Structural rearrangements occurring upon cofactor binding in the *Mycobacterium smegmatis* β-ketoacyl-acyl carrier protein reductase MabA

Tanja Küssau,[a] Marion Flipo,[b] Niel Van Wyk,[a]‡ Albertus Viljoen,[a] Vincent Olieric,[c] Laurent Kremer[a,d] and Mickaël Blaise[a]*

[a]Institut de Recherche en Infectiologie de Montpellier (IRIM), Université de Montpellier, CNRS UMR 9004, 34293 Montpellier, France, [b]Université de Lille, INSERM, Institut Pasteur de Lille, U1177 – Drugs and Molecules for Living Systems, 59000 Lille, France, [c]Swiss Light Source, Paul Scherrer Institute, 5232 Villigen, Switzerland, and [d]INSERM, IRIM, 34293 Montpellier, France. *Correspondence e-mail: mickael.blaise@irim.cnrs.fr

In mycobacteria, the ketoacyl-acyl carrier protein (ACP) reductase MabA (designated FabG in other bacteria) catalyzes the NADPH-dependent reduction of β-ketoacyl-ACP substrates to β-hydroxyacyl-ACP products. This first reductive step in the fatty-acid biosynthesis elongation cycle is essential for bacteria, which makes MabA/FabG an interesting drug target. To date, however, very few molecules targeting FabG have been discovered and MabA remains the only enzyme of the mycobacterial type II fatty-acid synthase that lacks specific inhibitors. Despite the existence of several MabA/FabG crystal structures, the structural rearrangement that occurs upon cofactor binding is still not fully understood. Therefore, unlocking this knowledge gap could help in the design of new inhibitors. Here, high-resolution crystal structures of MabA from *Mycobacterium smegmatis* in its apo, NADP⁺-bound and NADPH-bound forms are reported. Comparison of these crystal structures reveals the structural reorganization of the lid region covering the active site of the enzyme. The crystal structure of the apo form revealed numerous residues that trigger steric hindrance to the binding of NADPH and substrate. Upon NADPH binding, these residues are pushed away from the active site, allowing the enzyme to adopt an open conformation. The transition from an NADPH-bound to an NADP⁺-bound form is likely to facilitate release of the product. These results may be useful for subsequent rational drug design and/or for *in silico* drug-screening approaches targeting MabA/FabG.

1. Introduction

Mycobacterium tuberculosis (*Mtb*), the aetiological agent of tuberculosis (TB), is one of the deadliest human pathogens and is responsible for about 10.5 million new infections and 1.4 million deaths annually (World Health Organization, 2016). This high rate of mortality, despite existing chemotherapy, is owing to the lack of an efficient vaccine, co-infection with HIV and the emergence of multidrug-resistant, extensively drug-resistant and, more recently, totally drug-resistant strains of *Mtb* that considerably limit treatment options. Therefore, the development of new therapies, including new anti-TB drugs, is needed more than ever. The high impermeability of the mycobacterial cell wall, however, represents a serious obstacle to the identification of active compounds. The hydrophobicity of the mycomembrane is essentially caused by the presence of mycolic acids, consisting of very long α-alkyl, β-hydroxy fatty acids, comprising up to 90 C atoms (Pawełczyk & Kremer,
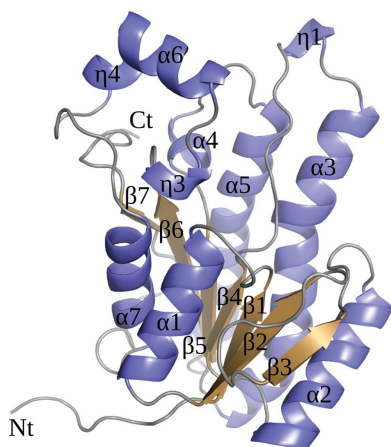
2014). The complex machinery responsible for mycolic acid biosynthesis involves numerous enzymes, some of which are the target of important antitubercular drugs, while others represent potent pharmacological targets that require further exploitation (Quémard, 2016).

The biosynthesis of mycolic acids is driven by two distinct elongation systems: the so-called type I and type II fatty-acid synthases (FAS I and FAS II, respectively; Bhatt, Molle *et al.*, 2007; Quémard, 2016; Pawełczyk & Kremer, 2014). FAS I consists of a single polypeptide, resembling the eukaryotic FAS (Boehringer *et al.*, 2013), which possesses all of the catalytic domains for the *de novo* synthesis of fatty acids and shows a bimodal activity that leads to the production of palmitic acid (C_{16}) and hexacosanoic acid (C_{26}), also referred to as the α-branch. In contrast to FAS I, FAS II is composed of multiple enzymes with discrete functions that interact to further elongate the C_{16} precursor generated by FAS I to produce meromycolic acids containing up to 56 C atoms. Meromycolates can be functionalized by the introduction of chemical modifications, which is performed by a diverse family of SAM-dependent methyltransferases. Following functionalization, meromycolates are condensed together with the α-branch by the polyketide synthase Pks13 to produce mycolic acids. Like the FAS II complexes found in other prokaryotes or in plants, the mycobacterial FAS II system involves four distinct enzymes that work successively and in an iterative manner to catalyze the elongation step of the meromycolic acid. In *Mtb* this process is initiated by MabA (*Rv1483*; designated FabG in other bacteria), a β-ketoacyl-acyl carrier protein (ACP) reductase that catalyzes the conversion of β-ketoacyl-ACP to β-hydroxyacyl-ACP (Marrakchi *et al.*, 2002; Banerjee *et al.*, 1998), which is subsequently converted by the heterodimeric (3R)-hydroxyacyl-ACP dehydratases HadAB and HadBC (*Rv0635–Rv0637*) to *trans*-2-enoyl-ACP (Sacco *et al.*, 2007). InhA (*Rv1484*), a *trans*-2-enoyl-ACP reductase, converts *trans*-2-enoyl-ACP to saturated acyl-ACP products (Quémard *et al.*, 1995; Banerjee *et al.*, 1994). Finally, the β-ketoacyl ACP synthases KasA (*Rv2245*) and KasB (*Rv2246*) elongate the acyl-ACP substrates by two additional C atoms to generate β-ketoacyl ACP fatty acids (Kremer *et al.*, 2002; Bhatt *et al.*, 2005; Bhatt, Fujiwara *et al.*, 2007), which are again reduced by MabA when entering a second round of the cycle. After each reductive cycle, the growing fatty-acyl chain is elongated by two additional C atoms. All four enzymes have been shown to be essential for mycobacterial viability, thereby representing potent drug targets. Isoniazid, the cornerstone of TB therapy, inhibits InhA, whereas thioacetazone and isoxyl, which were formerly used for TB treatment, inhibit the HadAB/BC complexes (Gannoun-Zaki *et al.*, 2013; Coxon *et al.*, 2013; Grzegorzewicz *et al.*, 2012, 2015). Several chemical entities such as thiolactomycin or indazole sulfonamide have been shown to inhibit KasA (Kapilashrami *et al.*, 2013; Kremer *et al.*, 2000; Abrahams *et al.*, 2016). However, no specific inhibitor of MabA has been reported.

FabG/MabA belongs to the short-chain dehydrogenases/reductase (SDR) superfamily, characterized by the presence of a central Rossmann fold and the use of NADH or NADPH as a cofactor. SDR proteins have been shown to undergo structural rearrangements upon binding of the cofactor. Based on high-resolution crystal structures of the apo and holo forms of *Mtb* MabA (MabA$_{Mtb}$), it was proposed that similar structural rearrangements are likely to occur in this protein (Cohen-Gonsaud *et al.*, 2002, 2005). However, our understanding of this mechanism still remains obscured by the fact that parts of the structure could not be modelled in the apo structure of MabA, while in the MabA–NADP$^+$ complex the nicotinamide group of the cofactor could not be modelled as it was not visible in the electron density. The difficulty in visualizing the flexible region covering the active site in the crystal structures of the SDR protein family, and in particular in FabG proteins, remains a recurrent issue. Recently, this region was built in the apo form of *Vibrio cholerae* FabG (Hou *et al.*, 2015). The complete apo and holo structures of *Pseudomonas aeruginosa* FabG have also been described, revealing a reduced mobility of this flexible region when the cofactor was bound (Cukier *et al.*, 2013). We have also shown that in the MabA-like protein MSMEG_6753 from *M. smegmatis*, a nonpathogenic mycobacterial species, the same region was very dynamic and the binding of NADP$^+$ was associated with important structural changes (Blaise *et al.*, 2017). Together, these two studies suggest that upon binding of the cofactor the so-called flexible loop might adopt a closed conformation.

