Davies, C., Heath, R.J., White, S.W. & Rock, C.O. The 1.8 angstrom crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from Esc...

ARTICLE *in* STRUCTURE · FEBRUARY 2000

Impact Factor: 5.62 · DOI: 10.1016/S0969-2126(00)00094-0 · Source: PubMed

4 AUTHORS, INCLUDING:

Richard J Heath
St. Jude Children's Research Hospital
47 PUBLICATIONS **3,471** CITATIONS
SEE PROFILE

Charles O Rock
St. Jude Children's Research Hospital
243 PUBLICATIONS **12,183** CITATIONS
SEE PROFILE

The 1.8 Å crystal structure and active-site architecture of β-ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*

Christopher Davies[1†], Richard J Heath[2], Stephen W White[1,3*] and Charles O Rock[2,3*]

Background: β-Ketoacyl-acyl carrier protein synthase III (FabH) initiates elongation in type II fatty acid synthase systems found in bacteria and plants. FabH is a ubiquitous component of the type II system and is positioned ideally in the pathway to control the production of fatty acids. The elucidation of the structure of FabH is important for the understanding of its regulation by feedback inhibition and its interaction with drugs. Although the structures of two related condensing enzymes are known, the roles of the active-site residues have not been experimentally tested.

Results: The 1.8 Å crystal structure of FabH was determined using a 12-site selenium multiwavelength anomalous dispersion experiment. The active site (Cys112, His244 and Asn274) is formed by the convergence of two α helices and is accessed via a narrow hydrophobic tunnel. Hydrogen-bonding networks that include two tightly bound water molecules fix the positions of His244 and Asn274, which are critical for the decarboxylation and condensation reactions. Surprisingly, the His244→Ala mutation does not affect the transacylation reaction suggesting that His244 has only a minor influence on the nucleophilicity of Cys112.

Conclusions: The histidine and asparagine active-site residues are both required for the decarboxylation step in the condensation reaction. The nucleophilicity of the active-site cysteine is enhanced by the α-helix dipole effect, and an oxyanion hole promotes the formation of the tetrahedral transition state.

Addresses: [1]Department of Structural Biology and [2]Department of Biochemistry, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA and [3]Department of Biochemistry, University of Tennessee, Memphis, Tennessee 38163, USA.

[†]Present address: School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK.

*Corresponding authors.
E-mail: stephen.white@stjude.org
charles.rock@stjude.org

Key words: acyl carrier protein, crystal structure, coenzyme A, condensing enzymes, fatty acids

Received: **13 September 1999**
Revisions requested: **6 October 1999**
Revisions received: **16 November 1999**
Accepted: **16 November 1999**

Published: **1 February 2000**

Structure 2000, **8**:185–195

0969-2126/00/$ – see front matter

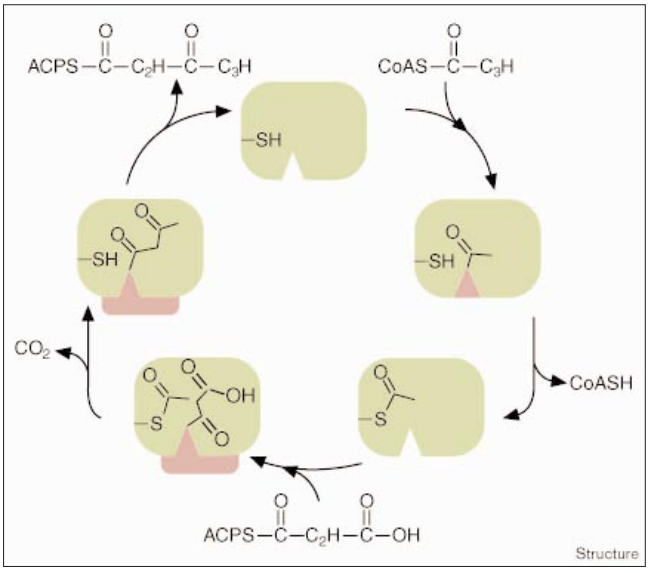
Introduction

The type II, or dissociated, fatty acid synthase system is the primary source of membrane fatty acids in bacteria and plants (for reviews, see [1,2]). This ubiquitous and highly conserved pathway has been studied most extensively in *Escherichia coli* and consists of a collection of individual enzymes that are encoded by separate genes. This general organization contrasts sharply with the type I fatty acid synthase system in animals in which a single multifunctional polypeptide catalyzes all of the reactions in the elongation pathway. This distinguishing characteristic makes the type II system a prime target for antibiotics [3–5] and amenable to manipulation to create transgenic plants with altered seed oil compositions [6]. The recent discovery of type II enzymes in *Plasmodium falciparum* suggests that this pathway might also be relevant to the development of antimalarial drugs [7]. The bacterial and plant polyketide synthases are closely related condensing enzymes that are involved in producing an array of useful natural products [8–10].

In the type II system, fatty acid elongation occurs in two-carbon steps by the Claisen condensation of malonyl-ACP with acyl-ACP (ACP is acyl carrier protein, a small, acidic soluble protein that shuttles the elongating chain between enzymes). Three enzymes catalyze these condensation reactions, FabB, FabF and FabH. FabH or β-ketoacyl-ACP synthase III initiates the process, and FabB and FabF carry out the elongation reactions in subsequent cycles of fatty acid elongation. All three enzymes are thought to operate by a similar catalytic mechanism but FabH is distinguished from the others by its selectivity for acetyl-CoA over acyl-ACP [11,12]. FabH operates via a ping–pong mechanism utilizing an acetyl-enzyme intermediate (Figure 1). FabB and FabF have a broad, overlapping substrate specificity and utilize a range of acyl-ACPs that are between 4 and 16 carbons long.

