

Exploiting the high-resolution crystal structure of *Staphylococcus aureus* MenH to gain insight into enzyme activity

Dawson *et al.*

RESEARCH ARTICLE

Open Access

Exploiting the high-resolution crystal structure of *Staphylococcus aureus* MenH to gain insight into enzyme activity

Alice Dawson, Paul K Fyfe, Florian Gillet and William N Hunter*

Abstract

Background: MenH (2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase) is a key enzyme in the biosynthesis of menaquinone, catalyzing an unusual 2,5-elimination of pyruvate from 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate.

Results: The crystal structure of *Staphylococcus aureus* MenH has been determined at 2 Å resolution. In the absence of a complex to inform on aspects of specificity a model of the enzyme-substrate complex has been used in conjunction with previously published kinetic analyses, site-directed mutagenesis studies and comparisons with orthologues to investigate the structure and reactivity of MenH.

Conclusions: The overall basic active site displays pronounced hydrophobic character on one side and these properties complement those of the substrate. A complex network of hydrogen bonds involving well-ordered water molecules serves to position key residues participating in the recognition of substrate and subsequent catalysis. We propose a proton shuttle mechanism, reliant on a catalytic triad consisting of Ser89, Asp216 and His243. The reaction is initiated by proton abstraction from the substrate by an activated Ser89. The propensity to form a conjugated system provides the driving force for pyruvate elimination. During the elimination, a methylene group is converted to a methyl and we judge it likely that His243 provides a proton, previously acquired from Ser89 for that reduction. A conformational change of the protonated His243 may be encouraged by the presence of an anionic intermediate in the active site.

Background

Menaquinone, vitamin K₂, is a lipid soluble quinone used by bacteria for electron transport. It is the sole quinone available to Gram-positive bacteria [1] whereas Gram-negative bacteria use ubiquinone under aerobic conditions, and menaquinone under anaerobic conditions [2]. Vitamin K, as menaquinone or the plant form phyloquinone, is a cofactor for γ -carboxylation of glutamate, a post-translational modification of proteins involved in blood coagulation and vascular biology and is an essential component of the mammalian diet [3].

There are two distinct biosynthetic routes to menaquinone. One, recently discovered, occurs in a small subset of bacteria such as *Helicobacter pylori* [4]. The starting point is chorismate supplied by the Shikimate pathway

and the intermediates include the inosine derivative futasoline. The most prevalent, classical biosynthetic route to menaquinone, has been most studied in *Escherichia coli* [1]. The starting point remains chorismate, which is isomerized to isochorismate then converted, in three enzyme-catalyzed stages, to *o*-succinyl-1-benzoate. Two further reactions, involving ATP and CoA, lead to a CoA-naphthalene derivative before a thioesterase [5] removes the CoA moiety to release 1,4-dihydroxy-2-naphthanoate. Prenylation, with products from the isoprenoid biosynthesis pathway, and finally a methylation reaction completes the production of menaquinone.

To attain a comprehensive understanding of menaquinone biosynthesis, a fundamental aspect of bacterial metabolism, it is imperative to characterize the molecular structures and properties of the relevant enzymes. Our research in this area is confined to the classical

* Correspondence: w.n.hunter@dundee.ac.uk
Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK

menaquinone biosynthetic route and we previously reported on MenB [6] and MenD [7,8]. We turn our attention now to MenH, the enzyme which catalyzes the third step in this pathway, a 2,5-elimination of pyruvate from 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate (SEPHCHC) to produce 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC, Figure 1).

There has been confusion about the reaction catalyzed by MenH, in large part due to misassignment of the enzymatic activities of MenB and MenD. Initial reports [9] described MenD as the enzyme that cleaved the pyruvate side chain from the isochorismate - α -ketoglutarate adduct SEPHCHC. MenB was assumed to cyclize and then cleave CoA from a naphthanyl derivative [10], and the product encoded by the *menH* gene had no function assigned. It was subsequently proven that MenB lacked such CoA esterase activity and then postulated that MenH could catalyze this reaction [1]. Indeed, many databases still list *menH* as encoding a thioesterase. More recently, Jiang *et al.* [11] showed that MenD is not a bifunctional enzyme but is SEPHCHC synthase [E.C. 2.2.1.9]. Structural and kinetic studies are compatible with this assignment [7,8]. Rigorous biochemical analysis then proved that MenH is actually responsible for the elimination of the pyruvate side chain (E.C. 4.2.99.20) and that previous studies had been compromised by the inherent reactivity of SEPHCHC, which slowly decomposes to SHCHC and pyruvate [12].

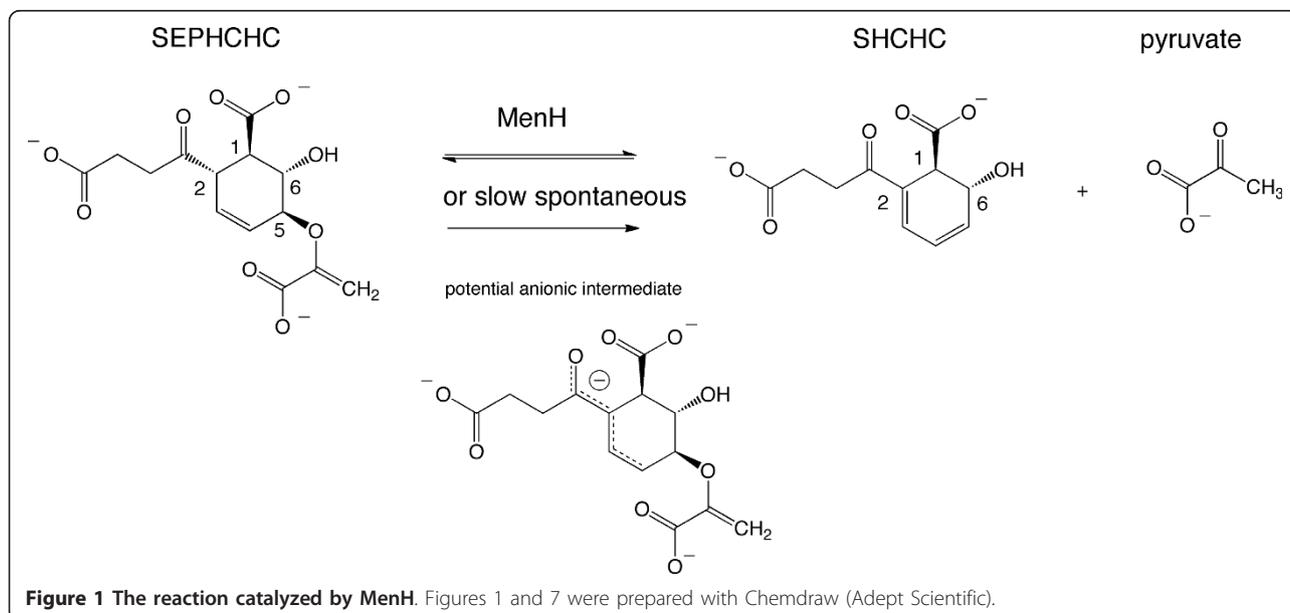
We now describe the high-resolution crystal structure of *Staphylococcus aureus* MenH (*SaMenH*), compare it to the structure of the *Vibrio cholerae* enzyme (*VcMenH*) deposited in the Protein Data Bank (PDB)

and discuss conserved residues with particular focus on what such residues contribute to the architecture of the active site. Consideration of data derived from site-directed mutagenesis and kinetic studies of the *E. coli* enzyme (*EcMenH*) [12,13] allow us to reassess the molecular features that influence specificity and reactivity of MenH and address two issues of the structure-mechanism relationship, which remain unclear. The assignment of an oxyanion-binding site formed by two amino acid side chains by Jiang *et al.* [13] is highly unusual and is revisited. Finally the origin of the proton that converts a methylene group to a methyl (Figure 1) is considered and a proton-shuttle mechanism, involving transfer from substrate to protein then back at a different position is proposed.