Considering the potential of MabA as an attractive drug target, a thorough comparison of the complete apo and holo crystal structures would be particularly helpful for subsequent target-based drug-discovery approaches as well as for the design of specific inhibitors and structure–activity relationship studies. Here, we investigated the structural dynamics that occur upon the binding of cofactor by *M. smegmatis* MabA (MabA$_{MSMEG}$).

2. Materials and methods

2.1. Cloning, protein expression and purification

The *mabA$_{MSMEG}$* (*MSMEG_3150*) gene was cloned into pET-30 Ek/LIC as described previously (Blaise *et al.*, 2017). The protein was expressed in *Escherichia coli* BL21 Rosetta 2 (DE3) pLysS cells (Novagen). Transformants were grown at 37°C and protein expression was induced with 1 m*M* isopropyl β-D-1-thiogalactopyranoside during exponential growth for 3 h. The cells were harvested by centrifugation at 6000*g* for 15 min, resuspended in lysis buffer (50 m*M* Tris–HCl pH 8, 400 m*M* NaCl, 5 m*M* β-mercaptoethanol, 1 m*M* benzamidine) and disrupted by sonication. The cell debris was centrifuged at 30 000*g* for 45 min and the supernatant was loaded by gravity onto nickel Sepharose beads (IMAC). A first wash step was performed using lysis buffer and a second wash used a buffer consisting of 50 m*M* Tris–HCl pH 8, 1 *M* NaCl, 5 m*M* β-mercaptoethanol. Proteins were eluted with 50 m*M* Tris–HCl pH 8, 400 m*M* NaCl, 5 m*M* β-mercaptoethanol, 250 m*M* imidazole. The eluted proteins were then dialyzed at 4°C overnight against 50 m*M* Tris–HCl pH 8, 200 m*M* NaCl, 5 m*M*

β-mercaptoethanol. To cleave the N-terminal tags, recombinant His-tagged *Tobacco etch virus* (TEV) protease was added to the protein solution at a ratio of 1 mg protease to 50 mg eluted protein. The proper removal of the N-terminal tags was assessed by Western blotting using anti-His and anti-S-tag primary antibodies. The tag-less proteins were then loaded onto an IMAC column pre-equilibrated with dialysis buffer. MabA$_{MSMEG}$ was collected in the flowthrough and concentrated to 5 mg ml^{-1}. The final purification step was performed on an ENrich SEC 650 size-exclusion chromatography column (Bio-Rad), which was eluted with 50 mM Tris–HCl pH 8, 200 mM NaCl, 5 mM β-mercaptoethanol. The purity of the protein was estimated to be 95% by Coomassie Blue-stained SDS–PAGE. Proteins obtained from these three chromatographic steps were used for determination of the kinetic constants as well as for crystallization experiments using the apo and NADPH-bound forms. The NADP$^+$-bound crystals were obtained using MabA$_{MSMEG}$ possessing its N-terminal tags and purified after a single step of IMAC followed by dialysis.

2.2. Enzymatic assays

NADPH and acetoacetyl-CoA were purchased from Sigma–Aldrich. The reductase activity was determined by monitoring the decrease in the optical density at 340 nm corresponding to the oxidation of NADPH to NADP$^+$. Kinetic constants for NADPH, acetoacetyl-CoA and β-keto-octanoyl-CoA were determined with a fixed concentration of one component (1 mM for NADPH and 2 mM for acetoacetyl-CoA) and a range of concentrations of the second substrate. Assays were performed in a volume of 20 µl to determine the kinetic constants for acetoacetyl-CoA and β-keto-octanoyl-CoA or in 80 µl for NADPH. Reaction mixtures consisted of 50 mM potassium phosphate buffer pH 7.0, the concentrations of substrates stated above and between 0.25 and 0.1 µM MabA$_{MSMEG}$. The enzymatic reactions were initiated by the addition of acetoacetyl-CoA after incubation for 5 min at 25°C. The oxidation of NADPH was followed by removing 1.5 µl aliquots of the reaction mixture for analysis in a NanoDrop 2000c spectrophotometer (path length 1 mm; Thermo Scientific) to determine the kinetic constants of the β-ketoacyl-CoAs and by continuous reading in an 80 µl quartz cuvette to determine the K_m for NADPH in a NanoDrop 2000c spectrophotometer (path length 10 mm; Thermo Scientific). Experiments were performed in triplicate. Data were fitted using nonlinear least-squares regression with *GraphPad Prism*.

2.3. Determination of molecular weight by size-exclusion chromatography

The oligomeric state of MabA$_{MSMEG}$ in solution was assessed on an ENrich SEC 650 gel-filtration column (Bio-Rad) eluted with 50 mM Tris–HCl pH 8, 200 mM NaCl, 5 mM β-mercaptoethanol at a flow rate of 0.5 ml min^{-1} at 4°C. Different concentrations of protein were assessed. The molecular weight was determined based on a calibration curve obtained using the Gel Filtration Markers Kit for Protein Molecular Weights 12 400–200 000 Da (Sigma–Aldrich) and dextran blue to assess the column void volume. The apparent mass was obtained by plotting the partition coefficient K_{av} against the logarithms of the molecular weights of standard proteins.

2.4. Crystallization

The apo MabA$_{MSMEG}$ crystals were grown in sitting drops in MR Crystallization Plates (Hampton Research) at 18°C by mixing 1.5 µl protein solution concentrated to 6 mg ml^{-1} with 1.5 µl reservoir solution consisting of 100 mM bis-tris propane pH 7, 2.7 M ammonium sulfate, 5% dimethylsulfoxide (DMSO) and using the microseeding technique. These crystals were briefly soaked in 100 mM bis-tris propane pH 7, 2.7 M ammonium sulfate, 5% DMSO, 20% glycerol prior to cryo-cooling in liquid nitrogen. The MabA$_{MSMEG}$–NADP$^+$ complex was co-crystallized at 7.7 mg ml^{-1} in the presence of 5 mM NADP$^+$ by mixing 0.8 µl protein solution with 0.8 µl reservoir solution consisting of 100 mM sodium citrate pH 5.6, 1.0 M lithium sulfate, 0.5 M ammonium sulfate. The MabA$_{MSMEG}$–NADPH crystals were obtained by mixing 0.8 µl protein solution concentrated to 7.5 mg ml^{-1} and containing 5 mM NADPH with 0.8 µl reservoir solution consisting of 0.1 M bis-tris pH 5.5, 0.2 M NaCl, 25% PEG 3350. All crystallization conditions for these complexes were obtained in sitting drops in 96-well SWISSCI MRC plates (Molecular Dimensions) at 18°C. Crystals obtained in the presence of cofactors were cryocooled without any additional cryoprotection.

2.5. Data collection, structure determination and refinement

The apo MabA$_{MSMEG}$ and MabA$_{MSMEG}$–NADP$^+$ data sets were collected on the ID30B and ID23-2 beamlines, respectively, at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. The MabA$_{MSMEG}$–NADPH data set was collected on the X06DA-PXIII beamline at the Swiss Light Source (SLS), Villigen, Switzerland. Data were processed with *XDS* and scaled and merged with *XSCALE* (Kabsch, 2010). As the crystal used to collect the MabA$_{MSMEG}$–NADPH data set was not cryoprotected, the ice-ring zones were excluded during data processing, which explains the relative low completeness of this data set. The MabA$_{MSMEG}$–NADP$^+$ structure was solved by molecular replacement using the structure of MabA$_{Mtb}$ as a model (PDB entry 1uzn; Cohen-Gonsaud *et al.*, 2002). The apo MabA$_{MSMEG}$ and MabA$_{MSMEG}$–NADPH structures were both solved by molecular replacement using MabA$_{MSMEG}$–NADP$^+$ as a model. All structures were solved using *Phaser* (McCoy *et al.*, 2007) from the *PHENIX* package (Adams *et al.*, 2010). *Coot* (Emsley *et al.*, 2010) was used for manual rebuilding, while structure refinement and validation were performed with the *PHENIX* package (Adams *et al.*, 2010). The statistics for data collection and structure refinement are displayed in Table 1. Figures were prepared with *PyMOL* (http://www.pymol.org).