As the initiator of the fatty acid elongation cycle, FabH is ideally positioned to play a major role as a regulator of the pathway. The levels of long-chain acyl-ACPs fluctuate as a

Figure 1



Mechanism of FabH. FabH from *E. coli* catalyzes a ping–pong reaction. Acetyl-CoA reacts with the free enzyme to form an acetyl-enzyme intermediate on Cys112. CoA is released from FabH followed by the binding of malonyl-ACP. The condensation reaction occurs releasing CO_2, and in the final step, acetoacetyl-ACP is released. FabH therefore catalyzes two half reactions. Acetyl-CoA:ACP transacylase reflects the formation of the acetyl-enzyme intermediate and the exchange of the acetyl group between the CoA and ACP acceptors. The second half reaction is malonyl-ACP decarboxylase which occurs in the absence of an acetyl-enzyme intermediate and the presence of the second substrate. In this schematic mechanism the enzyme is shown in green and the substrates in pink.

Table 1

Data collection and processing statistics.

	λ1	λ3	HR
λ (Å)	1.008	0.9789	1.000
Resolution (Å)	2.5	2.5	1.8
Reflections (measured)	99,978	183,671	230,489
Reflections (unique)	24,176	23,987	61,752
Multiplicity	4.1	7.7	3.7
Completeness (%) (outer shell)	99.9 (99.7)	99.7 (99.4)	97.4 (89.8)
R_{sym}* (outer shell)	3.5 (6.2)	4.9 (8.5)	5.6 (14.0)
Phasing power† (acentrics/centrics)		0.96/0.76	
R_{cullis}‡ (acentrics/centrics)		0.86/0.80	
R_{cullis}-anomalous		0.68	
FOMmlphare		0.455	
FOMdm		0.796	

Data were collected on beamline F2 at the Cornell High Energy Synchrotron Source (CHESS). Data were integrated with MOSFLM [37], and scaled and merged with SCALA [38]. Phasing calculations were performed using MLPHARE [42]. *$R_{sym} = \Sigma\Sigma | I_i - I_m | / \Sigma\Sigma I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of all symmetry-related reflections. †Phasing power = F_H / E_{RMS}. ‡$R_{cullis} = \Sigma | | F_{PH} \pm F_P | - F_H(calc) | / \Sigma | F_{PH} \pm F_P |$. F_P, F_{PH} and F_H are the protein, derivative and heavy-atom structure factors, respectively, and E_{RMS} is the residual lack of closure.

function of the rate of phospholipid synthesis, and these molecules act as feedback regulators of fatty acid biosynthesis [13–17]. FabH is one of the targets for these acyl-ACPs [18], and long-chain acyl-ACPs are the most potent FabH inhibitors *in vitro* [11]. Given that the activity of the type II pathway is absolutely essential for cell survival, the individual enzymes are each an Achilles heel in bacteria. Existing drugs that act against the pathway, such as isoniazid [3,19] and triclosan [4,5,20], block the enoyl-ACP reductase (FabI) component. Cerulenin is a fungal antibiotic that inhibits both FabB and FabF condensing enzymes by covalent modification of the active-site sulfhydryl [21,22]. Thiolactomycin is an example of another natural product that inhibits all condensing enzymes, including FabH [23,24].

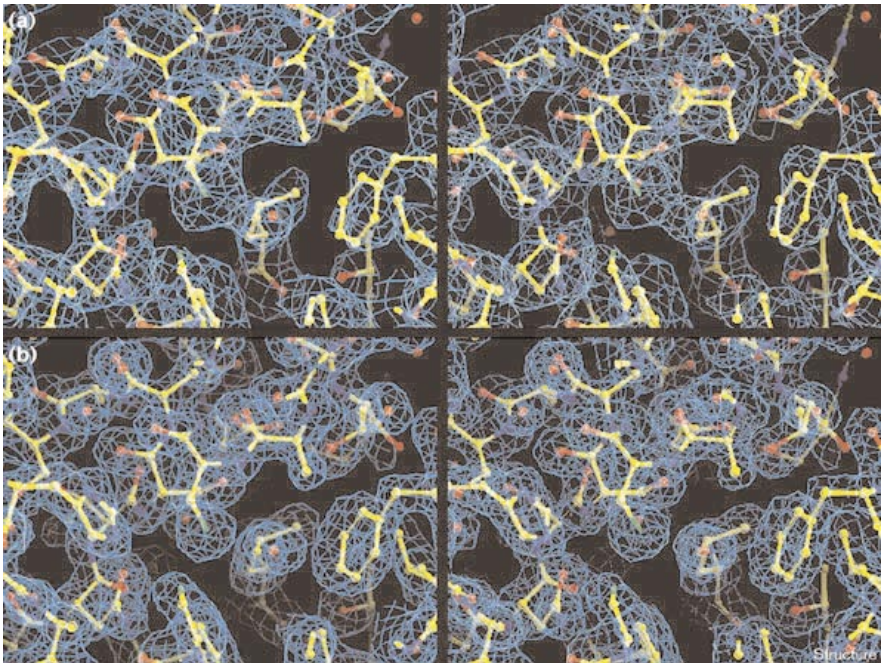
The crystal structures of two potentially related condensing enzymes, FabF [25] and chalcone synthase [26], have been determined. FabF is a non-essential enzyme in *E. coli* that is primarily responsible for the formation of *cis*-vaccenoyl-ACP [1]. Chalcone synthase plays an essential role in the production of plant phenylpropanoids, which are the precursors of a variety of useful natural products [10,26]. Roles for the active-site residues in the Claisen condensation reaction have been proposed in both cases but have not, as yet, been tested experimentally. We have determined the structure of FabH, and have analyzed the active-site residues to understand the catalytic mechanism of the Claisen condensation, and the regulation of the enzyme. An important goal of this work is to discover how existing antibiotics inhibit FabH activity and to provide a structural framework for the design of new drugs.