Results and Discussion

General comments and overall structure

An efficient bacterial expression system for a histidine-tagged *SaMenH* was constructed and supplied material for structural studies. Approximately 40 mg of protein was obtained from 1 L of *E. coli* culture. The enzyme crystallizes in the monoclinic space group *C2* with one molecule in the asymmetric unit and diffraction data to better than 2 Å were measured in-house. The structure was solved by molecular replacement using a model based on PDB entry 1r3d, which although listed as a protein of unknown function from *V. cholerae*, has been identified as MenH [12]. The sequence identity with *SaMenH* is about 20% and the best molecular replacement solution had a translational Z-score of 6.4, low for a correct solution. The initial electron and difference density maps were of poor quality but with defined



features that suggested a correct molecular replacement solution had been obtained. Analysis of molecular packing in the crystal lattice also engendered confidence in the solution therefore we proceeded with refinement. The maps gradually improved with rounds of careful model building and refinement. Once all ordered MenH residues were included in the model a search for water molecules, to be included in the refinement, was initiated. A conservative approach to water identification was adopted. The crystallographic statistics and model geometry indicate that the refinement has produced an acceptable high-resolution model (Table 1). At the end of the refinement there was continuous electron density for the entire polypeptide chain, residues 1-266, missing only the Gly-His residues that remain at the N-terminus following proteolytic cleavage of the histidine tag.

Approximately 65% of the *SaMenH* residues are in well-defined elements of secondary structure. These comprise 7 β -strands and 11 α -helices arranged in the α/β hydrolase fold [14,15], a central β -sheet with a marked twist, sandwiched between α -helices (Figure 2). Strand β 1 is antiparallel to the others. Helices α 1 and α 11 lie in the concave face of the β -sheet and α 4 to α 8 form an insertion in the canonical α/β hydrolase fold to create a domain that caps the active site. The interface between the cap and core domains, which involves 11 hydrogen bonds, occludes approximately 25% of the surface area of the cap domain itself ($\sim 1500 \text{ \AA}^2$ of $\sim 5900 \text{ \AA}^2$).

Table 1 Crystallographic statistics

Resolution range (\AA)	19.73 - 1.94
Unit cell dimensions <i>a,b,c</i> (\AA) β ($^\circ$)	76.6, 43.6, 71.7, 98.3
Space group	C2
No. reflections measured, unique	71,206, 17,242
Multiplicity ^a , Completeness (%)	4.1(3.3), 98.9 (93.3)
Wilson B (\AA^2)	24.5
$R_{\text{merge}}(\%)^b$, $\langle I^2(l) \rangle$	5.3 (37.3), 17.5 (3.2)
Protein residues, water molecules	266, 185
$R_{\text{work}}(\%)^c$, $R_{\text{free}}(\%)^d$	18.3, 24.1
DPI (\AA) ^e	0.18
Average B-factors (\AA^2)	
Overall, main chain, side chain, waters	30.6, 29.3, 31.8, 31.3
R.m.s.d bond lengths (\AA), angles ($^\circ$)	0.012, 1.25
Ramachandran plot analysis (%)	
Favorable, Outliers	97.7, 0

^aValues in parentheses refer to the highest resolution shell. 2.05 - 1.94 \AA ^b $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$; where $I_i(hkl)$ is the intensity of the i th measurement of reflection hkl and $\langle I(hkl) \rangle$ is the mean value of $I_i(hkl)$ for all i measurements. ^c $R_{\text{work}} = \sum_{hkl} |F_o| - |F_c| / \sum |F_o|$, where F_o is the observed structure factor and F_c is the calculated structure factor. ^d R_{free} is the same as R_{cyst} except calculated with a subset, 5%, of data that are excluded from refinement calculations. ^eDPI, Diffraction-component Precision Index [39].

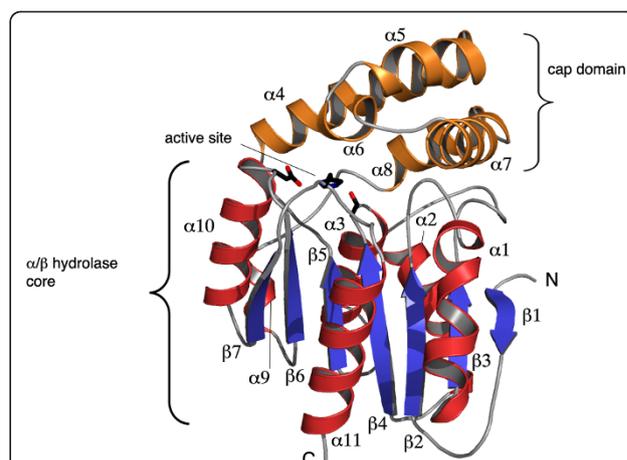


Figure 2 The structure of *SaMenH* and position of the active site. The polypeptide is depicted as a ribbon with helices of the core domain colored red and β -strands blue, helices of the cap domain are copper colored. N and C mark the amino and carboxy termini of the protein. The catalytic triad consisting of Ser89, His243 and Asp216 are shown as sticks and mark the active site. Figures 2, 3, 4 and 6 prepared using PyMOL[41] and Illustrator (Adobe Systems Inc.).

Members of the α/β hydrolase family display different oligomerization states. For example, the C-C bond hydrolase MphC is dimeric in both solution and crystalline states, with an extended β -sheet formed across the dimer interface [16] whilst others such as acetylcholine esterase are monomeric [17]. *SaMenH* behaves as a monomer in solution, as indicated by size-exclusion chromatography (data not shown). Analysis of the crystal structure using the Protein Interfaces, Surfaces and Assemblies server [18] indicated no significant areas of interaction to suggest oligomerization of MenH in the crystalline state (data not shown).