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

	Apo MabA$_{MSMEG}$	MabA$_{MSMEG}$–NADPH	MabA$_{MSMEG}$–NADP$^+$
Data-collection statistics			
Beamline	ID30B, ESRF	X06DA-PXIII, SLS	ID23-2, ESRF
Wavelength (Å)	0.979	1.0	0.872
Resolution range (Å)	34.1–1.70 (1.76–1.70)	39–1.45 (1.50–1.45)	47.2–2.40 (2.48–2.40)
Space group	$C121$	$P12_11$	$P2_12_12_1$
Unit-cell parameters			
a (Å)	96.8	59.1	52.1
b (Å)	68.2	69.3	69.3
c (Å	71.7	106.9	223.6
$\alpha = \gamma$ (°)	90	90	90
β (°)	123.7	91.8	90
Total reflections	213242 (20511)	923184 (96856)	214927 (21905)
Unique reflections	42269 (4171)	138568 (12845)	32703 (3240)
Multiplicity	5.0 (4.9)	6.7 (6.4)	6.6 (6.8)
Completeness (%)	98.7 (98.8)	89.1 (84.6)	99.9 (100.0)
Mean $I/\sigma(I)$	8.6 (1.3)	14.8 (1.05)	10.7 (3.5)
Wilson B factor (Å^2)	20.68	14.95	17.90
R_{meas}	0.132 (1.29)	0.097 (1.75)	0.166 (0.52)
CC$_{1/2}$	0.99 (0.57)	0.99 (0.68)	0.99 (0.86)
CC*	0.99 (0.85)	1.00 (0.90)	0.99 (0.96)
Refinement statistics			
Reflections used in refinement	42253 (4171)	136328 (12845)	32699 (3240)
Reflections used for R_{free}	2000 (198)	2007 (201)	2000 (198)
R_{work}	0.160 (0.300)	0.172 (0.442)	0.168 (0.220)
R_{free}	0.201 (0.310)	0.197 (0.444)	0.219 (0.310)
No. of non-H atoms			
Total	3929	7868	7518
Macromolecules	3520	6849	6845
Ligands	35	137	192
Solvent	374	882	481
Protein residues	478	928	934
R.m.s.d., bonds (Å)	0.006	0.009	0.002
R.m.s.d., angles (°)	0.74	1.02	0.38
Ramachandran favoured (%)	97.7	96.9	97.6
Ramachandran allowed (%)	2.1	3.1	2.4
Ramachandran outliers (%)	0.2	0.0	0.0
Rotamer outliers (%)	0.0	0.0	0.3
Average B factor (Å^2)			
Overall	27.5	21.9	21.4
Macromolecules	26.3	20.7	21.5
Ligands	62.4	27.8	20.8
Solvent	35.7	30.1	21.2
PDB code	5ovj	5ovk	5ovl

2.6. Synthesis of β-keto-octanoyl-CoA

The synthesis of β-keto-octanoyl-CoA was adapted from Kanchanabanca *et al.* (2013) and is described in the Supporting Information.

3. Results and discussion

3.1. Biochemical properties of MabA$_{MSMEG}$

Because FabG proteins can be found either as homodimers or homotetramers in solution, we determined the oligomeric state of MabA$_{MSMEG}$ by size-exclusion chromatography (Hou *et al.*, 2015; Marrakchi *et al.*, 2002). The elution peak at 12.9 ml for MabA$_{MSMEG}$ lacking the N-terminal tags corresponds to an apparent molecular weight of 87 kDa, which corresponds to a molecular weight between those of a homotrimer (theoretical molecular weight 79 kDa) and a homotetramer (theoretical molecular weight 106 kDa) (Fig. 1a). No change in

the oligomeric state was observed on decreasing the protein concentration (data not shown). Because we also report the crystal structure of MabA$_{MSMEG}$ with the N-terminal tags (in the NADP$^+$-bound structure), we also determined the oligomeric state of this protein in solution (Fig. 1b). This tagged form eluted at 12.6 ml, correlating with an estimated mass of 133 kDa, supporting a homotetrameric protein (the predicted molecular weight is 129 kDa). Taken collectively, these results are consistent with a homotetrameric form of MabA$_{MSMEG}$ in solution, which was further confirmed by analysis of the crystal packing of the different crystal forms obtained.

To check whether the purified protein was catalytically active, we next determined the steady-state kinetic properties of MabA$_{MSMEG}$. The sigmoidal nature of the plot of the velocity *versus* the NADPH concentration, while keeping the acetoacetyl-CoA (AcAcCoA) concentration constant, attests to the allosteric behaviour (Hill coefficient of 1.7 ± 0.25) of MabA$_{MSMEG}$ (Fig. 1c). In contrast, no significant cooperativity was observed when the concentration of NADPH was kept constant and the concentration of AcAcCoA (Fig. 1d) or β-keto-octanoyl-CoA (C$_8$-CoA; Fig. 1e) was increased. MabA$_{MSMEG}$ exhibits a K_m of 3.5 ± 0.42 mM and a k_{cat} of 1.14 s^{-1} for AcAcCoA. The enzyme has a higher affinity for C$_8$-CoA, with a K_m of 0.6 ± 0.12 mM and a faster catalytic cycle (k_{cat} = 5.72 s^{-1}). According to the values of the catalytic efficiency k_{cat}/K_m, MabA$_{MSMEG}$ has a preference for longer substrates, as it is 30 times more efficient at catalyzing the NADPH-dependent reduction of C$_8$-CoA than that of AcAcCoA. The catalytic efficiencies are in agreement with the previously reported kinetic constants of Mab$_{Mtb}$ (Marrakchi *et al.*, 2002). Moreover, it has also been shown that MabA$_{Mtb}$ presents allosteric behaviour (Silva *et al.*, 2008). Allosteric behaviour and positive cooperativity for NADPH have also been reported for the FabG enyzmes from *Staphylococcus aureus* and *V. cholerae* (Dutta *et al.*, 2012; Hou *et al.*, 2015).

3.2. Crystal structure of apo MabA$_{MSMEG}$

The apo MabA$_{MSMEG}$ structure was solved by molecular replacement as described in §2 and was refined to 1.7 Å resolution (Table 1). The asymmetric unit contained two monomers (Fig. 2a). Analysis of the crystal packing with the

PISA server (Krissinel & Henrick, 2007) predicts the formation of a stable homotetramer, corroborating the results of size-exclusion chromatography (Figs. 1*a* and 1*b*) and the quaternary structure of MabA$_{Mtb}$ (Cohen-Gonsaud *et al.*, 2002). All residues from Gly17 to His255 could be traced in both chains. The first 16 residues plus one glycine residue from the TEV protease cleavage site were not visible. Of note, an extra, but non-interpretable electron-density map in proximity to Arg107 in chain *A* (data not shown) could be seen but did not correspond to any component of the crystallization condition. Additionally, seven sulfate molecules were placed. In each monomer, one sulfate ion is sequestered by the Arg33, Arg55 and Ser57 side chains and occupies a position similar to that of the NADPH 2′-phosphate group in the adenosine ribose binding site (Fig. 2*b*). A similar occupancy of sulfate was described in the crystal structure of the apo form of the *Plasmodium falciparum* FabG protein (Wickramasinghe *et al.*, 2006).