Results

Crystallization and structure determination

Two crystal forms of FabH in space group $P2_12_12_1$ were grown using ammonium sulfate as the precipitant at pH 7.5 (see the Materials and methods section). The crystals that were used for the structural analysis have unit-cell dimensions a = 63.1 Å, b = 64.3 Å and c = 165.9 Å, diffract to 1.8 Å and contain a FabH dimer in the asymmetric unit. To perform a multiwavelength anomalous dispersion (MAD) analysis, selenomethionyl FabH was crystallized under the same conditions. Four wavelengths of MAD data around the selenium edge were collected from a single frozen crystal on beamline F2 at the Cornell High Energy Synchrotron Source (CHESS). The selenium positions were determined using direct methods (Shake-and-Bake), and the experimental electron-density map phased at 2.5 Å (Table 1) was of high enough quality to build the complete FabH dimer with no missing residues (Figure 2). Iterative rounds of refinement and model

Figure 2



Stereoviews of the electron-density maps of FabH. **(a)** The solvent-flattened MAD map calculated at 2.5 Å that was used to build the initial model. **(b)** The $2F_o–F_c$ calculated phased map generated from the final refined coordinates at 1.8 Å. The area shown in all panels is the active site of monomer A centered on Cys112 that projects from the top. On the right is Phe87′ that forms the floor of the oxyanion hole. On the left are His244 and Asn274, and at the bottom is the end of the CoA molecule including the terminal sulfur atom. Also shown are three important water molecules, one in the oxyanion hole (between Cys112 and Phe87′) and two at the top left that orient His244. Both maps are contoured at 1.0σ and displayed using the 'O' program [45].

building produced the final 1.8 Å structure with excellent statistics (Table 2). A total of 704 water molecules were fitted into the final $2F_o–F_c$ map (Figure 2). 'Tubes' of less well defined electron density in the two active-site tunnels of the dimer were resolved into CoA molecules.

The excellent anomalous signal from the selenium absorption peak (λ3) and the high quality of the data (Table 1) encouraged us to perform the phasing using only two wavelengths, λ3 and λ1 (low energy remote). These phases (Table 1) produced the high quality map that was used in the model building. A more conventional four-wavelength MAD phasing was also performed but the improvement in the experimental map was marginal at best. Correlation coefficients between the two-wavelength and four-wavelength experimental maps, and the final $2F_o–F_c$ map were 0.61 and 0.63, respectively.

Description of the structure

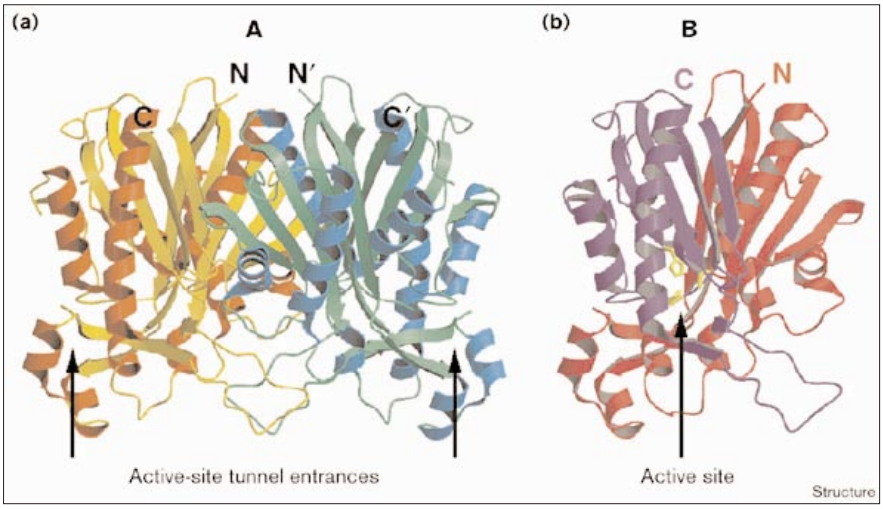
FabH is a dimeric, slab-shaped protein with approximate dimensions 45 Å × 55 Å × 80 Å (Figure 3). As will be discussed later, the overall architecture of the molecule is very similar to that of three other known protein structures, and the description will be brief. As viewed in Figure 3, the upper two thirds of the FabH monomer is a highly structured α/β motif, and the lower third is comprised of three extended loop regions. The α/β motif contains a clear internal duplication of two similar parts (1–171 and 172–317), and each part contains eight secondary structural elements in the order β1–α1–β2–α2–β3–α3–β4–β5. The two parts can be optimally superimposed by considering 58 equivalent residues within the β strands and α helices (root mean square deviation [rmsd] of 2.27 Å on α carbons). To reflect the duplication, the matching secondary structure elements within each part are labeled N or C and have the same numbers (Figure 4). The five β strands create a mixed β sheet; helices α1 and α2 are packed on one side

Table 2

Statistics of the final model.

