Structure of naphthoate synthase (MenB) from *Mycobacterium tuberculosis* in both native and product-bound forms

Jodie M. Johnston,[a] Vickery L. Arcus[a] and Edward N. Baker[a,b]*

[a]School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand, and [b]Department of Chemistry, University of Auckland, Private Bag 92019, Auckland, New Zealand

Correspondence e-mail: ted.baker@auckland.ac.nz

Mycobacterium tuberculosis, the cause of tuberculosis, is one of the most devastating human pathogens. New drugs for its control are urgently needed. Menaquinone, also known as vitamin K, is an essential cofactor that is required for electron transfer and the enzymes that synthesize it are therefore potential drug targets. The enzyme naphthoate synthase (MenB) from *M. tuberculosis* has been expressed in *Escherichia coli*, purified and crystallized both as the native enzyme and in complex with naphthoyl-CoA. Both structures have been determined by X-ray crystallography: native MenB at 2.15 Å resolution ($R = 0.203$, $R_{free} = 0.231$) and its napthoyl-CoA complex at 2.30 Å resolution ($R = 0.197$, $R_{free} = 0.225$). The protein structure, which has a fold characteristic of the crotonase family of enzymes, is notable for the presence of several highly flexible regions around the active site. The bound naphthoyl-CoA is only visible for one of the three molecules in the asymmetric unit and only partly rigidifies the structure. The C-terminal region of the protein is seen to play a critical role both in completion of the binding pocket and in stabilization of the hexamer, suggesting a link between oligomerization and catalytic activity.

1. Introduction

The availability of complete genome sequences for a large number of microorganisms offers unparalleled opportunities both to understand their biology and, in the case of pathogenic organisms, to develop new therapeutic agents to combat disease. *Mycobacterium tuberculosis* (*Mtb*), the cause of tuberculosis (TB), is one of the most serious human pathogens, being responsible for around two million deaths every year (Stokstad, 2000). Although drugs effective against replicating *Mtb* are available, a number of factors limit their usefulness. Among these are the long treatment times necessary (typically six months of multi-drug therapy; McKinney, 2000; Barry *et al.*, 2000; Duncan, 2003), with consequent compliance problems, the growing problem of multi-drug resistance (Barry *et al.*, 2000; Duncan, 2003), the synergy with HIV-AIDS (Barry *et al.*, 2000) and the ability of *Mtb* to enter a dormant or persistent state inside human tissues and emerge many years later as active infections (Parrish *et al.*, 1998; Wayne & Sohaskey, 2001).

The complete genome sequence for *Mtb* strain H37Rv was published in 1998 (Cole *et al.*, 1998), offering new possibilities for the development of anti-TB drugs. Among other initiatives, an international consortium of laboratories known as the International TB Structural Genomics Consortium (http://www.doe-mbi.ucla.edu/TB/) was set up to develop large-scale structure determination of *Mtb* proteins (Terwilliger *et al.*, 2003; Goulding *et al.*, 2003; Smith & Sacchettini, 2003). In this

context, we have targeted proteins involved in the biosynthesis of menaquinone (vitamin K2), an essential quinone used in electron-transfer pathways in the bacterium. The facts that menaquinone appears to be the only quinone used in *Mtb* (Collins & Jones, 1981) and that it is important for both aerobic and anaerobic processes (Meganathan, 2001) (the latter being relevant to the persistent state of *Mtb*; Wayne & Sohaskey, 2001), suggests that the enzymes of this pathway could be effective targets for anti-TB drug development; no such pathway exists in humans.

Menaquinone comprises an aromatic napthoquinone group with an attached polyisoprenyl substituent. Its biosynthetic pathway in bacteria typically comprises 6–7 steps (Meganathan, 2001), with genes for most of the enzymes responsible for these steps being located within a short section of the *Mtb* genome, between Rv0542c (*menE*) and Rv0558. The latter was originally annotated as *ubiE* and then changed to *menH*, but is more correctly annotated as *menG*, the methyltransferase that catalyses the last biosynthetic step (Johnston *et al.*, 2003). Here, we focus on the product of the open reading frame Rv0548c: MenB or naphthoate synthase (EC 4.1.3.36). This enzyme catalyses the closure of the naphthoyl ring (Fig. 1), with its substrate being the CoA ester of *O*-succinyl benzoic acid and its product the CoA ester of 1,4-dihydroxy-2-naphthoic acid (DHNA; Meganathan, 2001; Truglio *et al.*, 2003). The CoA moiety is expected to be cleaved subsequently by another enzyme, MenH (Meganathan, 2001; Palmer *et al.*, 1999), but no equivalent of this enzyme has yet been discovered in *Mtb*.

MenB is known to be a member of the crotonase super-family, although it has only 25–28% sequence identity at most with other members of this family. The various crotonase-family members are usually trimeric or hexameric and share a common fold with similarly located active sites, but catalyse diverse reactions and vary in the identity of some of their essential catalytic residues (Holden *et al.*, 2001; Gerlt & Babbitt, 2001). What their enzymatic reactions have in common is the binding of a CoA-linked substrate and the stabilization of an enolate anion involving a thioester carbonyl group through a structural feature of the protein referred to as an 'oxyanion hole' (Holden *et al.*, 2001; Gerlt & Babbitt, 2001; Xiang *et al.*, 1999). This common chemistry can often be recognized from two consensus sequences for the regions surrounding the oxyanion-hole residues (Holden *et al.*, 2001).