Comparisons with related structures

There are over 660 structures known for α/β hydrolase family members and they encompass a wide range of differing functions such as esterases, hydrolases, lipases, haloperoxidases and peptidases [14,17,19]. The fold is tolerant to insertions, with proteins ranging from 200 to 600 residues in length. The canonical α/β hydrolase structure contains a core composed of eight β -strands [14,15]. However, in *SaMenH* the core β -sheet consists of only seven strands due to the absence of the N-terminal strand observed in other family members. A common insertion in the canonical α/β hydrolase fold is a lid, or cap domain over the active site. In *SaMenH* the cap is formed by five helices, α 4 to α 8 (Figure 2). The cap domain is important in defining the active site size and aspects of substrate specificity as will be described. In some structures the cap domain appears to be flexible as

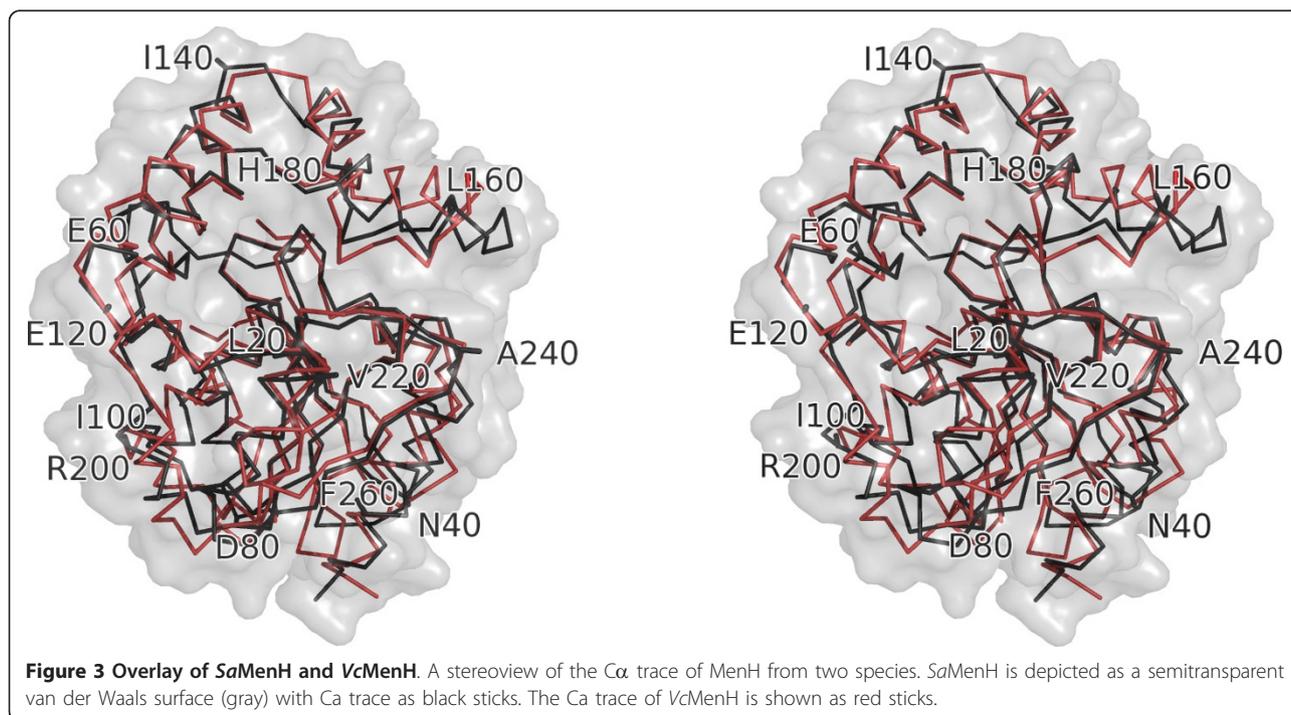
indicated by higher temperature factors. However, this is not the case in *SaMenH* (data not shown) where we see no correlation between the selected Translation-Libration-Screw (TLS) domains and the cap/core domains. This observation should be treated with caution since the *SaMenH* crystals contain a relatively low volume of bulk solvent (approximately 35%) and the molecular packing in the crystal lattice involves direct interactions of residues in the cap domain with three symmetry related molecules. The tight packing may therefore artificially restrict flexibility of the cap domain with respect to the core of the structure.

Analysis using Secondary Structure Matching [20] indicates that, as expected, *SaMenH* is similar to various α/β hydrolases. However, with the exception of *VcMenH*, there are no particularly close matches in the PDB. The structural homologues include haem-independent haloperoxidases, various esterases and hydrolases that overlay with *SaMenH* with root-mean-squared deviation (r.m.s.d.) values in the range 2.0-2.4 Å, all with approximately 230 aligned C α atoms and sequence identities around 20%. Similar results have been reported in structural comparisons of the α/β hydrolases, for example in an investigation of the carboxylesterase BioH from *E. coli* [21]. Of note is the Rv0554 protein from *Mycobacterium tuberculosis* recently investigated as a potential MenH orthologue, which falls within the range observed for the other α/β -hydrolases (r.m.s.d. 2.3Å for overlay of 227 C α positions) and has no activity with SEPHCHC [22].

A closer alignment is observed when comparing *SaMenH* with *VcMenH*, proteins which share 21% sequence identity. The r.m.s.d. is 1.8 Å for least-squares fit of 232 C α atoms (Figure 3). The loop linking β 3 and α 2, a region of high sequence variability in MenH orthologues (data not shown), is not modelled in the structure of *VcMenH* but is ordered in *SaMenH*. The largest difference between the two structures is in the orientation of the N-terminal region of β 7, part of the lid domain, where C α positions differ by up to 4 Å (Figure 3). This overlay highlights similarities in and around the active site that are likely important for enzyme activity and will be mentioned in the next section.

The architecture of the MenH active site and recognition of substrate

The residues that form the *SaMenH* catalytic triad, Ser89, Asp216 and His243 are positioned on one side of a water-filled tunnel formed between the cap and core domains (Figure 4). The residues that contribute to the active site are contributed from α 4, α 5, α 8, α 10, two turns between β 2- α 1 and β 4- α 3 respectively and the loops that follow β 5 and β 7 (Figure 2). The entrance to the active site is near Arg127 and the tunnel then extends around and over the catalytic center to Asn32, a distance of about 25 Å (data not shown). A complex network of hydrogen bonds, involving well ordered water molecules, serve to organize the residues around the active site. The electrostatic properties complement