Resolution range (Å)	20.0–1.8
Sigma cut-off applied	0.0
Number of reflections used in refinement	61,640
Percentage of reflections used in R_{free}	10.0
Completeness of data in resolution range (%)	97.2
Number of protein atoms	4782
Number of CoA atoms	94
Number of water molecules	704
R factor (%)	18.7
R_{free} factor (%)	22.5
Rms deviations from ideal stereochemistry	
Bond lengths (Å)	0.011
Bond angles (°)	1.8
Mean B factor (mainchain) (Å^3)	14.42
Rms deviation in mainchain B factor (Å^3)	0.733
Mean B factor (sidechains and waters) (Å^3)	22.91
Rms deviation in sidechain B factors (Å^3)	1.018
Ramachandran plot	
Residues in most favored region (%)	92.2
Residues in additionally allowed regions (%)	7.0
Residues in generously allowed regions (%)	0.7
Residues in disallowed regions (%)	0.0

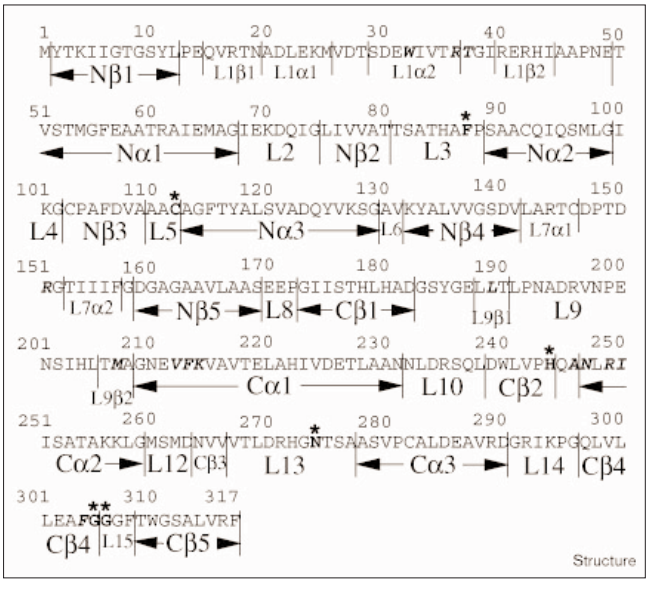
Figure 3



The overall structure of the FabH molecule. **(a)** Ribbon representation of the dimer showing the elements of secondary structure. The N and C termini and the active-site tunnel entrances are labeled. **(b)** A ribbon view of the FabH monomer colored to emphasize the two repeated substructures. The three active-site residues, Cys112, His244 and Asn274, are shown in yellow at the center of the molecule. The N and C termini are labeled and colored appropriately. The figure was produced using MOLSCRIPT [48].

and helix α3 is on the opposite side. Within the monomer, the two substructures are related by a pseudo-twofold axis, which brings the two α3 helices together at the center of the molecule with the α1/α2 pair at the periphery. The three extended loops in the lower part of the monomer (L1, L7 and L9) contain additional elements of secondary structure, and their labels reflect that they are outside of the repeated α/β motif (Figure 4).

Figure 4



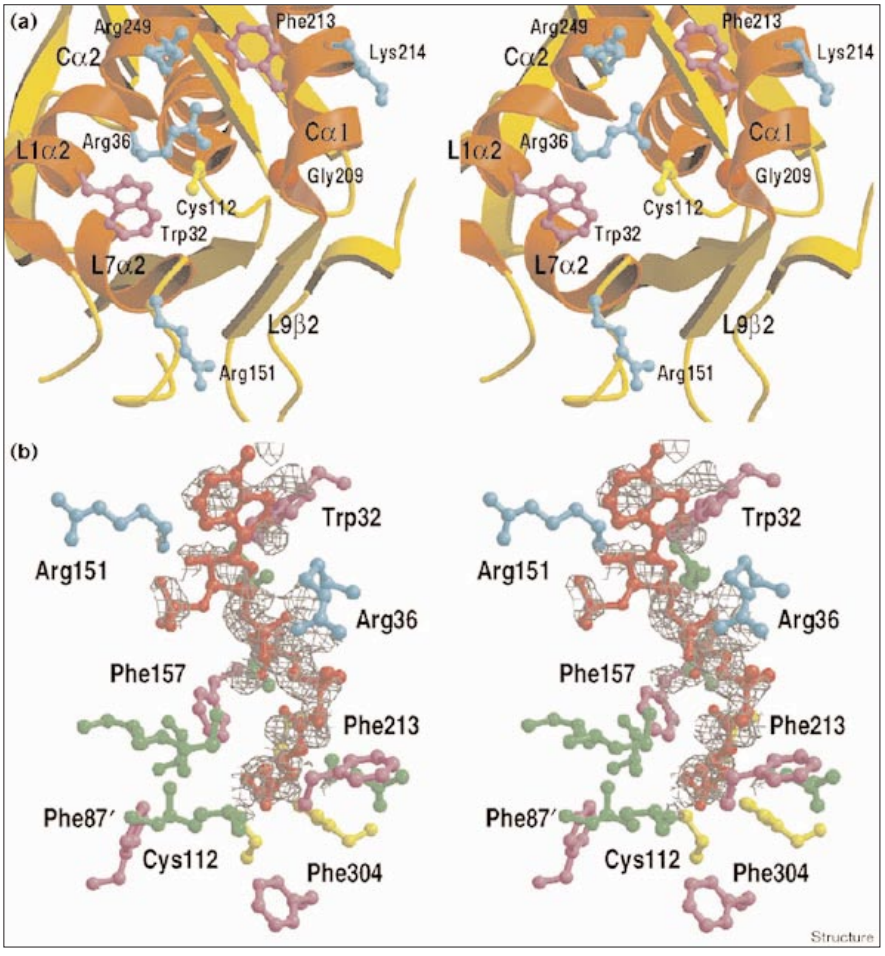
The amino acid sequence and secondary structure elements of FabH. The nomenclature used to label the β strands and α helices emphasizes the internal structural repeat in the protein. Corresponding elements in the two halves have the same label, and each has the prefix N or C. The loops are numbered sequentially, and additional β strands and α helices within the loops that are additions to the repeated substructures are indicated. Important conserved residues at the active site are in bold and marked with an asterisk. Conserved residues that form the active-site tunnel and mediate CoA binding are in bold and italicized. Conserved residues reflect amino acids that are identical in seven FabH protein sequences analyzed from bacteria and plants.