Here, we present the crystal structure of MenB from *M. tuberculosis* both as the native enzyme and in complex with its presumed reaction product, the CoA ester of DHNA, referred to here as naphthoyl-CoA (NCoA). While this work was nearing completion, another crystal structure of *Mtb* MenB, alone and in complex with acetoacetyl-CoA, was published (Truglio *et al.*, 2003). The crystal structures we present here are in different space groups, however, and with a different ligand and offer further insights into the role of flexible elements in the MenB structure and in the crotonase superfamily generally.

2. Materials and methods

2.1. Cloning, expression and purification

The predicted open reading frame for MenB (Rv0548c) was amplified from genomic DNA using standard PCR protocols. Sequencing confirmed the fidelity of the resulting product. The gene was subcloned into the expression vector pProEX (Life Technologies/Invitrogen), *Escherichia coli* BL21(DE3) pRI cells were transformed with this plasmid and MenB was expressed as an N-terminal His$_6$-tagged protein. Cells were grown at 310 K for 5–6 h, induced at 297 K with isopropyl β-D-thiogalactopyranoside (IPTG) and allowed to grow overnight at this temperature. After centrifugation, the cells were lysed by sonication and the His-tagged MenB was purified from the supernatant by Ni^{2+}-affinity chromatography using a 5 ml Hi-Trap column (Amersham Biosciences). MenB bound very tightly to this column, with 450–750 mM imidazole being required for elution. The protein used for preparation of native MenB crystals was further purified, without removal of the His tag, by passage through a Superdex 200 size-exclusion chromatography column (Amersham Biosciences); in later purifications, however, for example those used for preparation of the NCoA complex, this step was omitted. The protein was concentrated and showed excellent monodispersity (Cp/R$_H$ = 10–18%) when checked by dynamic light scattering (DynaPro 200, Protein Solutions).

2.2. Crystallization

All crystallization trials were carried out at 291 K using the sitting-drop or hanging-drop vapour-diffusion methods. The initial search for crystallization conditions made use of commercially available crystallization screens (Hampton Research) and our own in-house screens formulated using orthogonal arrays (Kingston *et al.*, 1994). The initial conditions were then optimized by fine screening around the successful conditions. Native MenB crystals were obtained in two different crystal forms. Form 1 crystals were grown by mixing equal volumes of protein solution (6–11 mg ml^{-1} protein in 50 mM Tris–HCl pH 8.0, 0.15 M NaCl) and precipitant {0.2 M 3-[4-(2-hydroxyethyl)-1-piperazinyl]propane sulfonic acid (EPPS)–KOH buffer pH 7.9, 16% methoxyPEG (mPEG) 5000}. Native form 2 crystals, which were used to obtain the initial structure solution, were grown in the same way but with 15% mPEG 5000. Form 2 crystals were also used for preparation of the NCoA complex and were grown in essentially the same way except that the protein solution also
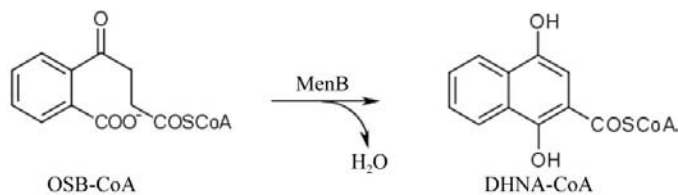

Figure 1
Reaction catalysed by MenB. Ring closure converts the CoA ester of *O*-succinyl benzoic acid (OSB) to that of 1,4-dihydroxy-2-naphthoic acid (DHNA).

Table 1
Data-collection statistics.

Values for the highest resolution shell are given in parentheses.

	Native MenB	MenB–NCoA complex
Space group	$P4_32_12$	$C2$
Unit-cell parameters (Å, °)	$a = b = 80.3$, $c = 257.6$	$a = 112.1$, $b = 114.8$, $c = 88.0$, $\beta = 123.4$
Mosaicity (°)	0.4	0.7
Resolution range	30–2.15 (2.23–2.15)	25–2.30 (2.38–2.30)
Measured reflections	454665	647330
Unique reflections	47221 (4635)	41361 (4070)
R_{merge}	0.114 (0.772)	0.139 (0.635)
Multiplicity	9.6	15.6
Completeness	100.0 (99.9)	99.9 (99.3)
Mean $I/\sigma(I)$	22.6 (3.0)	18.8 (3.0)

Table 2
Refinement and model details.

Values for the highest resolution shell are given in parentheses.

	Native MenB	NCoA complex
Resolution limits (Å)	30–2.15	25–2.30
No. of reflections	45158 (6262)	39895 (5584)
R factor	0.203	0.197
R_{free}	0.231	0.225
Model details		
Protein atoms	5599	6092
Water molecules	212	211
Other molecules/ions	1 EPPS	1 EPPS, 1 partial NCoA
R.m.s. deviations from standard geometry		
Bond lengths (Å)	0.007	0.008
Bond angles (°)	1.3	1.3
Average B factors (Å^3)		
Protein atoms (A, B, C chains)	30.1, 28.1, 27.3	33.9, 32.7, 30.2
Water molecules	31.6	32.5
Ligands		
EPPS	24.8	28.6
NCoA	—	57.8
Residues in most favoured region of Ramachandran plot (%)	90.5	89.7

contained ∼0.5 M imidazole and the concentration of mPEG 5000 in the precipitant was 20%. The form 1 crystals were tetragonal, space group $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = b = 80.3$, $c = 257.6$ Å and a V_M value (Matthews, 1968) of 1.8 Å^3 Da^{-1}, consistent with the presence of three MenB molecules (MW = 37.8 kDa including the polyhistidine tag) in the crystal asymmetric unit and a solvent content of 32.3%. The form 2 crystals were monoclinic, space group $C2$, unit-cell parameters $a = 112.1$, $b = 114.8$, $c = 88.0$ Å, $\beta = 123.4°$, with $V_M = 2.1$ Å^3 Da^{-1}, consistent with three MenB molecules per asymmetric unit and a solvent content of 40.7%.