The overall structure adopts a Rossmann fold, with a central β-sheet made up of seven parallel β-strands surrounded by α-helices, as typically seen in the SDR superfamily (Kavanagh *et al.*, 2008). The region comprising residues 195–217 between strand β6 and helix α6, which is named the flexible region and is disordered in most apo FabG structures, could be traced (Fig. 2*c*). Additionally, residues 98–107 and 150–157 near the catalytic triad that are absent in the apo MabA$_{Mtb}$ structure (PDB entry 1uzm; Cohen-Gonsaud *et al.*, 2002) could be modelled. Since MabA$_{MSMEG}$ and MabA$_{Mtb}$ share 83% primary-sequence identity, it is not surprising that the apo MabA$_{MSMEG}$ and apo MabA$_{Mtb}$ structures are very similar, with a root-mean-square deviation (r.m.s.d.) of 0.92 Å over 212 C$^\alpha$ residues. Nonetheless, differences were noticed in the region 195–217 (187–209 in MabA$_{Mtb}$) located between strand β6 and helix α6, with the residues adopting different secondary structures (Fig. 2*d*). In the apo MabA$_{MSMEG}$ structure residues 195–197 adopt a 3$_{10}$-helix followed by a coil,
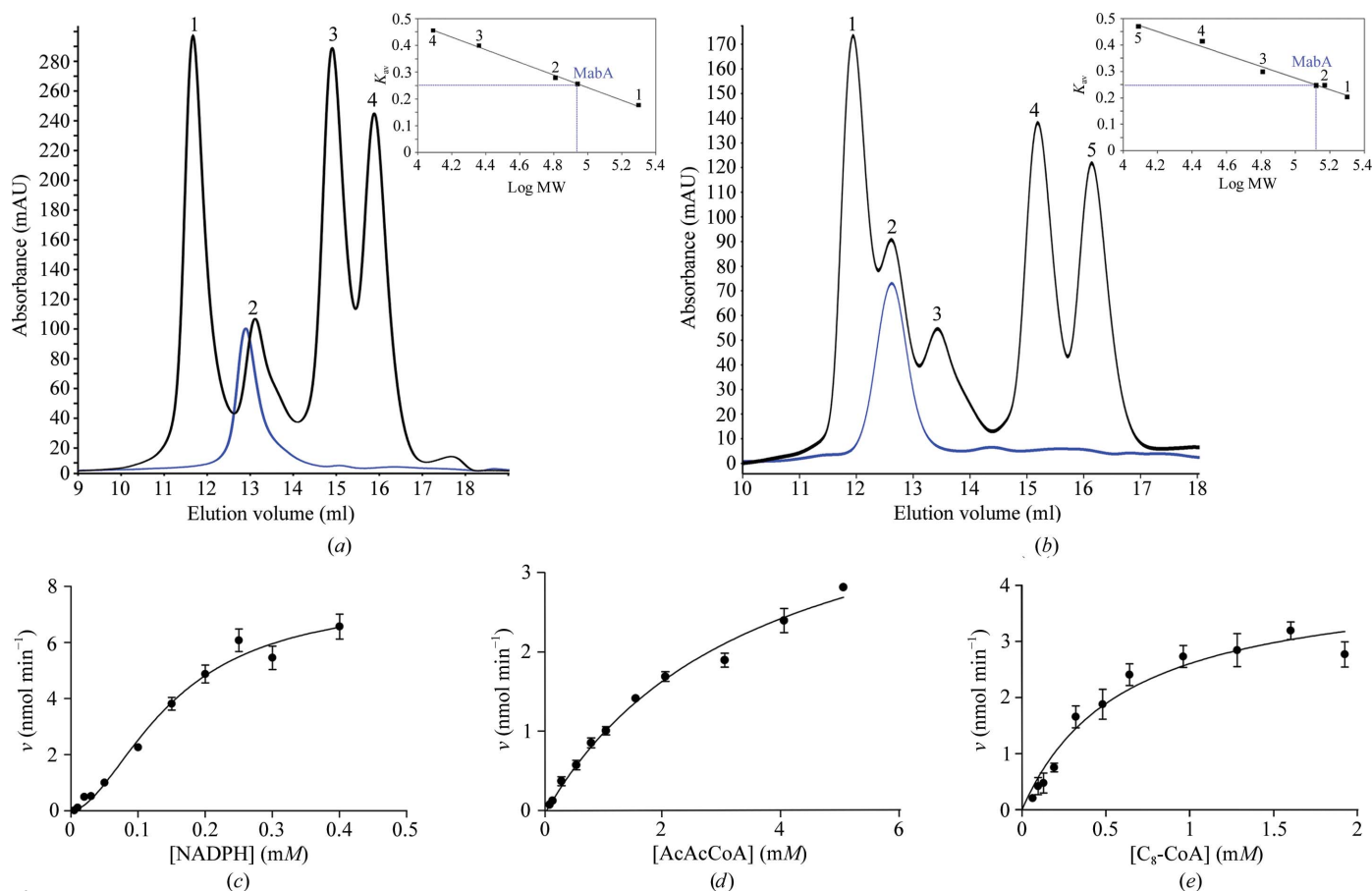

Figure 1
Biochemical characterization of MabA$_{MSMEG}$. (*a*) Oligomeric state of MabA$_{MSMEG}$ with tags removed determined by size-exclusion chromatography. The elution profile of the proteins used for calibration is displayed as a black line and the elution profile of MabA$_{MSMEG}$ is shown in blue. Calibration was established using β-amylase (1; 200 kDa), bovine serum albumin (2; 66 kDa), carbonic anhydrase (3; 29 kDa) and cytochrome *c* (4; 12.4 kDa), which eluted with estimated volumes of 11.6, 13.2, 15 and 15.8 ml, respectively. The void volume was estimated at 8.8 ml using dextran blue. MabA$_{MSMEG}$ was concentrated to 5 mg ml^{-1}. The MabA$_{MSMEG}$ elution peak at 12.90 ml corresponds to a calculated molecular weight of 87 kDa. (*b*) Oligomeric state of MabA$_{MSMEG}$ with its N-terminal tags determined by size-exclusion chromatography. The elution profile of the proteins used for calibration is displayed as a black line and the elution profile of MabA$_{MSMEG}$ is shown in blue. Calibration was established using β-amylase (1; 200 kDa), alcohol dehydrogenase (2; 150 kDa), bovine serum albumin (3; 66 kDa), carbonic anhydrase (4; 29 kDa) and cytochrome *c* (5; 12.4 kDa), which eluted with estimated volumes of 11.9, 12.6, 13.4, 15.2 and 16.1 ml, respectively. The void volume was estimated at 8.9 ml using dextran blue. The MabA$_{MSMEG}$ elution peak at 12.6 ml corresponds to a calculated molecular weight of 133 kDa. Steady-state kinetic properties of MabA$_{MSMEG}$ for (*c*) NADPH (with a fixed concentration of AcAcCoA), (*d*) acetoacetyl-CoA and (*e*) β-keto-octanoyl-CoA. Error bars indicate the standard error.

while residues 200–208 form an α-helix that is then followed by another 3_{10}-helix (residues 209–213). However, in the apo MabA$_{Mtb}$ structure two α-helices comprising residues 187–193 and 196–205 are formed.

Surprisingly, no shift of the catalytic triad was observed, in contrast to that previously seen in the apo form of MabA$_{Mtb}$ (PDB entry 1uzm). The catalytic triad Ser148, Lys165 and Tyr161 in MabA$_{MSMEG}$ is already positioned in an active conformation, as found in the FabG–NADP$^+$ structure from *E. coli* (Price *et al.*, 2004; Fig. 2e). Such an active-site conformation in the absence of any cofactor is not unique and has already been reported in the apo structures of FabG from *V. cholerae* (PDB entry 3rro; Hou *et al.*, 2015) and the FabG-like protein MSMEG_6753 from *M. smegmatis* (PDB entry 5t2v; Blaise *et al.*, 2017).

As there is no complete structure of MabA$_{Mtb}$ bound to cofactor available and since the cofactor could only be

partially modelled in the MabA$_{Mtb}$–NADP$^+$ complex (Cohen-Gonsaud *et al.*, 2005), we also crystallized both NADPH-bound and NADP$^+$-bound forms of MabA$_{MSMEG}$. This not only helped to identify all of the residues involved in cofactor recognition, but also provided new insights into the dynamics of the flexible 195–217 region.