The monomer–monomer interface within the dimer buries 3222 Å^2 of surface area and is primarily composed of four elements. The first involves the outer strand of the N-sheet (Nβ3) that associates with its dimeric partner across the interface in an antiparallel fashion to create a ten-stranded β sheet that traverses through the center of the entire molecule. The second element is loop L9 that reaches across to the other monomer, and the third is helix Nα3 that makes complementary interactions with its dimeric partner across the twofold axis towards the 'top' of the molecule. The final interface element involves residues 84–103 and is centered on helix Nα2 (Figure 4).

The active site

Cys112 provides the essential thiol group in the enzymatic reaction of condensing enzymes [12,21,25]. In FabH, Cys112 is positioned at the bottom of a tunnel measuring 20 Å deep and 5 Å in diameter. Four α-helical elements, L7α2, the N termini of Cα1 and Cα2, and the C terminus of L1α2, together with β strand L9β2, form the entrance to the cavity (Figure 5a). Four conserved basic residues (Arg36, Arg151, Lys214 and Arg249) and two conserved aromatic residues (Trp32 and Phe213) flank the entrance to the active site. This surface, therefore, has mixed hydrophobic and electropositive character, which sharply contrasts with the majority of the FabH surface that is generally electronegative. Gly209 is completely conserved because the presence of a sidechain would obstruct access to the active site. Conserved hydrophobic residues (Figure 4) line the inner surface of the tunnel. Finally, three completely conserved residues (Cys112, His244 and

Figure 5



The active-site tunnel of FabH. **(a)** A stereoview showing a close-up of the tunnel entrance and the associated conserved residues. The coloring of the ribbon diagram corresponds to Figure 3a. The tunnel entrance is formed by four α helices, L1α2, L7α2, Cα1 and Cα2 and one β strand, L9β2. Four basic residues shown in blue (Arg36, Arg151, Arg249 and Lys214) surround the entrance, and two aromatic residues shown in magenta are on either side (Trp32 and Phe213). Note that the conserved Gly209, shown as a red sphere, allows access to the tunnel. Cys112 in the active site lies at the bottom of the tunnel in yellow, and is positioned precisely at the N terminus of helix Nα3. The figure was produced using MOLSCRIPT [48]. **(b)** A stereoview showing the electron density (gray) and the fitted CoA molecule (red) in the active-site tunnel of monomer A. The electron density is contoured at 0.85σ, slightly below the protein density shown in Figure 2. There is no obvious density for the adenine ring, and its position is arbitrary. A number of important protein sidechains are shown. The three active-site residues at the base of the tunnel including Cys112 are shown in yellow, and the conserved hydrophobic and phenylalanine residues that line the tunnel are shown in green and magenta, respectively. The tunnel provides a suitable hydrophobic environment for the pantetheine moiety, however, and the conserved residues Arg36 (blue), Arg151 (blue) and Trp32 (magenta) are available for salt bridge and stacking interactions with the 3′,5′-ADP moiety. The figure was produced using Bobscript [49].