2.3. Data collection

All crystals were flash-frozen for data collection by soaking in cryoprotectant (mother liquor plus 15% 2-methyl-2,4-pentanediol) immediately prior to placement in a stream of cold N_2 gas (110 K). For preparation of the NCoA complex, the cryoprotectant was supplemented with 13 mM NCoA using a freshly prepared solution and the crystal was soaked for 3.5 min. Native data for MenB were collected at the National Synchrotron Light Source (NSLS), Brookhaven on beamline X8C at a wavelength of 1.1000 Å. Data for the initial form 2 native crystal and for the NCoA complex were collected with a MAR 345 image plate on a Rigaku RU-H3R rotating-anode generator equipped with Osmic mirrors using Cu $K\alpha$ radiation ($\lambda = 1.5418$ Å). In each case, the raw data were processed and scaled using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Highly redundant data sets were collected (multiplicities of 9.6 and 15.6) at the expense of some increase in R_{merge} values. Full data-collection statistics are shown in Table 1.

2.4. Structure determination

The native MenB structure was solved by molecular replacement using *AMoRe* (Navaza, 1994), with a search model that comprised the complete trimer of enoyl-CoA hydratase (Engel *et al.*, 1996; PDB code 1dub). This structure has 28% sequence identity with MenB from *M. tuberculosis* and was used without removing any side chains, but after removal of all solvent molecules. The structure solution was based on a lower resolution (2.7 Å) native MenB data set that

had been obtained from a form 2 ($C2$) crystal (data not shown). When the higher resolution (2.15 Å) tetragonal data set became available, the solution was readily transferred to this crystal form and at the same time confirmed that the space group was $P4_32_12$ rather than its enantiomorph $P4_12_12$. Initial model building took place with *ARP/wARP* (Perrakis *et al.*, 2001) and returned a model that comprised 557 residues out of the 942 expected for the MenB trimer. Manual model building in *O* (Jones *et al.*, 1991) into the *SIGMAA*-weighted $2F_o - F_c$ map generated by *ARP/wARP* then gave a model that comprised 660 residues, with $R = 0.351$ and $R_{free} = 0.357$ [the R_{free} test set (Brünger, 1992) having been chosen randomly at the start of refinement]. Further refinement took place with *CNS* (Brünger *et al.*, 1998), with cycles of simulated-annealing coordinate refinement and individual B-factor refinement interspersed with further manual model building in *O*. Tight NCS restraints were used initially, but later in refinement these were loosened and then removed. Water molecules were added using *WATERPICK* in *CNS*, but were only retained in the model if they had good spherical electron density, made appropriate hydrogen-bonding interactions with respect to both geometry and hydrogen-bonding partners and were not in regions where the protein structure was in doubt. One EPPS molecule was added at the centre of the trimer in three alternative conformations, each having one-third occupancy. The final model comprised 717 residues of the expected 942, 212 water molecules and one EPPS molecule and gave final values of R and R_{free} of 0.203 and 0.231, respectively, for all data to 2.15 Å resolution. Full refinement details are given in Table 2.

The structure of the NCoA complex of MenB was solved using *MOLREP* (Vagin & Teplyakov, 1997) with the native structure as a search model and was refined with *CNS* using similar protocols to those employed for the native structure. Moderate NCS restraints and grouped B factors (two per residue) were used early in refinement, but the NCS restraints

were later loosened significantly as it became apparent that there were differences between the three molecules in the asymmetric unit. Individual B factors were then also assumed and refined. Water molecules were added as for the native structure. A partial NCoA molecule, lacking the naphthoyl ring and β-mercaptoethylamine moiety, was modelled into molecule B. As in the native crystals, an EPPS molecule in three conformations was found at the centre of the trimer. The final model consisted of 779 residues, 211 water molecules, one EPPS molecule and one partial NCoA molecule and gave final values for R and R_{free} of 0.197 and 0.225, respectively, for all data to 2.3 Å resolution. Full refinement details are given in Table 2.

For both structures, the protein geometry was monitored regularly with *PROCHECK* (Laskowski *et al.*, 1993), with the final models having 90.5 and 89.7% of residues in the most favoured regions of the Ramachandran plot, with no outliers. Intersubunit interactions were analysed using the *Protein–Protein Interaction Server* (Jones & Thornton, 1995, 1996; http://www.biochem.ucl.ac.uk/bsm/PP/server).

3. Results

3.1. Overview of the crystal structures

In both crystal forms presented here, the tetragonal ($P4_32_12$) crystals of the native MenB and the monoclinic ($C2$) crystals of its NCoA complex, essentially the same hexameric species is seen. This corresponds with dynamic light-scattering measurements and gel-filtration results, both of which are consistent with the presence of a hexamer in solution (data not shown). Both crystal forms contain three MenB molecules in the asymmetric unit, assembled into a tightly associated trimer, with the hexamer generated in each case by the application of a crystallographic twofold axis. This hexamer also corresponds closely to that reported by Truglio and coworkers for their monoclinic ($P2_1$) crystals of *M. tuberculosis* MenB (Truglio *et al.*, 2003), in which the asymmetric unit contains two complete hexamers. Although the crystal packing between hexamers is somewhat different in each case, all three structures show similar patterns of disordered regions, described below, which are clearly a characteristic feature of the MenB structure and a result of inherent flexibility.