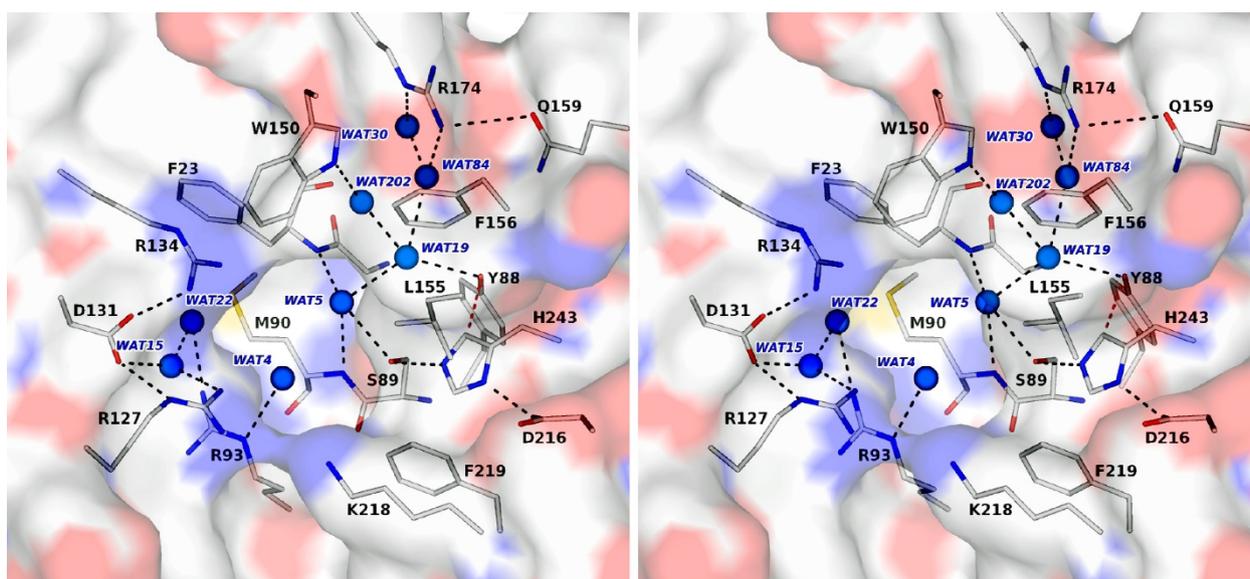


Figure 4 Stereoview of the active site, catalytic triad and oxyanion-binding site. *SaMenH* is depicted as semi transparent van der Waals surface colored according to atom type (C grey, N blue, O red, S yellow). Selected amino acid side chains are shown as sticks colored according to atom type. The main chain of Met90 and Gly22-Phe23 are included to mark the oxyanion-binding site. Marine colored spheres represent selected water molecules. Black dashed lines mark potential hydrogen bonding interactions, the C-H...O interaction between Tyr88 and His243 is represented by a dashed red line.

the negatively charged substrate, which carries three carboxylate groups, at positions C1, C2, C5 and a hydrophobic component (Figure 1). A cavity is formed to the side of this tunnel near Phe23, adjacent to the catalytic machinery (Figure 4).

As is typical of the α/β hydrolase fold, the catalytic serine (Ser89) occurs in a GX SXG motif (Figure 5) located on a tight turn, often termed the 'nucleophile elbow', between a central β -strand and a helix. The two glycine residues help create the tight turn, which is also stabilized by a hydrogen bond formed between the main chain carbonyl of Ser89 and amide nitrogen of Gly92

(data not shown). Asp216 is located on a second tight turn, between $\beta 6$ and $\alpha 10$, and is shielded from the active site by Leu155 and Phe219. The hydrogen bonding interaction between Asp216 and His243, helps to position the imidazole to interact with the hydroxyl of Ser89 in the apo-structure. Ser89 OG also accepts a hydrogen bond donated from a particularly well-ordered water molecule ($B = 7 \text{ \AA}^2$, where B is the Debye-Waller or isotropic temperature factor, $8\pi^2 \langle u^2 \rangle$ with u the atomic displacement parameter), which in turn accepts two hydrogen bonds from the main chain amides of Phe23 and Met90.

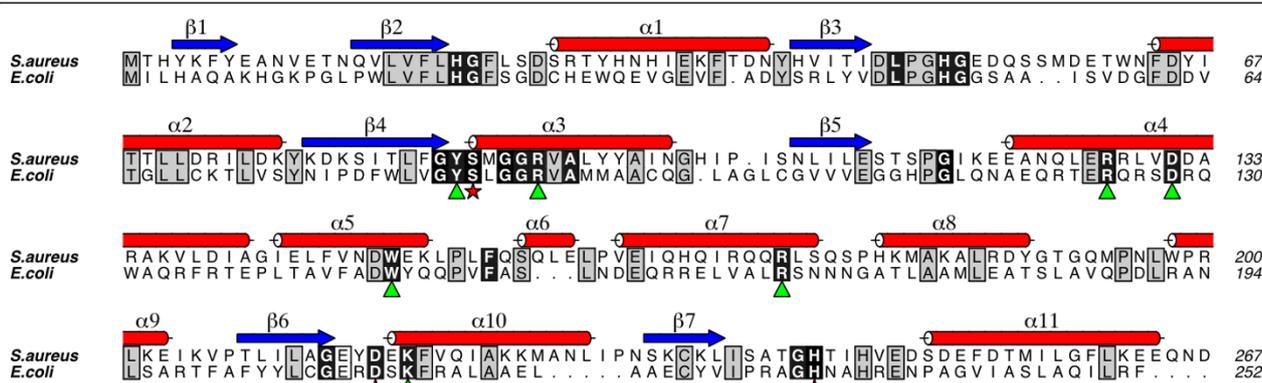


Figure 5 A sequence alignment of *SaMenH* and *EcMenH*. Residues outlined in black are 90% conserved in 101 MenH sequences. Residues outlined in grey are conserved in *SaMenH* and *EcMenH*. Residues marked with a star form the catalytic triad and those residues were mutated by Jiang et al.[13] in *EcMenH* are marked with a triangle. This figure was prepared using MUSCLE [42] and ALINE [43].

The active site has a pronounced hydrophobic patch created by Trp150, Phe156, Leu153, Leu155 and Phe219, and in addition Phe23 and Leu186 (not shown) help form the base of the site (Figure 4). The position of Trp150 is determined by van der Waals interactions with Phe23 and Phe146 (not shown) on one side of the indole, Phe156 on the other side and the side chain of Arg134 abuts CZ3. This orientation of the Trp150 side chain places NE1 directed towards the active site to form a hydrogen bond with an ordered water molecule (Figure 4) that contributes to forming a network of solvent molecules in the active site. Phe219, Leu155 and Phe156 provide a hydrophobic environment around the side chain of His243 helping to hold it in place and providing a microenvironment that may increase the pK_a .

The side chain of Tyr88 is positioned by van der Waals interactions with Gly22 C α , Ser89 C β , Glu112 CD and His246 CE1. His246 is held in place by hydrogen bonds formed between ND1 and NE2 with Glu112 OE1 and Thr29 OG1 respectively (data not shown). As a result, Tyr88 OH is placed to participate in hydrogen bonds with Thr244 OG1 and an ordered water molecule, ($B = 18 \text{ \AA}^2$). A distance of 3.2 \AA separates His243 CD2 from Tyr88 OH and the geometry is compatible with the presence of a C-H...O hydrogen bond (Figure 4) [23]. The C-H...O hydrogen bond is a weak interaction but may contribute stability to the positive-charge distribution on the protonated histidine and in so doing support a catalytic function. Such a role has been invoked in enzymes such as trypanothione reductase, which is dependent on proton transfer from histidine [24].

There are six basic residues in the active site (Arg93, Arg127, Arg134, Arg174, Lys218 and the catalytic His243) and a single acidic residue, Asp131 (Figure 4). Lys218 is on the periphery of the active site with the aliphatic component of the side chain lying on the side chain of Phe219. Arg127 and Arg134, together with Asp131 are opposite Lys218. Arg127 forms both a direct and a water-mediated hydrogen-bonding association with Asp131, which also accepts a hydrogen bond from Arg134. The water that bridges Arg127 and Asp131, Wat15 ($B = 17 \text{ \AA}^2$) accepts a hydrogen bond donated by the amide of Thr191. Arg93 is held in place by forming a hydrogen bond to the main chain carbonyl of Tyr189 and to two water molecules in the active site, Wat22 ($B = 21 \text{ \AA}^2$) and Wat4 ($B = 22 \text{ \AA}^2$). Arg174 is about 10 \AA from the catalytic center lining the polar tunnel formed at the interface of the cap and core domains. Arg174 NH2 donates a hydrogen bond to Gln159 OE1 and to four water molecules, two of which are shown in Figure 4.