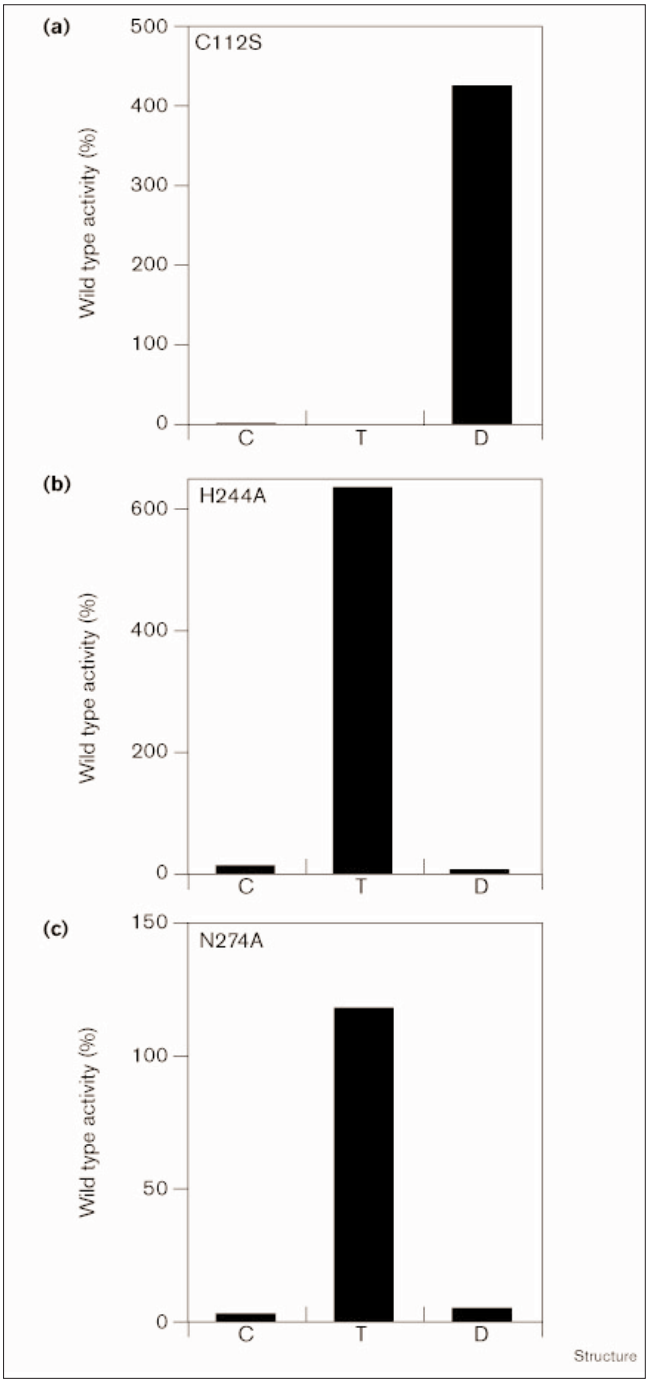
Asn274) lie at the bottom of the tunnel forming the active-site constellation. The deep active-site tunnel is principally created by the extended loop regions and this unusual feature is stabilized by interactions involving a number of conserved, charged and polar residues. These include aspartate residues 30, 107, 140 and 159, arginine residues 40, 42 and 144, Asn19, Gln245 and Ser279. A number of water molecules also have important structural roles in this region.

After refinement, the non-protein electron density remaining in the active-site tunnel is interpreted as bound CoA with partial occupancy (Figure 5b). The CoA density extends from the terminal sulfur atom to the ribose sugar ring (Figure 5b). The conserved hydrophobic residues in the tunnel that contact 4′-phosphopantetheine are shown in Figures 4 and 5b. Thr37 and Asn247 are also highly conserved in the tunnel and might have a role in binding. The phosphate groups of CoA form salt-bridge interactions with the conserved Arg36 and possibly Arg151, which places the adenine group in a position to allow stacking contact with the conserved Trp32 (Figure 5b). The adenine is not visible in the electron density but its potential interaction with Trp32 appears to be compromised by symmetry-related molecules in the unit cell.

Site-directed mutagenesis of active-site residues

The roles of the conserved active-site residues in the enzyme mechanism were investigated by constructing three FabH mutants, FabH(C112S), FabH(H244A) and FabH(N274A) (mutants named using single-letter amino acid code). Each mutant protein was analyzed by circular dichroism to verify that the spectrum matched that of the native protein (data not shown). The mutants were assayed for their ability to carry out the overall condensation reaction with acetyl-CoA and malonyl-ACP. The activity of each was compared with the wild-type protein, which has a specific activity of 739 ± 30 nmol/min/μg. All of the mutants were catalytically impaired, confirming their identification as important elements of the active site. The FabH(C112S) protein had an activity that was 0.1% of the wild-type activity (Figure 6a), FabH(H244A)

Figure 6



Catalytic activities of the FabH mutants. **(a)** FabH(C112S). **(b)** FabH(H244A). **(c)** FabH(N274A). The FabH proteins were assayed for the overall condensation reaction, and the transacylation and decarboxylation half reactions as described in the Materials and methods section. For each reaction, the data are presented as a percentage of the rate of the wild-type enzyme determined in the same assay. Specific activities of wild-type FabH in the three assay were: condensation, 739 ± 30 nmol/min/μg; transacylation, 4.1 ± 0.15 nmol/min/μg; and decarboxylation, 40 ± 0.45 nmol/min/μg.

had 7.7% activity (Figure 6b) and FabH(N274A) had 1.7% activity (Figure 6c).