3.2. Molecular structure

The MenB polypeptide of 314 residues has the classic crotonase fold (Engel *et al.*,

1996) comprising a large N-terminal domain of about 250 residues followed by a smaller C-terminal domain (Fig. 2). The N-terminal domain has two mostly parallel β-sheets (β-strands A0'–A5 and β-strands B2–B4) that are oriented orthogonally to each other, together with a number of helices. The parallel β-strands A1, A2, A3 and A4 are connected by helices and loops of varying length, giving four 'spiral' connections, and are preceded by strand A0 antiparallel to A1. MenB also has an N-terminal extension to the core fold, comprising a short 3_{10}-helix and a short β-strand A0' that is antiparallel to A0. At the periphery of the N-terminal domain is an antiparallel pair of helices, H5 and H6, joined by β-strand B5, which play a significant role in trimerization. The N-terminal domain is completed by a long helix H7 that packs against the main β-sheet and leads into the three final α-helices, H8, H9 and H10, which comprise the C-terminal domain, residues 244–
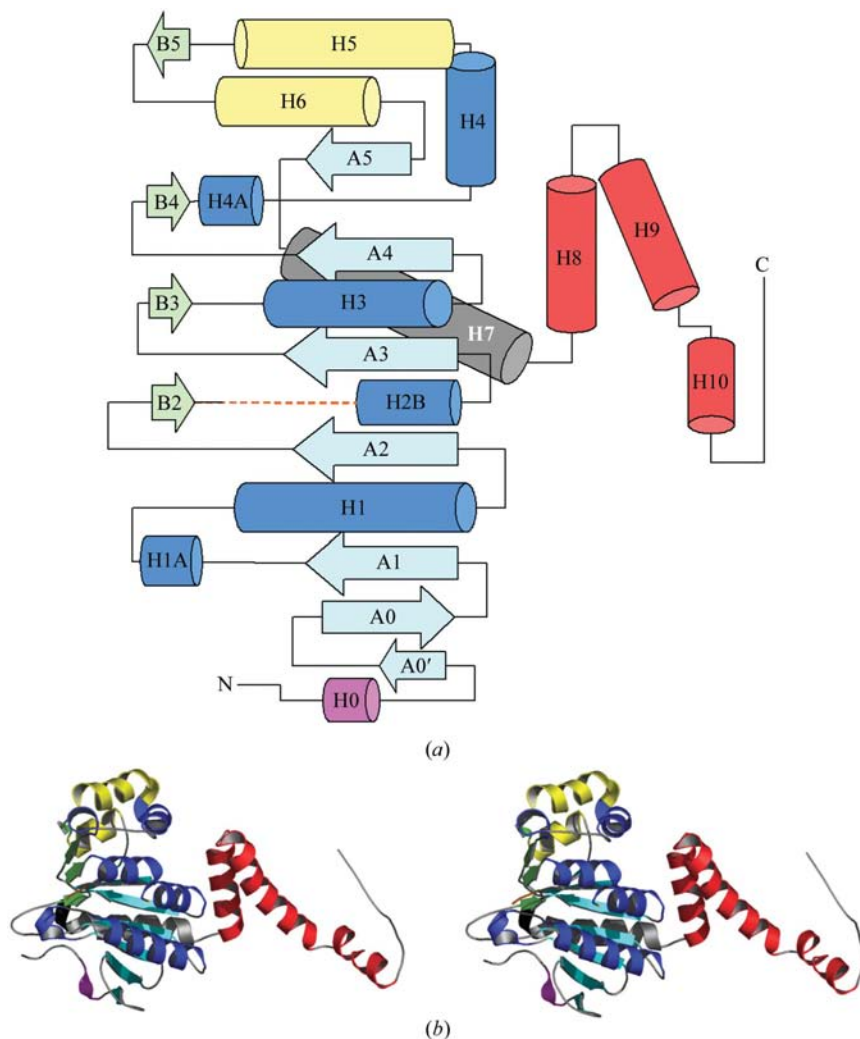

(a)

(b)

Figure 2
Structure of the MenB monomer. (*a*) Topology diagram, with β-strands shown as arrows and α-helices as cylinders. The two β-sheets of the N-terminal domain are shown in cyan and green, helices H5 and H6, which contribute to trimer formation, are in yellow and H8, H9 and H10 of the C-terminal domain are shown in red. The orange broken line denotes the disordered region that connects B2 and H2B. (*b*) Stereo ribbon diagram of the MenB monomer, with secondary structures coloured as in (*a*). This and subsequent figures were generated with *PyMol* (http://www.pymol.org).

314. Helices H8 and H9 pack against each other in an anti-parallel fashion, with H8 resting against the spiral domain, and H9 is followed by H10 and a final extended section of polypeptide.

Several sections of the polypeptide chain have no interpretable electron density and are assumed to be disordered, representing regions of flexibility in the molecule. These include, in both structures, the N-terminal residues 1–17 plus the attached His tag, approximately 25–30 residues between β-strand B2 and helix H2B, starting around Gln107 and extending to about Arg133, and approximately 30 residues at the C-terminus, from about Ala286 onwards. Even at a contour level of 0.5σ in $2F_o - F_c$ maps, only broken uninterpretable density could be seen in these regions. There are individual variations between the molecules, however. In the native MenB structure, an additional section, residues 188–190 at the end of helix H4A, is missing from molecule C, although it is present with high B factors in molecules A and B. In the NCoA complex, the full C-terminal domain to Phe314 could be modelled in molecule B and the most of helix H10, up to Arg296, could be modelled in molecule A.

In most respects, the present structures agree very closely with the MenB structure of Truglio *et al.* (2003). The three monomers in the asymmetric unit of the NCoA complex, for example, superimpose on those of the structure of Truglio and coworkers with root-mean-square (r.m.s.) differences in C^α positions of 0.33, 0.27 and 0.30 Å, respectively, over all

modelled residues. The trimer agrees almost equally well (r.m.s. difference of 0.35 Å over all 757 modelled C^α-atom positions), but the hexamer, which in the present structures is generated by crystallographic symmetry as opposed to the non-crystallographic symmetry in the structure of Truglio and coworkers, agrees slightly less well (r.m.s. difference of 0.46 Å over 1514 C^α-atom positions). This may have some impact on the degree of disorder observed; see below.