The overall sequence identity between MenH proteins is relatively low and a comparison highlights key residues in and around the active site (Figure 5). Of 256

orthologues in the Kyoto Encyclopedia of Genes and Genomes (KEGG) [25], 101 are unique when filtered at a 90% sequence identity level. Within these sequences, 23 residues are conserved in at least 90% of them. The catalytic triad is conserved in all but one sequence where aspartate is replaced by glutamate. Eight of the remaining strictly conserved residues are glycine, which are all located at the end of secondary structure elements or in loops and may be important in defining the overall structure, for example in the "elbow" motif in the active site. The acidic Asp131, five of the nine hydrophobic residues and five out of the six basic residues listed above as contributing to the architecture of the active site are strictly conserved. Three of the hydrophobic residues, Met90, Phe146, Leu155 are replaced by valines or leucine. Arg134 is not a highly conserved residue, being substituted by Trp131 in *EcMenH* for example (Figure 5).

The structural overlay of *SaMenH* and *VcMenH*, discussed above, indicates that the conserved residues in the active site adopt similar orientations and only the side chain of Arg174 (*SaMenH* numbering) differs significantly (data not shown). In *SaMenH*, Arg174 is oriented towards $\alpha 5$. In *VcMenH*, the corresponding residue, Arg176, adopts a different rotamer and the side chain is directed towards $\alpha 1$ whilst Tyr156 occupies the site of the *SaMenH* Arg174 guanidinium (data not shown). This observation suggests a degree of conformational freedom deep in the active site that might influence the position of the cap domain.

Attempts to soak ligands such as pyruvate, salicylic acid and the substrate SEPHCH into crystals of *SaMenH*, and so derive structures directly relevant to substrate recognition resulted in the loss of diffraction. Co-crystallization experiments were also carried out but either the crystals did not grow or when available and subsequently analyzed, there was no electron density that could be attributed to these ligands (data not shown). In the absence of any experimentally determined structures we sought to prepare a simple model of the enzyme-substrate (ES) complex. We were guided by seeking a position of SEPHCH in the active site in which there was little steric hindrance and where the chemical properties of substrate and active site were complementary to each other. The potential for His243 to act as a base and acquire the C2 proton direct from substrate was recognized. However, when the substrate was positioned to facilitate the interaction of SEPHCH C2 with His243 NE2 there were severe steric clashes with the protein. The positions of the Leu155 and Phe156 side chains in particular appear to compromise such a model. We cannot rule out the possibility that gross structural changes might change the active site to facilitate the use of His243 as an active site base for the

MenH reaction but note that there is no evidence to support such conformational changes to the catalytic residues utilized within the α/β hydrolase family of enzymes. An alternative model was constructed by placing SEPHCHC C2 3.3 Å from Ser89 OG. The ligand was simply rotated about C2 to minimize steric clash with the enzyme (Figure 6) and, as will be described, the model suggests plausible chemical interactions to stabilize the substrate in the enzyme active site.

Our ES model provides a template to consider aspects of substrate recognition and MenH catalysis. In this we expect that important features relevant to substrate recognition and catalysis be conserved in MenH sequences and the α/β -hydrolase fold as discussed above. We noted also that the position of water molecules might identify where functional groups of substrate could bind. A thorough analysis, combining kinetic studies with the use of site-directed mutagenesis studies has been applied to seven conserved MenH residues using the *E. coli* enzyme (Figure 5). The kinetic results were interpreted in light of a homology model of *EcMenH* based on *VcMenH* combined with the use of docking calculations [12,13]. Our *SaMenH* ES model (Figure 6), which should be consistent with the kinetic data, indeed corroborates certain aspects of the previous study but there are significant differences in the models, which we will detail.

Our ES model suggests that Arg127 is placed to interact with the C6 hydroxyl and or the C5 carboxylate. Lys218 could also interact with the C5 carboxylate and Arg93 is well placed to form a salt-bridge interaction with the C1 carboxylate. The substrate C2 carboxylate, the succinyl substituent, is positioned in the solvent filled channel and placed to accept hydrogen bonds

donated from Tyr88 OH and Trp150 NE1. The hydrophobic component, C3, C4 and the methylene group are positioned over towards the hydrophobic patch on the active site near Leu155 and Phe156. Mutations of the conserved Arg93, Arg127 and Arg174 to alanine leads to increased K_M values and reduced catalytic properties. This is compatible with a role of these arginine residues in binding and orientating the substrate in the active site. Arg174 does not interact with substrate in the ES model but we cannot rule out a conformational change that facilitates an interaction between the guanidinium and the C2 carboxylate. The side chain of Arg174 is implicated in positioning Trp150 and disruption of the structure in this part of the active site might contribute to the compromised level of enzyme activity. Asp131 interacts with Arg127 and mutation of the corresponding residue, Asp128, in *EcMenH* to alanine leads to an eight-fold increase in K_M [13] consistent with localized disruption to substrate binding.

The mechanism of typical α/β hydrolases involves the formation of an oxyanion intermediate stabilized in the active site by an "oxyanion-binding hole" usually formed by two main chain amides. One amide is contributed from the residue immediately following the catalytic serine, the second can vary in location but often resides on the loop linking a β -strand and helix. Jiang *et al.* [13] proposed that the oxyanion hole is located between the side chains of Tyr88 and Trp150 (*SaMenH* numbering) based on results of computational docking calculations. This would be an extremely unusual situation not previously observed in this enzyme family. In *SaMenH* the main chain amides of Met90 and Phe23 are well placed to form the oxyanion hole; indeed a well-ordered water molecule is observed within hydrogen bonding distance

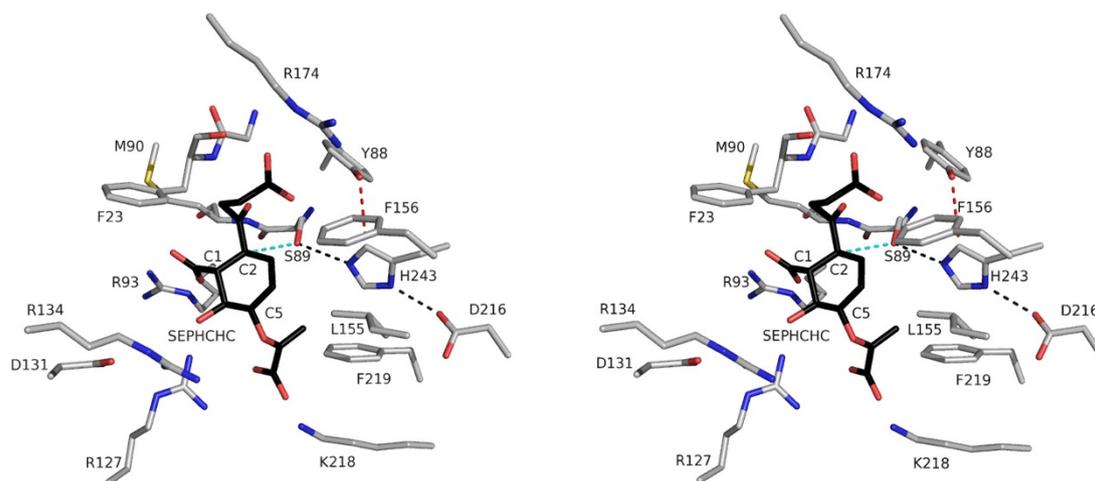


Figure 6 A model for substrate binding to *SaMenH*. A stereoview of the ES complex model. SEPHCHC is depicted with C positions colored black. MenH residues are shown as in Figure 4. Trp150 is not shown for the purpose of clarity. The chemical structure modelled represents Stage I in Figure 7.