The roles for these residues in the overall mechanism were defined by assaying the mutants for the partial reactions of FabH (Figure 1). The transacylation step was assayed by following the rate of transfer of an acetyl group from a CoA donor to an ACP acceptor via the acetyl enzyme. The wild-type protein had a specific activity of 4.1 ± 0.15 nmol/min/μg under the standard conditions established for this assay. The transacylase activity of FabH(C112S) was below the level of detection in this assay (<0.05 nmol/min/μg) (Figure 6a). The FabH(H244A) and FabH(N274A) mutants remained competent for transacylation, actually exhibiting increased rates of 634% and 118%, respectively, of the wild-type enzyme (Figures 6b,c). The decarboxylation half reaction was assayed by measuring the conversion of [2-^{14}C]malonyl-ACP to [^{14}C]acetyl-ACP (in the absence of acetyl-CoA). The wild-type enzyme had a specific activity of 40 ± 2.0 nmol/min/μg under the standard malonyl-ACP decarboxylase assay conditions. FabH(C112S) had a significantly increased decarboxylase activity; 425% higher than wild-type (Figure 6a). Both FabH(H244A) and FabH(N274A) were severely impaired, exhibiting <5% of the wild-type rate (Figures 6b,c). All three residues, therefore, were required for the overall condensation reaction but His244 and Asn274 were specifically required for decarboxylation and Cys112 was essential for transacylation.

Discussion

Enzyme mechanism: the transacylation step

The generation of an Sγ thiolate anion on Cys112 is required to attack the acetyl-CoA substrate to form the acetyl-FabH intermediate. Thus, the active-site environment would be predicted to decrease the pK_a of Cys112 for the enzyme to function at physiological pH. In other enzymes that form acyl-enzyme intermediates, adjacent histidine residues often promote the unusual reactivity of cysteine residues. For example, His159 in papain is 3.01 Å from Cys25 and enhances the nucleophilicity of the cysteine thiol [27]. In our structure, we also observe a histidine residue (His244) adjacent to Cys112. In FabH, the distance between the Nε of His244 and the Sγ of Cys112 is either 3.58 Å or 4.06 Å because of slight differences in the position of the Sγ in the two monomers. The FabH histidine, therefore, would not be as effective in lowering the pK_a of cysteine as the histidines in other proteins. This conclusion is supported by the fact that the transacylase activity remains robust in the FabH(H244A). A supportive role for His244, however, cannot be ruled out. The pH optima of FabH and FabH(H244A) were between pH 7.5 and pH 8.0, but FabH was relatively more active than FabH(H244A) at pH 6.0 in the transacylase assay (data not shown).

An alternative explanation for the unusual reactivity of Cys112 is suggested by the location of this residue at the N terminus of Nα3. We attribute the enhanced nucleophilicity of the Sγ of Cys112 to the α-helix dipole effect [28]. The tight turn of β strand Nβ3 into helix Nα3 (L5) is stabilized by hydrogen bonds between the backbone carbonyl of Ala111 and the amide of Gly114 causing the Sγ of Cys112 to be ideally positioned to receive the full benefit of a half unit of positive charge generated by the strong dipole of Nα3 (Figure 7a). Experiments with model peptides have shown that the location of cysteine at the N terminus of an α helix results in a ΔpK_a of –1.6, thus lowering the pK_a of the thiol from an unperturbed pK_a of 8.8 to a pK_a of 7.2 [29]. In the case of FabH, the helix–dipole interaction with Cys112 would generate a strong enough nucleophile to catalyze the thioester exchange reaction responsible for the formation of the acetyl-FabH intermediate.

Enzyme mechanism: the oxyanion hole

Another important component of the FabH active site is an arrangement of hydrogen-bond donors (termed an 'oxyanion hole') [30] that accommodates the negatively charged carbonyl oxygen in the tetrahedral transition state arising from reactions with the acetyl-enzyme intermediate. This feature in FabH encompasses the backbone amide groups of Cys112 and Gly306, the cavity formed by Gly305 and the sidechain of Phe87′ from the opposite monomer (Figure 7a). Phe87′ represents the only active-site communication between the two halves of the dimer. Model building suggests that the tetrahedral intermediate predicted to arise from reactions involving the acetyl-FabH intermediate would be accommodated in this cavity, with the oxyanion forming hydrogen-bonding interactions with the amide nitrogens. Phe87 is highly conserved in FabH enzymes, and it is possible that it also has a role in stabilizing the oxyanion transition state. This location for the oxyanion hole is supported by electron density within the cavity (Figure 2). The density appears to be principally a water molecule making hydrogen-bonding interactions with the two amide nitrogens, but the shape does suggest that a low occupancy acetyl group might also be present. This suggests that the FabH-CoA complex contains a small population of the acetyl-FabH intermediate in the crystal with the CoA still bound.

Gly305 and Gly306 are part of a conserved glycine-rich sequence within loop L15 that is fixed by a hydrogen-bonding network involving structural water molecules. The absence of sidechains in this tight turn contributes to the formation of the oxyanion cavity. The amide nitrogens forming the FabH oxyanion hole have the same overall geometry as the classical oxyanion cavity in serine proteases [30], although in FabH the architecture is most similar to the structure of the oxyanion hole in acetylcholinesterase in which a Gly–Gly–Gly loop from another strand of the protein contributes the second amide nitrogen to the oxyanion hole [31]. The oxyanion hole is predicted to be important in the deacetylation of the enzyme during the transacylation reaction and in the attack on the acetyl-FabH intermediate by the carbanion that forms on the 2-carbon of malonate in the condensation reaction (Figure 7c).