The MenB hexamer can be described as a dimer of trimers (Fig. 3). The trimer is formed by interactions across two surfaces. The first surface primarily involves residues from helices H4 and H5 of the N-terminal domain together with the preceding helix H4A and the H4A–H4 connecting strand. This packs against a second surface which comprises residues from helices H8 and H9 of the C-terminal domain, together with some residues from H6, H7 and neighbouring loops of the N-terminal domain. Taking the more complete NCoA complex as the reference, a total of ~3050 Å^2 of solvent-accessible surface is buried in each monomer–monomer interface of the trimer (~1525 Å^2 per monomer or 10.5% of the total monomer surface, taking the most complete molecule, B, as the reference). This figure is an underestimate, however, since some disordered residues are also likely to be part of the interface.

The hexamer is formed by two trimers packed 'back-to-back', with monomers of one trimer sitting almost directly over monomers of the other: A over B', B over A' and C over
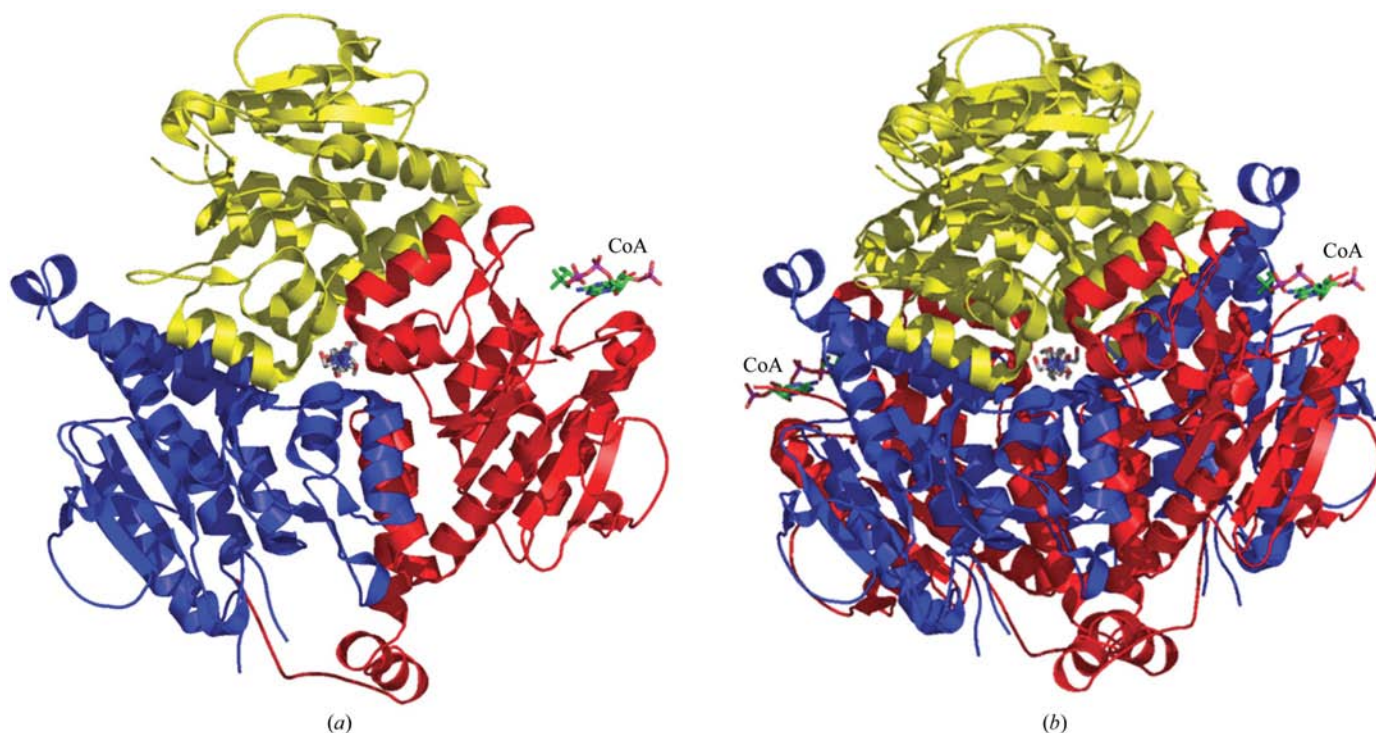

Figure 3
Oligomerization of MenB, shown for the NCoA complex. (*a*) The MenB trimer, with molecule A shown in blue, molecule B in red and molecule C in yellow. The bound EPPS molecule at the centre of the trimer is modelled in three conformations (see text) and the CoA ligand is shown in the active site of monomer B; both ligands are shown in stick mode. (*b*) The MenB hexamer, looking down the non-crystallographic threefold axis, with the MenB subunits coloured as in (*a*). The 'top' trimer is related to the trimer underneath by a crystallographic twofold axis that results in B' lying underneath A, A' underneath B and C' underneath C. The CoA molecule can be seen to occupy a shallow surface cleft formed between molecules B and A' (and A and B').

C' (Fig. 3b). Most of the trimer–trimer contacts, however, are made by residues from the C-terminal domain of each monomer, comprising the helix pair H8/H9 and the more extended structure formed by H10 and the C-terminus. These contacts are substantial, burying at least 3450 Å^2 (24% of the total surface area) per monomer; this figure is derived from monomer B, which has a complete C-terminal domain. The interface is mostly (62%) hydrophobic and clearly indicates a very stable hexamer.