3.3. Crystal structures of MabA$_{MSMEG}$ complexed with NADP$^+$ and NADPH

The crystal structure of the MabA$_{MSMEG}$–NADP$^+$ complex could be solved and refined to 2.4 Å resolution (Table 1). The asymmetric unit contained the biological tetramer of MabA$_{MSMEG}$ (Fig. 3a). Electron density for the four NADP$^+$ cofactors is well defined in all the monomers and the cofactors could therefore be entirely modelled (Fig. 3b). Rebuilding was near to completion only for chain A, where we were able to
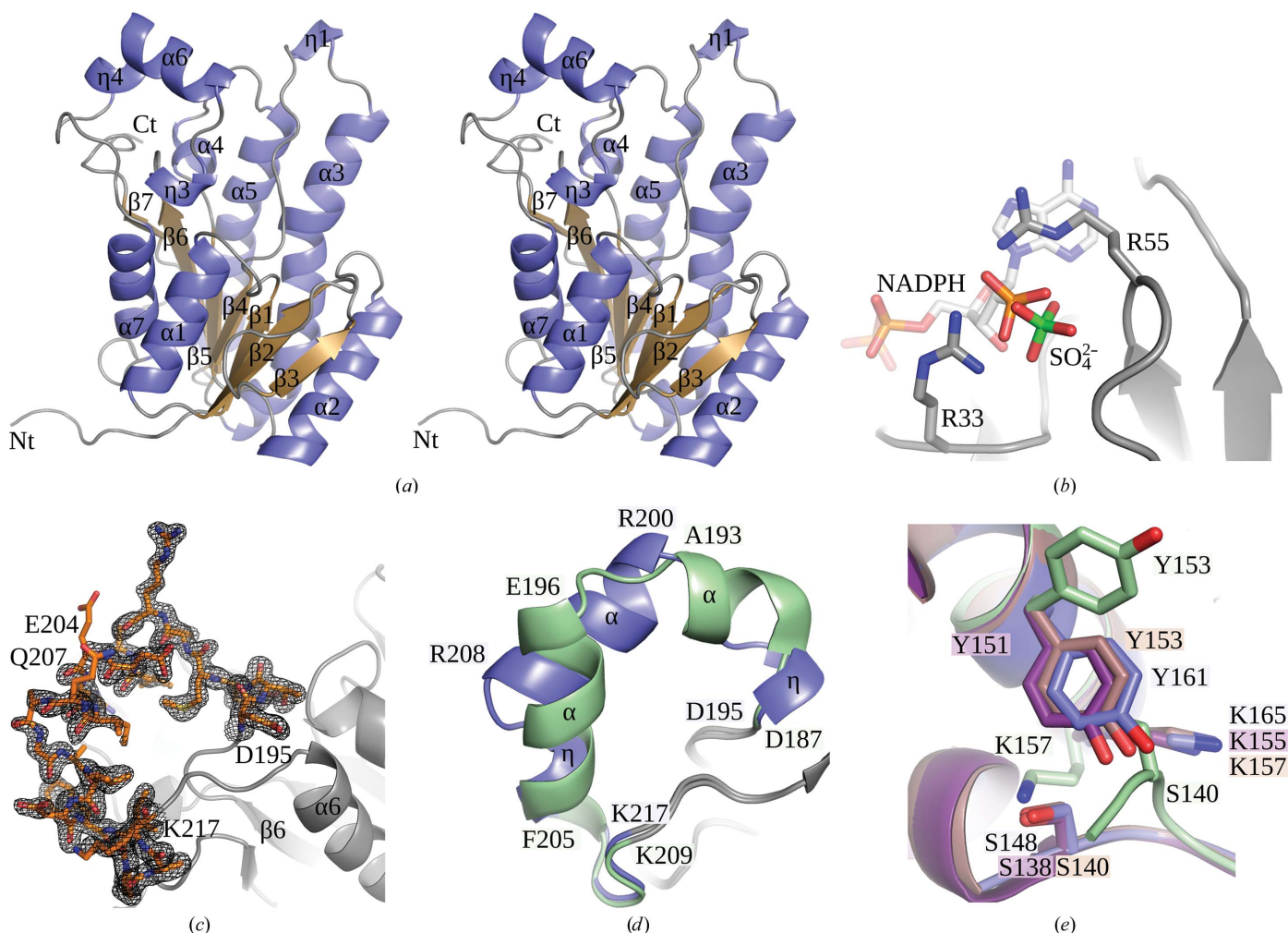

Figure 2
Crystal structure of the apo form of MabA$_{MSMEG}$. (a) Overall structure and stereoscopic view of apo MabA$_{MSMEG}$. Helices are represented as blue cartoons, whereas strands and coils are shown in sand and grey, respectively. α, β and η followed by a number indicate α-helices, β-strands and 3_{10}-helices, respectively. (b) Superposition of apo MabA$_{MSMEG}$ and NADPH-bound MabA$_{MSMEG}$ showing a sulfate ion in a position that is comparable to that of the phosphate group of NADPH in the MabA$_{MSMEG}$–NADPH crystal structure. (c) $2mF_o - DF_c$ electron-density map contoured at the 1σ level corresponding to residues Asp195–Lys217 represented as orange sticks. All residues are well defined in the electron density, with the exception of the side chains of Glu204 and Gln207. (d) Comparison of the flexible region in the apo forms of MabA$_{MSMEG}$ (slate blue) and MabA$_{Mtb}$ (pale green; PDB entry 1uzm). (e) Positions of the catalytic residues found in apo MabA$_{MSMEG}$ (slate blue), apo MabA$_{Mtb}$ (pale green), MabA$_{Mtb}$–NADP$^+$ (brown) and *E. coli* FabG–NADP$^+$ (purple; PDB entry 1q7b).

model residues 15–255 (Fig. 3*c*), while in chains *B*, *C* and *D* residues 197–208, 197–202 and 197–209, respectively, could not be modelled. Importantly, all of the amino acids involved in cofactor binding are seen in chain *A*, allowing a detailed interaction map with NADP[+] to be built (Fig. 3*d*). NADP[+] is tightly bound by its nicotinamide group to the side chain of Arg205 and the main chains of Gly192 and Ile194, while Pro191 interacts by hydrophobic interaction. The side chains
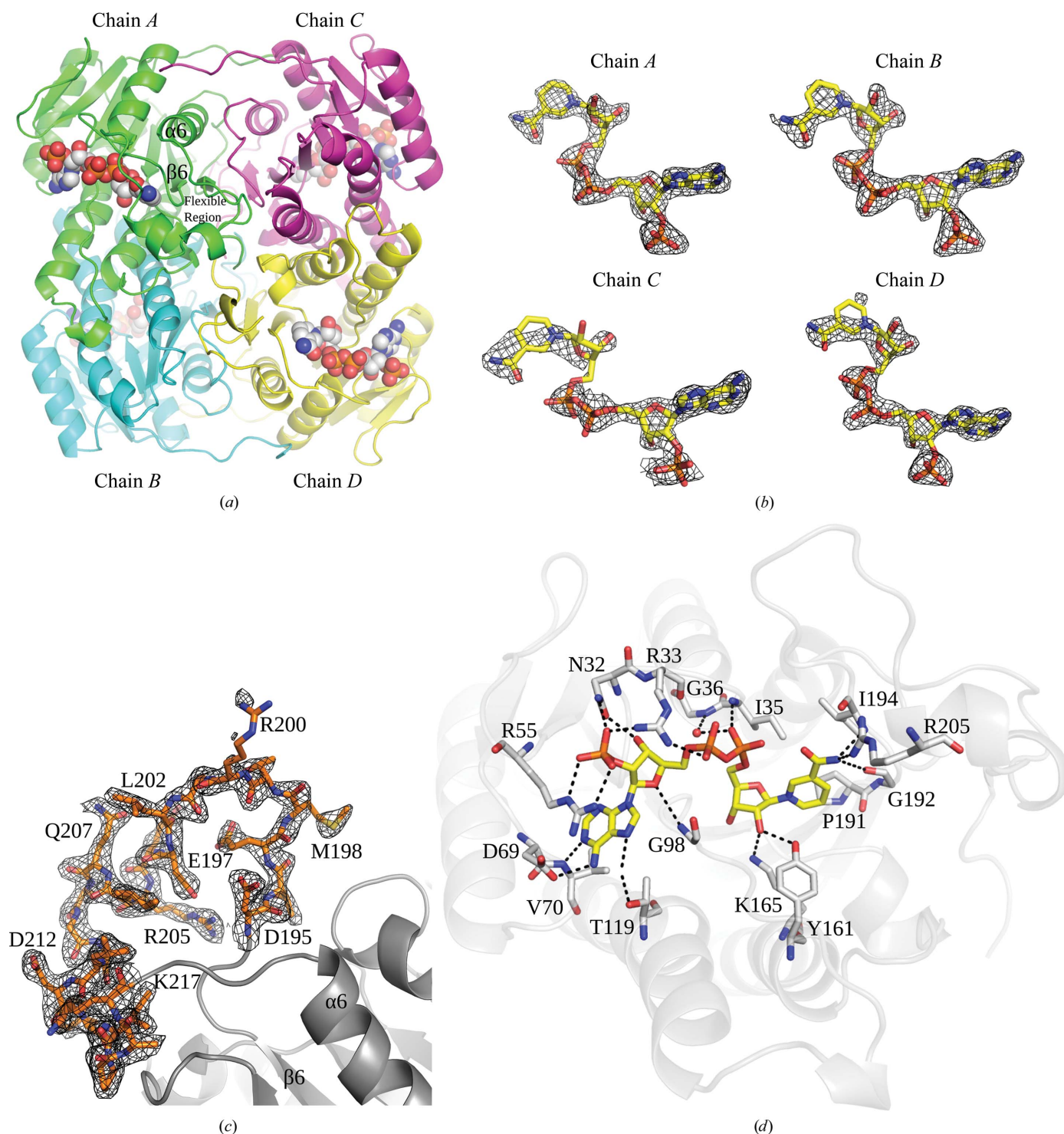

Figure 3
Crystal structure of MabA$_{MSMEG}$ bound to NADP[+] (*a*) The overall structure is represented as a cartoon, the location of the flexible loop between helix α6 and strand β6 is indicated and the NADP[+] molecules are represented as spheres. (*b*) $F_o - F_c$ simulated-annealing OMIT map contoured at the 3.5σ level of the NADP[+] molecules in each subunit, which are shown as sticks. (*c*) $2mF_o - DF_c$ electron density of residues 195–217 displayed as sticks; only the side chain of Arg200 is not well defined. (*d*) NADP[+]-binding site. NADP[+] is shown as yellow sticks and all of the residues involved in interactions are displayed as white sticks. Water molecules are represented by red spheres and dashed lines indicate hydrogen bonds.

of Tyr161 and Lys165 are in contact with O2 of the ribose. The pyrophosphate group of NADP$^+$ binds the main chains of Ile35 and Gly36 as well as the side chain of Arg33. The ribose of the adenosine interacts with the main chain of Gly98 and the side chain of Asn32. The latter also contacts the phosphate group of adenosine. The adenine group is bound by the side chain of Thr119 via a water molecule, and also by the side chain of Asp69 and the main chain of Val70. The phosphate group of adenosine is also involved in salt bridges to the guanidinium groups of Arg33 and Arg55.