of both nitrogen atoms in addition to the side chain of Ser89 (Figure 4). A similar situation occurs in *VcMenH*, with the corresponding residues being Ser91, Leu92 and Leu24. Structural alignment with other α/β hydrolases where the oxyanion hole has been identified support our interpretation [26,27]. In our ES model, a carbonyl oxygen of the succinyl substituent, where a negative charge might build up, overlays on the position of this ordered water molecule (Figure 6). That being the case we suggest a different role for the conserved tryptophan (*SaMenH* Trp150). Mutation of the equivalent residue in *EcMenH*, Trp147 to phenylalanine, had significantly less effect on catalysis than a Trp147 to alanine mutation [13]. This further suggests that the side chain is not actually involved in the formation of the oxyanion hole. Our ES model suggests that Trp150 stabilizes the hydrophobic part of the active site by interactions with Phe146, Phe156 and Leu153 and that Trp150 NE1, in partnership with Tyr88 OH, donate hydrogen bonds to bind the carboxylate of the succinyl side group of the substrate (Figure 6). Alteration of Tyr88 to phenylalanine increases K_M slightly but reduces catalytic efficiency to a marked degree. This would support the earlier proposal that interaction with His243 might influence catalysis.

A proposed mechanism for MenH

All α/β hydrolases contain a catalytic triad consisting of a nucleophile, an acid and a histidine. In many cases, the reaction proceeds in a similar fashion to that of the serine proteases, where the substrate is subject to nucleophilic attack by serine, for example prolyl oligopeptidase [28]. The histidine and aspartate components of the catalytic triad are essential for *EcMenH* activity; mutation of the serine resulted in a considerably less active, but still functional enzyme, suggesting a mechanism distinct from most other α/β hydrolases [12,13] but similar to the C-C bond hydrolases MhpC and BhpD. Mutation of the active site serine in MhpC reduces but does not abolish enzyme activity and additional kinetic

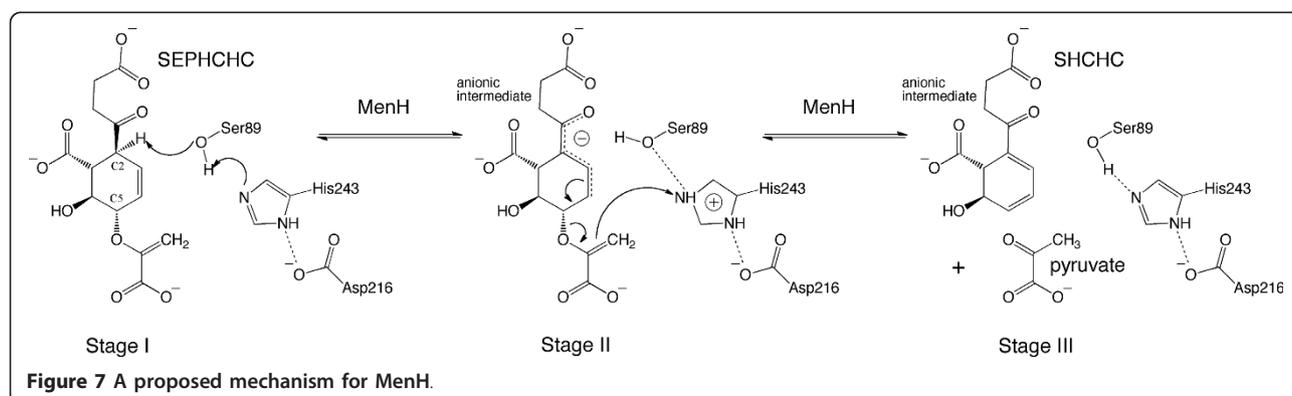
studies lead to the assignment of a general base catalysis mechanism [29]. The non-essentiality of the active site serine in MenH suggests that it may also adopt a general base mechanism.

The reaction catalyzed by MenH can be described in three stages (Figure 7). In stage I, the catalytic triad structure positions His243 to abstract a proton from Ser89 OG and so prepare a seryl oxide which in turn supports removal of the acidic C2 proton from SEPHCHC. An anionic intermediate is generated. Stage II is breaking the C5-O bond to release the pyruvate side chain; a common step in biosynthetic reactions that use chorismate. However, in most cases, the product is aromatic [30] and MenH is unusual in generating a partially saturated product. Nevertheless, the stability expected when a conjugated system is present is likely a major driving force in pyruvate elimination from C5 subsequent to the abstraction of the proton from C2. In stage III, methylene is reduced to the pyruvate methyl by addition of a proton and the product then leaves the active site.

Whereas mutation of the active site serine does not completely destroy MenH activity, alteration of the catalytic His243 is catastrophic. The explanation may reside in the combination of two factors. Firstly, the substrate has a relatively labile acidic C2 proton irrespective of whether a seryl oxide is nearby to acquire it or not. The difference is in the rate of proton loss. Secondly, His243 is primed by association with Asp216 and Tyr88 to abstract a proton from any nearby hydroxyl group or a water molecule, which can then be passed on to reduce the methylene to a methyl. An attraction between this intermediate and the protonated His243 could assist a movement of the amino acid side chain into a position whereby donation of a proton to the methylene group is possible.

Conclusions

The high-resolution crystal structure of *SaMenH*, an α/β hydrolase, has been determined and detailed comparisons carried out with proteins of related fold. The



MenH active site properties complement those of the substrate, which is negatively charged and possesses significant hydrophobic character. The enzyme:substrate complex has been modelled and a potential mechanism, exploiting previously published data combining site-directed mutagenesis with kinetic studies of *EcMenH*, is proposed. The reaction may be initiated by proton abstraction from the substrate by an activated Ser89, part of a catalytic triad with Asp216 and His243. The driving force for pyruvate elimination is provided by the formation of a conjugated system. During pyruvate elimination, a methylene group is converted to a methyl and it is possible that His243 provides a proton, acquired from Ser89 for that reduction.