The C-terminal domain from each monomer performs a dual function, stabilizing the hexamer and completing the active site of another monomer. Thus, for monomer B the H8/H9 helices insert between monomers A' and B' below, with H8 and one face of H9 interacting with B' and the other face of H9 plus H10 capping the active site of molecule A', after which the C-terminal residues 298–314 fold back to lie over the surface of monomer B'. Similarly, the C-terminal domain of monomer A inserts between B' and C', capping the active site of B' and packing against C', and the C-terminal region of monomer C inserts between C' and A', capping the active site of C' and packing against A'.

In both structures a large region of additional electron density was present at the centre of the trimer, with a large globular head associated with the side chains of Arg202 from each of the three monomers and a long tail extending down the local threefold axis. This was modelled as a molecule of the crystallization buffer 3-[4-(2-hydroxyethyl)-1-piperazinyl]propane sulfonic acid (EPPS), with its negatively charged sulfonic acid head group interacting via well ordered water molecules with the positively charged Arg202 side chains. The head group has approximate threefold symmetry, but the tail takes on three separate conformations. It is unlikely that the binding of the EPPS molecule has functional significance, but the hole through the centre of the trimer is the location of bound metal ions in other crotonase-family members, e.g. methylmalonyl-CoA decarboxylase (Benning et al., 2000), dienoyl-CoA isomerase (Modis et al., 1998) and 6-oxo-camphor hydrolase (Whittingham et al., 2003).

3.3. Active site and mode of naphthoyl-CoA binding

The active site is identified by the bound NCoA molecule (only visible in molecule B of the NCoA complex), which binds very similarly to acetoacetyl-CoA in the MenB structure of Truglio et al. (2003). The NCoA molecule is bound in a U-shaped configuration in a shallow surface cleft flanked by

three sections of polypeptide chain: residues 56–62, 95–107 and 157–159. Specific interactions involve the carbonyl O atoms of Ser103 and Gly105 which bind the adenine N atoms N1 and N6, Arg58 which binds the pyrophosphate and Lys95 which binds the ribose phosphate. Although only the adenine ring, phosphoribose, pyrophosphate and pantetheine moieties of the NCoA can be fitted into electron density (Fig. 4a), the attached thioacyl group can be modelled, with its thioester carbonyl O atom hydrogen bonded by the peptide NH groups of Gly105 and Gly161. These residues belong to the two conserved sequence motifs, FXXGGD (residues 101–106) and GGGH (residues 160–163), that identify the oxyanion hole in crotonase-family enzymes, polarizing the C=O bond and stabilizing the resulting anion (Holden et al., 2001).

Beyond the position of the thioester group is a deep and largely enclosed pocket (Fig. 4b) which is expected to be the binding site for the naphthoyl moiety and the site where the enzymatic chemistry takes place (Truglio et al., 2003). This pocket, which we refer to as the aromatic binding pocket, is
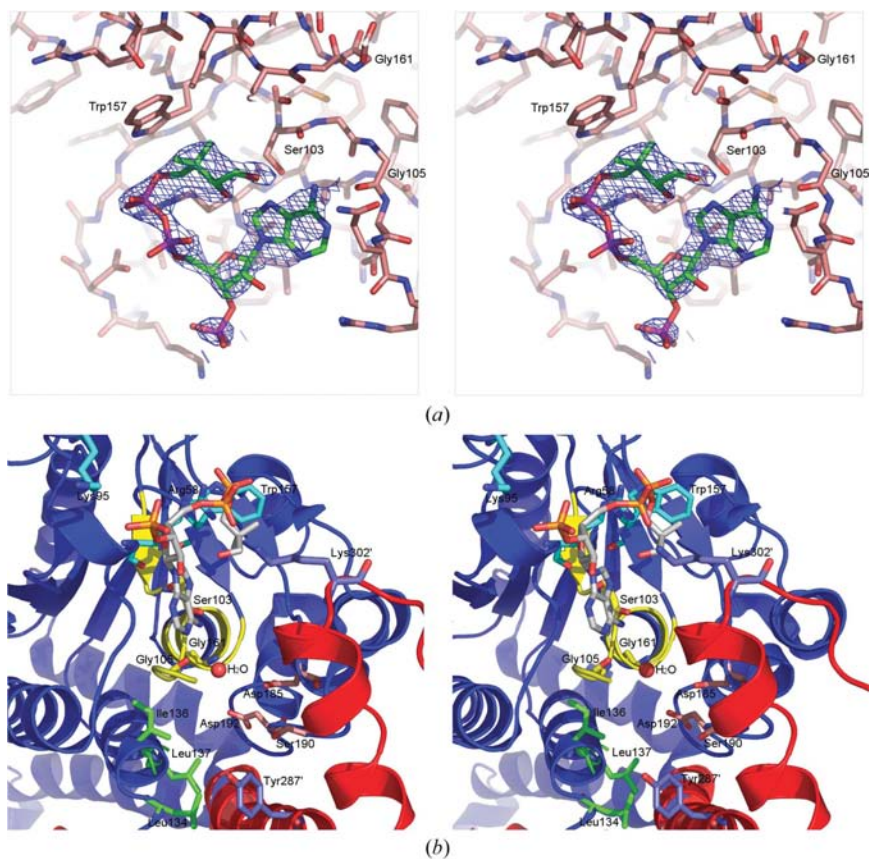

(a)

(b)