Moreover, we also solved the co-crystal structure of MabA$_{MSMEG}$ bound to NADPH (Table 1). The asymmetric unit consists of four monomers that all enclose an NADPH molecule. The monomers present disordered regions. Only one monomer, chain A, is almost complete: residues 14–255

could be traced including the flexible region 195–217 (Fig. 4a). In chains B, C and D, in addition to the N-terminal part that is also missing in chain A, residues 197–209, 196–210 and 199–210, respectively, from the flexible region were not visible in the electron-density map. Furthermore, although it is clear from the simulated-annealing OMIT map that all of the subunits are occupied by NADPH, only the NADPH in chain B could be entirely traced (Fig. 4b). As chain A was almost completely rebuilt but with partial NADPH occupancy, and as full NADPH occupancy was observed in chain B, we could map the NADPH-binding site by reporting the interactions that occur in these two subunits (Figs. 4c and 4d). As a similar positioning of NADPH and NADP$^+$ is observed, the residues contacting NADP$^+$ also contact NADPH. Additional residues involved in cofactor binding could be identified, notably
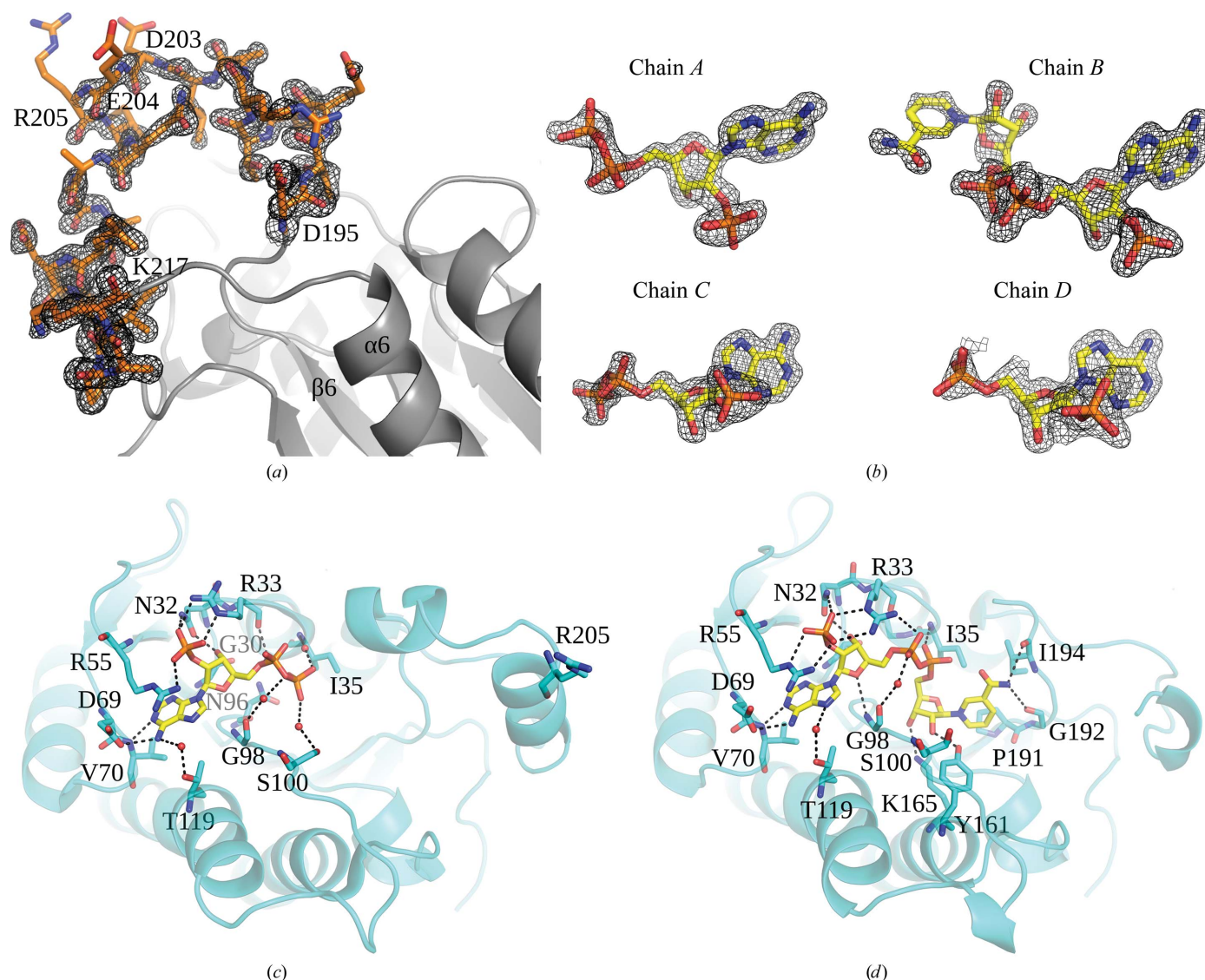

Figure 4
Crystal structure of MabA$_{MSMEG}$ bound to NADPH. (a) The $2mF_o - DF_c$ electron density (1σ level) is well defined for the main and side chains of most residues, with the exception of Asp203, Glu204 and Arg205, the side chains of which were only visible at a lower contour. (b) $F_o - F_c$ simulated-annealing OMIT map contoured at the 3.5σ level of the NADPH molecules in each subunit, which are shown as sticks. Only NADPH from chain B could be rebuilt entirely. (c) NADPH-binding site in chain A. (d) NADPH-binding site in chain B. All residues surrounding the NADPH-binding site are shown as sticks. Water molecules are represented by red spheres and dashed lines indicate hydrogen bonds.

Ser100 that contacts the ribose. Despite the relatively well conserved interaction network in the two cofactor-bound structures, a significant structural change was discerned. The

side chain of Arg205 that was involved in recognition of the nicotinamide group of NADP⁺ appears too far away to contact the NADPH cofactor (Figs. 4c and 4d). This structural rearrangement is triggered by a change in the positioning of the flexible region.

3.4. Cofactor binding-induced structural rearrangements

The structures of the apo, NADPH-bound and NADP⁺-bound forms of MabA$_{MSMEG}$ are similar overall. Superposition of the NADP⁺-bound and NADPH-bound structures with the apo form generates an r.m.s.d. of 0.64 Å over 232 and 229 C$^\alpha$ atoms, respectively, while the MabA$_{MSMEG}$–NADPH and MabA$_{MSMEG}$–NADP⁺ structures are slightly more distant, with an r.m.s.d. of 1.05 Å over 234 C$^\alpha$ atoms. The main structural difference remains in the different conformations of residues 195–215. The holo MabA$_{MSMEG}$ structures appear in a more open state than the apo-form structure. These different conformations are mainly triggered by residues 195–215, which display large movements and undergo a secondary-structure reorganization (Fig. 5a). In the MabA$_{Mtb}$ structure, NADP⁺ also seems to induce a structural change since the flexible region is more ordered in the NADP⁺-bound form than in the apo form (Cohen-Gonsaud et al., 2002, 2005).