Methods

Expression and purification of *SaMenH*

Chemicals, of the highest quality available, were sourced from Sigma-Aldrich and VWR International except where stated otherwise. The *menH* gene was amplified by PCR from *S. aureus* genomic DNA (strain American Type Culture Collection 35556) using 5'-**CATATGA-CACACTACAAGTTTTATGAAGC**-3' as the forward primer and 5'-**GGATCCTTAGTCATTTTGCTCCTCC**-3' as the reverse primer (Thermo Scientific). These oligonucleotides contain 5' *NdeI* and 3' *BamHI* sites indicated in bold. The PCR product was ligated into TOPO-BLUNT-II (Invitrogen) and then subcloned into a modified pET15b vector (Novagen), which produces an N-terminally hexa-histidine tagged protein with a Tobacco Etch Virus (TEV) protease site. The resulting expression vector was heat shock transformed into *E. coli* strain BL21 (DE3) (Stratagene). Cell cultures were grown in 1 L volumes using Luria-Bertani media supplemented with 50 $\mu\text{g mL}^{-1}$ of carbenicillin at 37°C to an optical density of 0.6. The cultures were cooled to room temperature, expression induced using 0.5 mM isopropyl- β -D-thiogalactopyranoside, and growth continued for 16 hours at room temperature.

Cells were harvested by centrifugation for 25 minutes at 40,000 g at 4°C, resuspended in lysis buffer (50 mM Tris-HCl, 250 mM NaCl, 20 mM imidazole pH 7.5) containing DNase I (0.1 mg), and a single tablet of a cocktail of EDTA-free protease inhibitors (Roche), and lysed using a French press at 1000 psi. Insoluble debris was separated by centrifugation at 39,000 g for 25 minutes, at 4°C. The supernatant was loaded onto a 5 mL HisTrap HP column (GE) precharged with Ni^{2+} . A linear concentration gradient of imidazole was applied, with *SaMenH* eluting at approximately 160 mM imidazole. Fractions were analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and those containing *SaMenH* were pooled and incubated with TEV

protease at 30°C for three hours. The untagged protein was separated from the TEV protease, tag and residual tagged protein using a second Ni^{2+} charged HisTrap column; the untagged protein still bound weakly to the column, eluting at approximately 80 mM imidazole. The untagged *SaMenH* was then further purified using a Superdex 200 26/60 gel filtration column (GE), equilibrated using 50 mM Tris-HCl, 250 mM NaCl pH 7.5. The protein eluted as a single species with mass of approximately 30 kDa consistent with that of a monomer. Selected fractions were pooled, dialyzed into 20 mM Tris-HCl, 50 mM NaCl, pH 7.5 and concentrated for crystallization. The protein concentration was estimated spectrophotometrically using a theoretical extinction coefficient of $32,890 \text{ M}^{-1} \text{ cm}^{-1}$ (ProtParam) [31].

Crystallization and data collection

Initial screening was performed using sitting drop vapor diffusion with standard sparse matrix screens. Multiple hits were obtained in conditions containing polyethylene glycol (PEG). Further tests with a PEG screen (Qiagen) yielded a diffraction quality crystal from reservoir conditions 0.1 M sodium HEPES pH 7.5, 25% w/v PEG 4000 in a drop consisting of 1 μL of protein solution (5 mg mL^{-1}) and 1 μL of reservoir. Rod shaped crystals took several weeks to appear at 20°C. A crystal (approximately $0.5 \times 0.08 \times 0.08 \text{ mm}^3$) was mounted directly from the drop, without any additional cryoprotection, into a flow of cooled nitrogen at -173°C. Data were collected in-house using a Rigaku Micromax 007 rotating anode ($\text{Cu K}\alpha \lambda = 1.5418 \text{ \AA}$) R-Axis IV⁺⁺ image plate system. The data were integrated using XDS [32] and scaled with Scala [33]. Statistics are presented in Table 1.

Structure solution and refinement

The search model for molecular replacement calculations was derived from PDB entry 1r3d, the MenH orthologue from *V. cholerae*. Non-conserved side chains were truncated to alanine using Chainsaw [34], and molecular replacement carried out in Phaser [35]. Initial rigid body refinement gave an R-factor of 55%. Rounds of model manipulation using Coot [36] interspersed with refinement using Refmac5 [37] were used to complete the protein model; addition of water molecules then completed the refinement. The R_{work} and R_{free} values converged at 18.3 and 24.1% respectively. TLS refinement using three domains (analysis used the TLS server [38]) was included as part of the refinement. The three domains consisted of residues 2-45, 46-119 and 120-265. The Diffraction-component Precision Index (DPI), as defined by Cruickshank [39], and the Ramachandran plot statistics, as calculated in Molprobit [40], are given in Table 1.

Protein Data Bank accession numbers

Coordinates and structure factors have been deposited with accession code 2xmz.

Abbreviations

CoA: coenzyme A; E.C.: Enzyme Commission; PEG: polyethylene glycol; PDB: Protein Data Bank; r.m.s.d.: root mean square deviation; *Sa*: *Staphylococcus aureus*; SEPHCHC: 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate; SHCHC: 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate; TEV: Tobacco Etch Virus; TLS: Translation/Libration/Screw; Vc: *Vibrio cholerae*.

Acknowledgements

Supported by awards from BBSRC [Structural Proteomics of Rational Targets grant BBS/B/14434], and The Wellcome Trust [grant numbers WT082596 and WT083481]. We thank Rachel Morgan and Jennifer Fleming for useful discussions, and Zhihong Guo for a gift of SEPHCHC.

Authors' contributions

AD cloned, expressed, purified and crystallized SaMenH, refined the structure and drafted the manuscript. PKF collected data and aided in the refinement. FG aided in protein purification and crystallization. WNH conceived of the study and drafted the manuscript. All authors read and approved the final manuscript.