Figure 4
Ligand binding and the active site of MenB. (a) Stereoview of the electron density for the NCoA molecule bound in the active site of molecule B. The electron density is from a $2|F_o| - |F_c|$ map contoured at a level of 0.7σ. (b) Stereoview of the active-site region of MenB, shown for molecule A (the most complete molecule) of the NCoA complex. The partial NCoA molecule at the top, drawn in stick mode, is in the position in which it was experimentally modelled in molecule B. The conserved regions that form the oxyanion hole in crotonase-family enzymes are shown in yellow and a water molecule (labelled) is bound at this site in MenB. The helices H9 and H10 from the C-terminal domain of molecule B' are shown in red, capping the proposed aromatic binding pocket. Key side chains that contribute to this binding pocket are shown in stick mode: Ile136 and Leu137 are in green, Asp185, Ser190 and Asp192 in red and Tyr287 in blue.

flanked by Ile136 and Leu137 from helix H2B and by residues from helices H3 and H4A and the H4A–H4 loop; the latter loop provides the potential catalytic residues Ser190 and Asp192. Mutation of Asp192 has in fact been shown to abolish all catalytic activity of MenB (Truglio *et al.*, 2003). The binding pocket is capped by helix H9 (residues 266–288) of a different monomer, with Gly280, Thr283 and Tyr287 forming part of the lining of the pocket. This is only shown clearly in molecules *A* and *B* of the present structure, however, where the active site is capped by H9 from molecules *B'* and *A'*, respectively, from the opposing trimer; in molecules *C* and *C'*, H9 is not fully modelled.

4. Discussion

The structure of naphthoate synthase (MenB) from *M. tuberculosis* has now been determined in three different crystal forms: the two reported here and that of Truglio *et al.* (2003). All show very similar structures, with similar patterns of flexible or disordered regions, which are clearly a molecular property and not significantly influenced by crystal packing. Two of the disordered regions, the N-terminal 15–20 residues and the B2–H2B loop (residues 107–130), are on the surface and in full contact with bulk solvent, so their flexibility is not unexpected. The B2–H2B loop region also varies widely in sequence in crotonase-family enzymes and is the site of frequent insertions and deletions, consistent with it being a variable loop projecting into solution; *M. tuberculosis* MenB contains an eight-residue insertion in this loop relative to most other MenB and crotonase-family enzymes.

What is of particular interest, however, is the flexibility of the parts of the structure that are involved in function and in oligomerization. At either end of the flexible B2–H2B loop are Gly105, which provides part of the oxyanion hole, and residues 133–137, which contribute Ile136 and Leu137 to the substrate-binding pocket. Residues 185–192, which also help form the substrate-binding pocket, have high B values (50–70 Å^2) in the uncomplexed structure (with 188–190 completely missing in molecule *C*), but substantially lower B values, reduced in each molecule by about 20 Å^2, in the NCoA complex. More unexpected is the disorder in the C-terminal region, where helices H8 (254–267), H9 (266–287) and H10 (289–297) contribute both to hexamer formation and to the active site: H9 caps the aromatic binding pocket, H10 interacts with residues 187–189 from helix H4A in the active-site region, and H8, H9 and H10 are all involved in trimer–trimer contacts. Helix H8 and most of H9, up to Leu285, is well ordered in all molecules in both structures, but the rest of H9 plus H10 can only be modelled in molecules *A* and *B* of the NCoA complex and the C-terminal residues 298–314 can only be modelled in molecule *B* of the complex.

Two clear trends are evident. Firstly, even though we cannot model the whole of the naphthoyl-CoA molecule and no density at all can be seen for the naphthoyl group, flexibility is markedly reduced by ligand binding; in the NCoA complex structure, B values around the active site are lower by $\sim$20 Å^2 and some structural elements that could not be seen in the native structure can now be clearly modelled. Secondly, the formation of the aromatic binding pocket is correlated with oligomerization. Not only is the active site of each subunit capped by H9 and H10 from a subunit of the opposing trimer, but Leu134 from the flexible B2–H2B connection also contacts H9 and residues 187–189 from the flexible H4A–H4 loop contact H10. These contacts provide a mechanism whereby changes in the hexamer can be relayed to the active site and *vice versa*.

Why is the disorder in the C-terminal region more pronounced in the present structure than in that of Truglio *et al.* (2003)? One possibility is that the constraints of the crystallographic symmetry that generates the hexamer in both the two present structures do not allow it to 'relax' to its most favoured form. When the NCoA complex trimer is superimposed onto a trimer of the structure of Truglio and coworkers, the rest of the hexamer does not exactly superimpose; a slight rotation of the other trimer results in molecule *B'* corresponding quite well but *A'* and *C'* noticeably less well, resulting in displacements of 0.5–0.7 Å. Although the differences are small, they do correlate with the extents of disorder seen in the present structures, with the disorder in the C-terminal region also impinging on the ordering of the parts of the active site it contacts.

The tendency of the C-terminal region of MenB towards disorder signals an inherent flexibility in this region, which is used in diverse ways in other crotonase-family enzymes. Whereas all family members have the same fold, including helices H8, H9 and H10 of the C-terminal domain, a hinge following H8 allows various orientations of H9 and H10. Thus, in methylmalonyl CoA decarboxylase (Benning *et al.*, 2000) and enoyl CoA isomerase (Mursula *et al.*, 2001) the C-terminus is oriented such that H9 and H10 'cap' the active site of the same subunit. In contrast, in enoyl-CoA hydratase (crotonase; Engel *et al.*, 1996, 1998), 4-chlorobenzoyl-CoA dehalogenase (Benning *et al.*, 1996), AUH protein (Kurimoto *et al.*, 2001) and dienoyl-CoA isomerase (Modis *et al.*, 1998) the C-terminal region is oriented such that H9 and H10 cap the active site of the neighbouring monomer in the trimer and in MenB the C-terminal region is oriented in a different manner again to cap the active site of a monomer from the other trimer in the hexamer. Owing to this intertrimer swapping of the C-terminus seen in the MenB structures and the significance of this in complete active-site formation, it is unlikely that this enzyme would be active in anything but a hexameric quaternary arrangement.