To address whether and how these structural changes influence substrate binding, a C$_{16}$ fatty-acid molecule was modelled into the MabA$_{MSMEG}$ active site by superposition of our structures with the crystal structure of InhA bound to *trans*-2-hexadecenoyl-(*N*-acetylcysteamine)-thioester and NADH (PDB entry 1bvr; Rozwarski et al., 1999). The proper positioning of the catalytic triad in apo MabA$_{MSMEG}$ (Fig. 2e) is however not compatible with binding of the cofactor because of
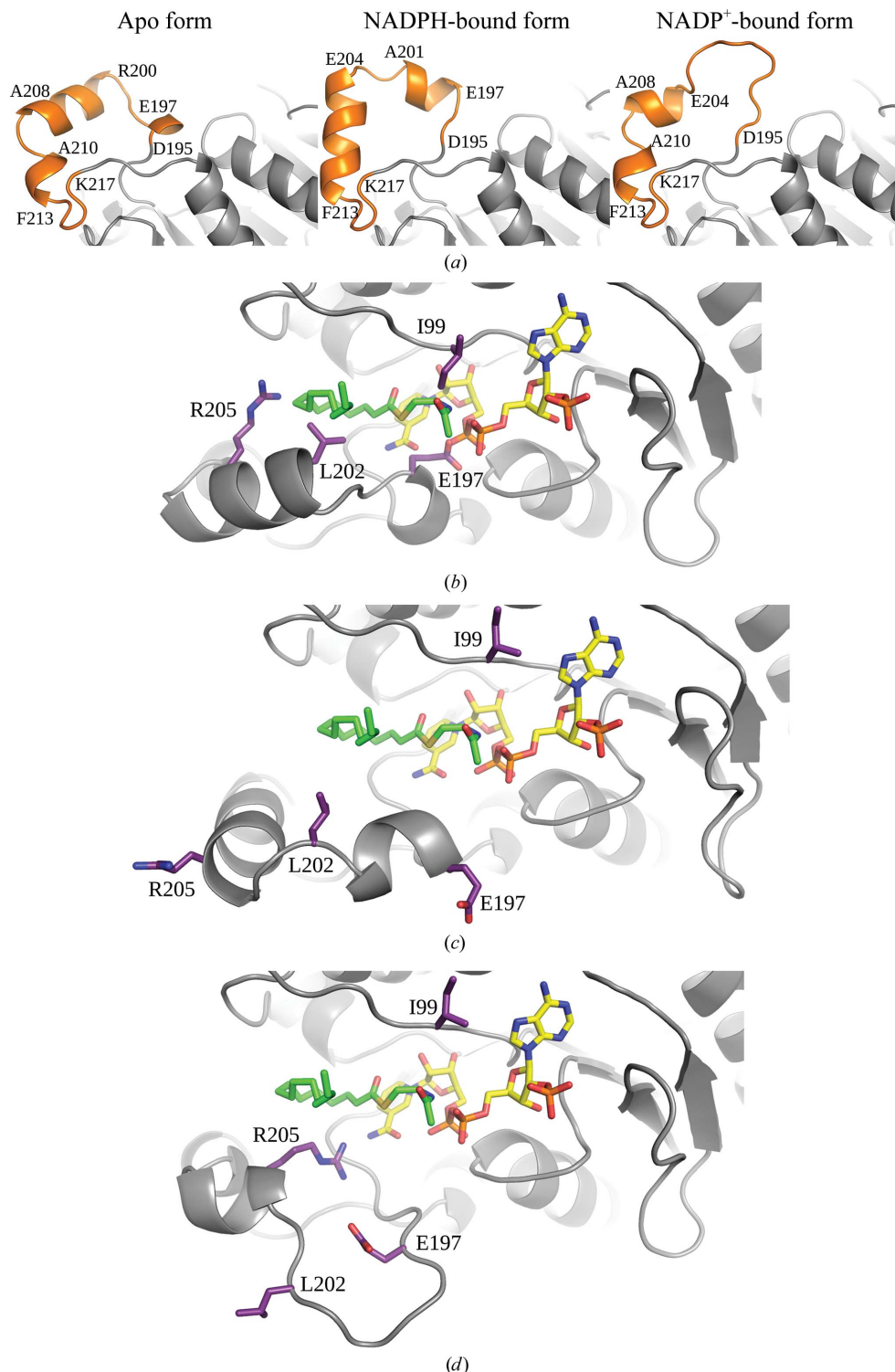
Figure 5
Dynamics of residues 195–217 upon the binding of cofactor. (a) Residues 195–217 coloured orange adopt different conformations in the apo form (left panel), the NADPH-bound form (middle panel) and the NADP⁺-bound form (right panel). (b), (c) and (d) represent docking models of the C$_{16}$ fatty acid in the active site of apo, NADPH-bound and NADP⁺-bound forms of MabA$_{MSMEG}$, respectively. Key residues for substrate/cofactor binding are represented as purple sticks.

the presence of the Glu197 side chain that provides steric hindrance of the pyrophosphate group of NADPH. Docking with the fatty acid shows that numerous residues in the apo form of MabA$_{MSMEG}$ would trigger clashes with the lipid. The side chains of Arg205, Leu202 and Ile99 substantially block the lipid-binding pocket (Fig. 5b). In contrast, in the NADPH-bound structure, both the lipid- and the NADPH-binding cavities are accessible owing to the shift of the side chain of Ile99. Since the side chains of Arg205 and Leu202 are pushed away from the lipid cavity and because the side chain of Glu197 no longer interferes with the pyrophosphate group of NADPH, accommodation of the cofactor in its binding pocket is possible (Fig. 5c). In this conformation, residues 197–201 and 204–213 form two α-helices (Fig. 5a, middle panel), whereas they form 3_{10}-helices and one shorter α-helix in the apo form (Fig. 5a, left panel).

The active site of the NADP$^+$-bound form appears much more accessible as some of the residues move more than 10 Å away from their positions in the apo form. The flexible region adopts different secondary structures compared with the structures of the apo or NADPH-bound forms. Residues 195–

203 now form a loop, while residues 204–208 form an α-helix that appears to be much shorter than in the NADPH-bound and apo forms (Fig. 5a, right panel). This helix is moved towards the substrate-binding site, which positions the Arg205 side chain below the lipid and allows its interaction with the nicotinamide group of the cofactor (Fig. 5d).

Comparison of these three structures allows sequential structural rearrangements to be proposed, in which the apo MabA$_{MSMEG}$ would initially be in a closed state. The flexible loop would then open to bind NADPH and the lipid (a β-ketoacyl-ACP) and after catalysis adopts an even more open state to release the product (a β-hydroxyacyl-ACP) as well as the oxidized cofactor. This scenario is in agreement with the kinetic analysis performed on MabA$_{Mtb}$, where the authors proposed a sequential binding order for the cofactor and the substrate (Silva *et al.*, 2008). The important changes observed between the NADPH-bound and NADP$^+$-bound structures are surprising as the two molecules only differ by one H atom, but are consistent with other FabG crystal structures. In the *S. aureus* FabG structure bound to NADPH (PDB entry 3sj7; Dutta *et al.*, 2012) the position of the flexible region is similar to our observation, while in the *Listeria monocytogenes* FabG structure bound to NADP$^+$ (PDB entry 4jro; Center for Structural Genomics of Infectious Diseases, unpublished work) this region displays a more open conformation (Figs. 6a and 6b). Nevertheless, the movements observed when comparing the three structures of MabA$_{MSMEG}$ seem to differ from those of other FabG proteins, as described for instance for the *P. aeruginosa* FabG structures (Cukier *et al.*, 2013). Indeed, NADPH binding induced a transition from a low-ordered to a higher ordered loop and an inward movement of this loop, albeit moderate, towards the cofactor-binding site (Fig. 6c). In MabA$_{MSMEG}$ binding of the cofactors leads to a reordering of the flexible loop with secondary-structure rearrangements, but in contrast to *P. aeruginosa* FabG an outward movement of the loop is observed (Fig. 6d). The reason for these mechanistic differences may be explained by the fact that MabA can bind to much larger substrates, namely the mero-mycolic acids, while FabG from other bacterial species binds to much smaller acyl chains. The catalytic domain of MabA may
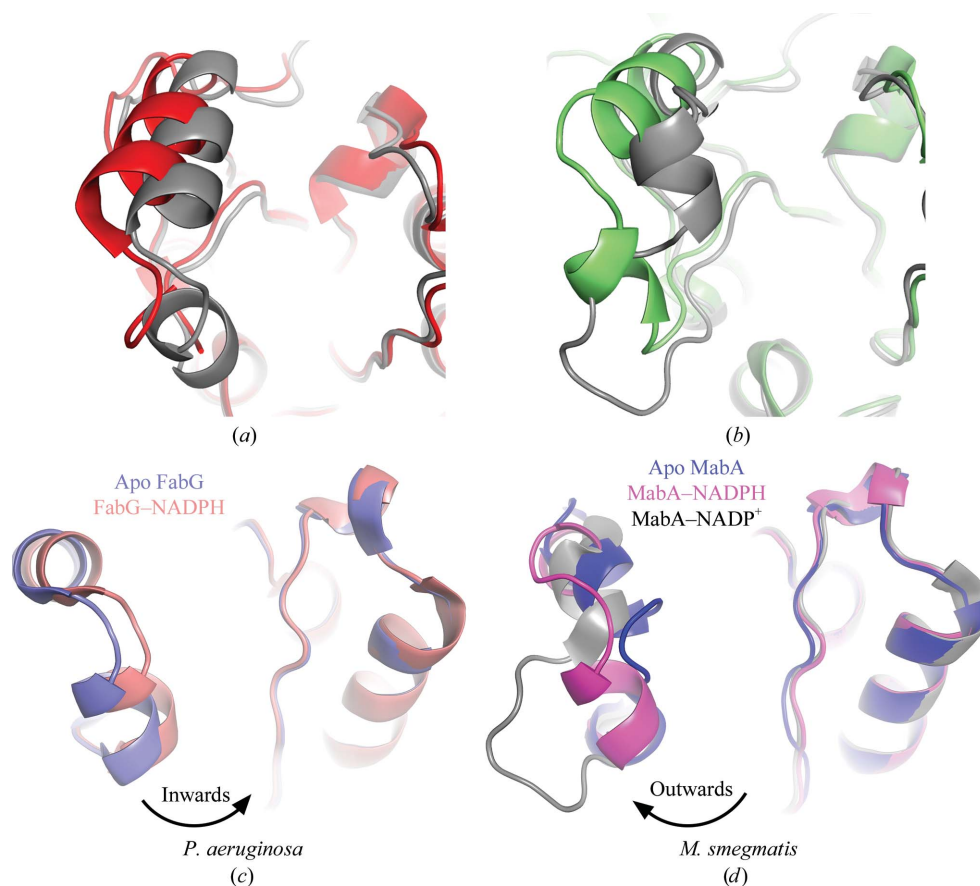

