, S. T. & Barrell, B. G. (1998). *Novartis Found. Symp.* **217**, 160–177.

Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.

DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.

De Voss, J. J., Rutter, K., Schroeder, B. G., Su, H., Zhu, Y. & Barry, C. E. III (2000). *Proc. Natl Acad. Sci. USA*, **97**, 1252–1257.

Dhatwalia, R., Singh, H., Oppenheimer, M., Karr, D. B., Nix, J. C., Sobrado, P. & Tanner, J. J. (2012). *J. Biol. Chem.* **287**, 9041–9051.

Djordjevic, S., Dong, Y., Paschke, R., Frerman, F. E., Strauss, A. W. & Kim, J. J.-P. (1994). *Biochemistry*, **33**, 4258–4264.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Finnegan, S., Agniswamy, J., Weber, I. T. & Gadda, G. (2010). *Biochemistry*, **49**, 2952–2961.

Fu, G., Yuan, H., Wang, S., Gadda, G. & Weber, I. T. (2011). *Biochemistry*, **50**, 6292–6294.

Ghisla, S. & Thorpe, C. (2004). *Eur. J. Biochem.* **271**, 494–508.

Gomes, B., Fendrich, G. & Abeles, R. H. (1981). *Biochemistry*, **20**, 1481–1490.

Guy, J. E., Whittle, E., Moche, M., Lengqvist, J., Lindqvist, Y. & Shanklin, J. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 16594–16599.

Kastrinsky, D. B., McBride, N. S., Backus, K. M., LeBlanc, J. J. & Barry, C. E. III (2010). *Comprehensive Natural Products II: Chemistry and Biology*, edited by L. Mander & H.-W. Liu, Vol. 1, pp. 65–145. Oxford: Elsevier. doi:10.1016/B978-008045382-8.00029-0.

Kim, J.-J. P., Wang, M. & Paschke, R. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 7523–7527.

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

Krithika, R., Marathe, U., Saxena, P., Ansari, M. Z., Mohanty, D. & Gokhale, R. S. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 2069–2074.

Lee, H.-J. K., Wang, M., Paschke, R., Nandy, A., Ghisla, S. & Kim, J.-J. P. (1996). *Biochemistry*, **35**, 12412–12420.

Lennon, B. W., Williams, C. H. Jr & Ludwig, M. L. (1999). *Protein Sci.* **8**, 2366–2379.

Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* D**64**, 125–132.

Luo, M., Arentson, B. W., Srivastava, D., Becker, D. F. & Tanner, J. J. (2012). *Biochemistry*, **51**, 10099–10108.

Lyubimov, A. Y., Heard, K., Tang, H., Sampson, N. S. & Vrielink, A. (2007). *Protein Sci.* **16**, 2647–2656.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

McAndrew, R. P., Wang, Y., Mohsen, A.-W. A., He, M., Vockley, J. & Kim, J.-J. P. (2008). *J. Biol. Chem.* **283**, 9435–9443.

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

Miethke, M. & Marahiel, M. A. (2007). *Microbiol. Mol. Biol. Rev.* **71**, 413–451.

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

Nagpal, A., Valley, M. P., Fitzpatrick, P. F. & Orville, A. M. (2006). *Biochemistry*, **45**, 1138–1150.

Nishina, Y., Sato, K., Hazekawa, I. & Shiga, K. (1995). *J. Biochem.* **117**, 800–808.

Orru, R., Dudek, H. M., Martinoli, C., Torres Pazmiño, D. E., Royant, A., Weik, M., Fraaije, M. W. & Mattevi, A. (2011). *J. Biol. Chem.* **286**, 29284–29291.

Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.

Rostkowski, M., Olsson, M. H. M., Søndergaard, C. R. & Jensen, J. H. (2011). *BMC Struct. Biol.* **11**, 6.

Sassetti, C. M. & Rubin, E. J. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 12989–12994.

Satoh, A., Nakajima, Y., Miyahara, I., Hirotsu, K., Tanaka, T., Nishina, Y., Shiga, K., Tamaoki, H., Setoyama, C. & Miura, R. (2003). *J. Biochem.* **134**, 297–304.

Schiffer, A., Fritz, G., Kroneck, P. M. & Ermler, U. (2006). *Biochemistry*, **45**, 2960–2967.

Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). *Acta Cryst.* D**60**, 1355–1363.

Snow, G. A. (1965). *Biochem. J.* **97**, 166–175.

Snow, G. A. (1970). *Bacteriol. Rev.* **34**, 99–125.

Tan, T. C., Pitsawong, W., Wongnate, T., Spadiut, O., Haltrich, D., Chaiyen, P. & Divne, C. (2010). *J. Mol. Biol.* **402**, 578–594.

Thorpe, C. & Kim, J.-J. P. (1995). *FASEB J.* **9**, 718–725.

Tiffany, K. A., Roberts, D. L., Wang, M., Paschke, R., Mohsen, A.-W. A., Vockley, J. & Kim, J.-J. P. (1997). *Biochemistry*, **36**, 8455–8464.

Todone, F., Vanoni, M. A., Mozzarelli, A., Bolognesi, M., Coda, A., Curti, B. & Mattevi, A. (1997). *Biochemistry*, **36**, 5853–5860.

Wanders, R. J. A., Denis, S., Ruiter, J. P. N., IJlst, L. & Dacremont, G. (1998). *Biochim. Biophys. Acta*, **1393**, 35–40.

Wang, W. Z., Fu, Z. J., Zhou, J. Z., Kim, J.-J. P. & Thorpe, C. (2001). *Biochemistry*, **40**, 12266–12275.

Watanabe, K., Khosla, C., Stroud, R. M. & Tsai, S.-C. (2003). *J. Mol. Biol.* **334**, 435–444.

Wenz, A., Thorpe, C. & Ghisla, S. (1981). *J. Biol. Chem.* **256**, 9809–9812.

Zeng, J., Deng, G., Yu, W. & Li, D. (2006). *Bioorg. Med. Chem. Lett.* **16**, 1445–1448.

Zhang, Y.-M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O. & White, S. W. (2001). *J. Biol. Chem.* **276**, 8231–8238.

Zumla, A., George, A., Sharma, V., Herbert, N. & Baroness Masham of Ilton (2013). *Lancet*, **382**, 1765–1767.