Finally, there is the question of why the naphthoyl group is not seen in the NCoA complex. A similar result was obtained by Truglio *et al.* (2003), who reported that only the CoA portion could be seen after soaking their MenB crystals in NCoA. We envisage two possibilities. Since NCoA is the expected product of the reaction catalysed by MenB, it may be that unfavourable contacts are made by the naphthoyl group, destabilizing binding and favouring product release. Alternatively, since no gene corresponding to *menH* has yet been found in the *M. tuberculosis* genome, it is possible that *Mtb* MenB performs the reactions that are carried out by both

research papers

MenB and MenH in *E. coli* and *Bacillus subtilis*: ring closure to create the naphthoyl moiety (the MenB reaction) and cleavage of the naphthoyl group from NCoA (the MenH reaction). While this would certainly explain why no naphthoyl group could be visualized, it is unlikely to be the case as Truglio *et al.* (2003) have given experimental evidence that the product of the MenB-catalysed reaction is NCoA.

References

Barry, C. E. III, Slayden, R. A., Sampson, A. E. & Lee, R. E. (2000). *Biochem. Pharmacol.* **59**, 221–231.

Benning, M. M., Haller, T., Gerlt, J. A. & Holden, H. M. (2000). *Biochemistry*, **39**, 4630–4639.

Benning, M. M., Taylor, K. L., Liu, R.-Q., Yang, G., Xiang, H., Wesenberg, G., Dunaway-Mariano, D. & Holden, H. M. (1996). *Biochemistry*, **35**, 8103–8109.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

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

Collins, M. D. & Jones, D. (1981). *Microbiol. Rev.* **45**, 316–354.

Duncan, K. (2003). *Tuberculosis*, **83**, 201–207.

Engel, C. K., Kiema, T. R., Hiltunen, J. K. & Wierenga, R. K. (1998). *J. Mol. Biol.* **275**, 847–859.

Engel, C. K., Mathieu, M., Zeelan, J. P., Hiltunen, J. K. & Wierenga, R. K. (1996). *EMBO J.* **15**, 5135–5145.

Gerlt, J. A. & Babbitt, P. C. (2001). *Annu. Rev. Biochem.* **70**, 209–246.

Goulding, C. W., Perry, L. J., Anderson, D., Sawaya, M. R., Cascio, D., Apostol, M. I., Chan, S., Parseghian, A., Wang, S.-S., Wu, Y., Cassano, V., Gill, H. S. & Eisenberg, D. (2003). *Biophys. Chem.* **105**, 361–370.

Holden, H. M., Benning, M. M., Haller, T. & Gerlt, J. A. (2001). *Acc. Chem. Res.* **34**, 145–157.

Johnston, J. M., Arcus, V. L., Morton, C. J., Parker, M. W. & Baker, E. N. (2003). *J. Bacteriol.* **185**, 4057–4065.

Jones, S. & Thornton, J. M. (1995). *Prog. Biophys. Mol. Biol.* **63**, 31–59.

Jones, S. & Thornton, J. M. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 13–20.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kingston, R. L., Baker, H. M. & Baker, E. N. (1994). *Acta Cryst.* D**50**, 429–440.

Kurimoto, K., Fukai, S., Nureki, O., Muto, Y. & Yokoyama, S. (2001). *Structure*, **9**, 1253–1263.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

McKinney, J. D. (2000). *Nat. Med.* **6**, 1330–1333.

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

Meganathan, R. (2001). *Vitam. Horm.* **61**, 173–218.

Modis, Y., Filppula, S. A., Novikov, D. K., Norledge, B., Hiltunen, J. K. & Wierenga, R. K. (1998). *Structure*, **6**, 957–970.

Mursula, A. M., van Aalten, D. M. F., Hiltunen, J. K. & Wierenga, R. K. (2001). *J. Mol. Biol.* **309**, 845–853.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Palmer, D. R. J., Garrett, J. B., Sharma, V., Meganathan, R., Babbitt, P. C. & Gerlt, J. A. (1999). *Biochemistry*, **38**, 4252–4258.

Parrish, N. M., Dick, J. D. & Bishai, W. R. (1998). *Trends Microbiol.* **6**, 107–112.

Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001). *Acta Cryst.* D**57**, 1445–1450.

Smith, C. V. & Sacchettini, J. C. (2003). *Curr. Opin. Struct. Biol.* **13**, 1–7.

Stokstad, E. (2000). *Science*, **287**, 2391.

Terwilliger, T. C., Park, M. S., Waldo, G. S., Berendzen, J., Hung, L.-W., Kim, C.-Y., Smith, C. V., Sacchettini, J. C., Bellinzoni, M. & Bossi, R. (2003). *Tuberculosis*, **83**, 223–249.

Truglio, J. J., Theis, K., Feng, Y., Gajda, R., Machutta, C., Tonge, P. J. & Kisker, C. (2003). *J. Biol. Chem.* **278**, 42352–42360.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Wayne, L. G. & Sohaskey, C. D. (2001). *Annu. Rev. Microbiol.* **55**, 139–163.

Whittingham, J. L., Turkenburg, J. P., Verma, C. S., Walsh, M. A. & Grogan, G. (2003). *J. Biol. Chem.* **278**, 1744–1750.

Xiang, H., Luo, L., Taylor, K. L. & Dunaway-Mariano, D. (1999). *Biochemistry*, **38**, 7638–7652.