Received: 13 January 2011 Accepted: 22 April 2011

Published: 22 April 2011

References

- Meganathan R: Biosynthesis of menaquinone (Vitamin K₂) and ubiquinone (coenzyme Q): a perspective on enzymatic reactions. *Vitam Horm* 2001, **61**:173-218.
- Bekker M, Kramer G, Hartog AF, Wagner MJ, de Koster CG, Hellingwerf KJ, de Matos MJ: Changes in the redox state and composition of the quinone pool of *Escherichia coli* during aerobic growth. *Microbiology* 2007, **153**:1974-1980.
- Bandyopadhyay PK: Vitamin K-dependent γ -glutamylcarboxylation: an ancient posttranslational modification. *Vitam Horm* 2008, **78**:157-184.
- Hirasaka T, Furihata K, Ishikawa J, Yamashita H, Itoh N, Seto H, Dairi T: An alternative menaquinone biosynthetic pathway operating in microorganisms. *Science* 2008, **321**:1670-1673.
- Widhalm JR, van Oostende C, Furt F, Basset GJ: A dedicated thioesterase of the hotdog-fold family is required for the biosynthesis of the naphthoquinone ring of vitamin K₁. *Proc Natl Acad Sci USA* 2009, **106**:5599-5603.
- Ulaganathan V, Agacan MF, Buetow LS, Tulloch LB, Hunter WN: Structure of *Staphylococcus aureus* 1,4-dihydroxy-2-naphthoyl-CoA synthase (MenB) in complex with acetoacetyl-CoA. *Acta Crystallogr Sect F Struct Biol & Cryst Comm* 2007, **63**:908-913.
- Dawson A, Fyfe PK, Hunter WN: Specificity and reactivity in menaquinone biosynthesis: the structure of *Escherichia coli* MenD (2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate synthase). *J Mol Biol* 2008, **384**:1353-1368.
- Dawson A, Chen M, Fyfe PK, Guo Z, Hunter WN: Structure and reactivity of *Bacillus subtilis* MenD catalyzing the first committed step in menaquinone biosynthesis. *J Mol Biol* 2010, **401**:253-264.
- Palaniappan C, Sharma V, Hudspeth ME, Meganathan R: Menaquinone (vitamin K₂) biosynthesis: evidence that the *Escherichia coli* menD gene encodes both 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylic acid synthase and alpha-ketoglutarate decarboxylase activities. *J Bacteriol* 1992, **174**:8111-8118.
- Sharma V, Suvarna K, Meganathan R, Hudspeth ME: Menaquinone (vitamin K₂) biosynthesis: sequence and expression of the menB gene from *Escherichia coli*. *J Bacteriol* 1992, **174**:5057-5062.
- Jiang M, Cao Y, Guo ZF, Chen M, Chen X, Guo Z: Menaquinone Biosynthesis in *Escherichia coli*: identification of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate as a novel intermediate and re-evaluation of MenD activity. *Biochemistry* 2007, **46**:10979-10989.
- Jiang M, Chen X, Guo ZF, Cao Y, Chen M, Guo Z: Identification and characterization of (1R,6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase in the menaquinone biosynthesis of *Escherichia coli*. *Biochemistry* 2008, **47**:3426-3434.
- Jiang M, Chen X, Wu XH, Chen M, Wu YD, Guo Z: Catalytic mechanism of SHCHC synthase in the menaquinone biosynthesis of *Escherichia coli*: identification and mutational analysis of the active site residues. *Biochemistry* 2009, **48**:6921-6931.
- Heikinheimo P, Goldman A, Jeffries C, Ollis DL: Of barn owls and bankers: a lush variety of α/β hydrolases. *Structure* 1999, **7**:R141-R146.
- Nardini M, Dijkstra BW: Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr Opin Struct Biol* 1999, **9**:732-737.
- Dunn G, Montgomery MG, Mohammed F, Coker A, Cooper JB, Robertson T, Garcia JL, Bugg TDH, Wood SP: The structure of the C-C bond hydrolase MhpC provides insights into its catalytic mechanism. *J Mol Biol* 2005, **346**:253-265.
- Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolov F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J: The alpha/beta hydrolase fold. *Protein Eng* 1992, **5**:197-211.
- Krissinel E, Henrick K: Inference of macromolecular assemblies from the crystalline state. *J Mol Biol* 2007, **372**:774-797.
- Carr PD, Ollis DL: α/β hydrolase fold: an update. *Prot Pep Lett* 2009, **16**:1137-1148.
- Krissinel E, Henrick K: Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr Sect D Biol Crystallogr* 2004, **60**:2256-2268.
- Sanishvili R, Yakunin AF, Laskowski RA, Skarina T, Evdokimova E, Doherty-Kirby A, Lajoie GA, Thornton JM, Arrowsmith CH, Savchenko A, Joachimiak A, Edwards AM: Integrating structure, bioinformatics, and enzymology to discover function: BioH, a new carboxylesterase from *Escherichia coli*. *J Biol Chem* 2003, **278**:26039-26045.
- Johnston JM, Jiang M, Guo Z, Baker EN: Structural and functional analysis of Rv0554 from *Mycobacterium tuberculosis*: testing a putative role in menaquinone biosynthesis. *Acta Crystallogr D Biol Crystallogr* 2010, **66**:909-917.
- Leonard GA, McAuley-Hecht K, Brown T, Hunter WN: Do C-H...O hydrogen bonds contribute to the stability of nucleic acid base pairs? *Acta Crystallogr D Biol Crystallogr* 1995, **51**:136-139.
- Bond CS, Zhang Y, Berriman M, Cunningham ML, Fairlamb AH, Hunter WN: Crystal structure of *Trypanosoma cruzi* trypanothione reductase in complex with trypanothione, and the structure-based discovery of new natural product inhibitors. *Structure* 1999, **7**:81-89.
- Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M: KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 2010, **38**:D355-D360.
- Longhi S, Mannesse M, Verheij HM, De Haas GH, Egmond M, Knoops-Mouthuy E, Cambillau C: Crystal structure of cutinase covalently inhibited by a triglyceride analogue. *Protein Sci* 1997, **6**:275-286.
- Schleberger C, Sachelaru P, Brandsch R, Schulz GE: Structure and action of a C-C bond cleaving α/β hydrolase involved in nicotine degradation. *J Mol Biol* 2007, **367**:409-418.
- Fulop V, Bocskai Z, Polgar L: Prolyl oligopeptidase: an unusual β -propeller domain regulates proteolysis. *Cell* 1998, **94**:161-170.
- Li JJ, Bugg TDH: Investigation of a general base mechanism for ester hydrolysis in C-C hydrolase enzymes of the α/β hydrolase superfamily: a novel mechanism for the serine catalytic triad. *Org Biomol Chem* 2007, **5**:507-513.
- Ziebart KT, Toney MD: Nucleophile specificity in anthranilate synthase, aminodeoxychorismate synthase, isochorismate synthase and salicylate synthase. *Biochemistry* 2010, **49**:2851-2859.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A: Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*. Edited by: Walker JM. Humana Press Inc., Totowa, NJ, USA; 2005:571-607.
- Kabsch W: Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* 1993, **26**:795-800.
- Evans P: Scaling and assessment of data quality. *Acta Crystallogr Sect D Biol Crystallogr* 2006, **62**:72-82.
- Stein N: CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J Appl Cryst* 2008, **41**:641-643.
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ: Phaser crystallographic software. *J Appl Cryst* 2007, **40**:658-674.

36. Emsley P, Cowtan K: **Coot: model-building tools for molecular graphics.** *Acta Crystallogr Sect D Biol Crystallogr* 2004, **60**:2126-2132.
37. Murshudov GN, Vagin AA, Dodson EJ: **Refinement of macromolecular structures by the maximum-likelihood method.** *Acta Crystallogr Sect D: Biol Crystallogr* 1997, **53**:240-255.
38. Painter J, Merritt EA: **TLSMD web server for the generation of multi-group TLS models.** *J Appl Cryst* 2006, **39**:109-111.
39. Cruickshank DW: **Remarks about protein structure precision.** *Acta Crystallographica Sec D Biol Crystallogr* 1999, **55**:583-601.
40. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC: **MolProbity: all-atom contacts and structure validation for proteins and nucleic acids.** *Nucleic Acids Res* 2007, **35**:W375-383.
41. DeLano WL: **The PyMOL Molecular Graphics System.** DeLano Scientific LLC, Palo Alto, CA; 2008 [<http://www.pymol.org>].
42. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Res* 2004, **32**:1792-1797.
43. Bond CS, Schuttelkopf A: **ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments.** *Acta Crystallogr Sect D Biol Crystallogr* 2009, **65**:510-512.

doi:10.1186/1472-6807-11-19

Cite this article as: Dawson et al.: Exploiting the high-resolution crystal structure of *Staphylococcus aureus* MenH to gain insight into enzyme activity. *BMC Structural Biology* 2011 **11**:19.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