Figure 6
Comparison of the crystal structures of MabA$_{MSMEG}$ with those of other FabG proteins. (a) Superposition of the flexible regions of the *S. aureus* FabG–NADPH complex (red; PDB entry 3sj7) and the MabA$_{MSMEG}$–NADPH structure (grey). (b) The open form of the *L. monocytogenes* FabG–NADP$^+$ complex (green; PDB entry 4jro) appears to be similar to that of MabA$_{MSMEG}$–NADP$^+$ (grey). (c, d) Comparison of the cofactor-induced movement in *P. aeruginosa* FabG and MabA$_{MSMEG}$. The arrows highlight the orientations of the movements.

require a more open form than conventional FabG proteins to accommodate/release its very bulky substrates and products.

4. Conclusion

Potent inhibitors of the *P. aeruginosa* FabG enzyme have recently been identified, emphasizing the interest in developing specific inhibitors of this essential class of enzymes (Cukier *et al.*, 2013). Despite being excellent antimicrobial targets owing to their crucial role in mycolic acid biosynthesis, to the best of our knowledge no specific inhibitors have yet been reported for mycobacterial MabA proteins. In this study, we provide a sequential and dynamic mechanism that comprises important structural rearrangements that occur during catalysis by MabA. This dynamic primarily involves a continuum of changes within a flexible region ranging from a closed to an open conformation. Access to these structural details will be very informative for subsequent *in silico* drug-screening approaches and structure–activity relationship studies, as well as for the design of optimal and effective inhibitors of *Mtb* and other pathogenic mycobacteria.

References

Abrahams, K. A. *et al.* (2016). *Nature Commun.* **7**, 12581.
Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K. S., Wilson, T., Collins, D., de Lisle, G. & Jacobs, W. R. (1994). *Science*, **263**, 227–230.
Banerjee, A., Sugantino, M., Sacchettini, J. C. & Jacobs, W. R. (1998). *Microbiology*, **144**, 2697–2704.
Bhatt, A., Fujiwara, N., Bhatt, K., Gurcha, S. S., Kremer, L., Chen, B., Chan, J., Porcelli, S. A., Kobayashi, K., Besra, G. S. & Jacobs, W. R. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 5157–5162.
Bhatt, A., Kremer, L., Dai, A. Z., Sacchettini, J. C. & Jacobs, W. R. (2005). *J. Bacteriol.* **187**, 7596–7606.
Bhatt, A., Molle, V., Besra, G. S., Jacobs, W. R. & Kremer, L. (2007). *Mol. Microbiol.* **64**, 1442–1454.
Blaise, M., Van Wyk, N., Banères-Roquet, F., Guérardel, Y. & Kremer, L. (2017). *Biochem. J.* **474**, 907–921.
Boehringer, D., Ban, N. & Leibundgut, M. (2013). *J. Mol. Biol.* **425**, 841–849.
Cohen-Gonsaud, M., Ducasse, S., Hoh, F., Zerbib, D., Labesse, G. & Quemard, A. (2002). *J. Mol. Biol.* **320**, 249–261.
Cohen-Gonsaud, M., Ducasse-Cabanot, S., Quemard, A. & Labesse, G. (2005). *Proteins*, **60**, 392–400.
Coxon, G. D., Craig, D., Corrales, R. M., Vialla, E., Gannoun-Zaki, L. & Kremer, L. (2013). *PLoS One*, **8**, e53162.
Cukier, C. D., Hope, A. G., Elamin, A. A., Moynie, L., Schnell, R., Schach, S., Kneuper, H., Singh, M., Naismith, J. H., Lindqvist, Y., Gray, D. W. & Schneider, G. (2013). *ACS Chem. Biol.* **8**, 2518–2527.
Dutta, D., Bhattacharyya, S. & Das, A. K. (2012). *Proteins*, **80**, 1250–1257.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Gannoun-Zaki, L., Alibaud, L. & Kremer, L. (2013). *Antimicrob. Agents Chemother.* **57**, 629–632.
Grzegorzewicz, A. E. *et al.* (2012). *J. Biol. Chem.* **287**, 38434–38441.
Grzegorzewicz, A. E., Eynard, N., Quémard, A., North, E. J., Margolis, A., Lindenberger, J. J., Jones, V., Korduláková, J., Brennan, P. J., Lee, R. E., Ronning, D. R., McNeil, M. R. & Jackson, M. (2015). *ACS Infect. Dis.* **1**, 91–97.
Hou, J., Zheng, H., Chruszcz, M., Zimmerman, M. D., Shumilin, I. A., Osinski, T., Demas, M., Grimshaw, S. & Minor, W. (2015). *J. Bacteriol.* **198**, 463–476.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 133–144.
Kanchanabanca, C., Tao, W., Hong, H., Liu, Y., Hahn, F., Samborskyy, M., Deng, Z., Sun, Y. & Leadlay, P. F. (2013). *Angew. Chem. Int. Ed.* **52**, 5785–5788.
Kapilashrami, K., Bommineni, G. R., Machutta, C. A., Kim, P., Lai, C.-T., Simmerling, C., Picart, F. & Tonge, P. J. (2013). *J. Biol. Chem.* **288**, 6045–6052.
Kavanagh, K. L., Jörnvall, H., Persson, B. & Oppermann, U. (2008). *Cell. Mol. Life Sci.* **65**, 3895–3906.
Kremer, L., Douglas, J. D., Baulard, A. R., Morehouse, C., Guy, M. R., Alland, D., Dover, L. G., Lakey, J. H., Jacobs, W. R., Brennan, P. J., Minnikin, D. E. & Besra, G. S. (2000). *J. Biol. Chem.* **275**, 16857–16864.
Kremer, L., Dover, L. G., Carrère, S., Nampoothiri, K. M., Lesjean, S., Brown, A. K., Brennan, P. J., Minnikin, D. E., Locht, C. & Besra, G. S. (2002). *Biochem. J.* **364**, 423–430.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Marrakchi, H., Ducasse, S., Labesse, G., Montrozier, H., Margeat, E., Emorine, L., Charpentier, X., Daffé, M. & Quémard, A. (2002). *Microbiology*, **148**, 951–960.
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.
Pawełczyk, J. & Kremer, L. (2014). *Microbiol. Spectr.* **2**, MGM2-0003-2013.
Price, A. C., Zhang, Y.-M., Rock, C. O. & White, S. W. (2004). *Structure*, **12**, 417–428.
Quémard, A. (2016). *Trends Microbiol.* **24**, 725–738.
Quémard, A., Sacchettini, J. C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W. R. & Blanchard, J. S. (1995). *Biochemistry*, **34**, 8235–8241.
Rozwarski, D. A., Vilchèze, C., Sugantino, M., Bittman, R. & Sacchettini, J. C. (1999). *J. Biol. Chem.* **274**, 15582–15589.
Sacco, E., Covarrubias, A. S., O'Hare, H. M., Carroll, P., Eynard, N., Jones, T. A., Parish, T., Daffé, M., Bäckbro, K. & Quémard, A. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 14628–14633.
Silva, R. G., Rosado, L. A., Santos, D. S. & Basso, L. A. (2008). *Arch. Biochem. Biophys.* **471**, 1–10.
World Health Organization (2016). *Global Tuberculosis Report 2016*. Geneva: World Health Organization.
Wickramasinghe, S. R., Inglis, K. A., Urch, J. E., Müller, S., van Aalten, D. M. F. & Fairlamb, A. H. (2006). *Biochem. J.* **393**, 447–457.