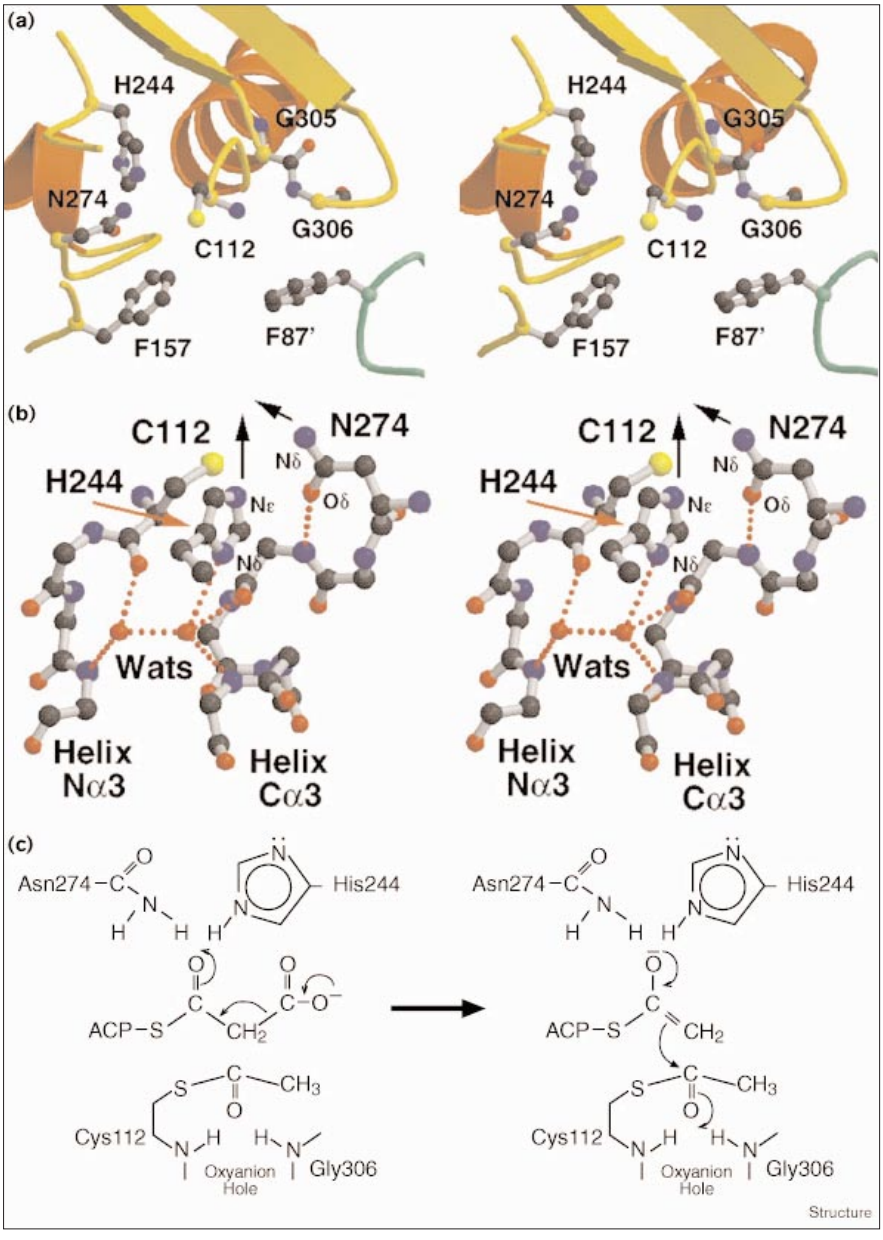
Enzyme mechanism: condensation reaction

The mutagenesis experiments show clearly that His244 and Asn274 are key active-site residues required for the condensation reaction, and are specifically involved in the decarboxylase half reaction (Figure 6). The orientations of both sidechains are precisely fixed by hydrogen-bonding interactions. As shown in Figure 7b, the Nδ of His244 is part of a water-mediated hydrogen-bonding cluster, and the sidechain Oδ of Asn274 is hydrogen bonded back to the mainchain amide of Ser276. The result of these interactions is that the Nε of His244 and the Nδ of Asn274 are 3.8 Å apart and point towards the same space in the active site (Figure 7b). Modeling a malonate group into the active site at the end of the 4′-phosphopantetheine suggests that the carbonyl group of the malonyl thioester would interact directly with this pair of adjacent nitrogen atoms (data not shown). The proposed FabH condensation mechanism is shown in Figure 7c. Asn274 and His244 interact with the incoming malonyl-ACP to facilitate the enolization of the substrate and the formation of a carbanion on the C2 of malonate by stabilizing the developing negative charge on the thioester carbonyl oxygen. This carbanion attacks the acetyl group that is attached to the Sγ of Cys112 and the ensuing tetrahedral transition state is stabilized by hydrogen-bond interactions with the amide nitrogen atoms of Cys112 and Gly306, which constitute the oxyanion hole. This transition state breaks down to form acetoacetyl-ACP, which is released from the enzyme. In the absence of an acetyl-FabH intermediate, the carbanion developing at the C2 of malonate picks up a proton from either Cys112 or the solvent leading to decarboxylation and the release of acetyl-ACP.

Enzyme mechanism: evolutionary implications

An intriguing feature of the active site is that it spans equivalent locations of the repeated substructures. Cys112 is at the N terminus of helix Nα3 in one substructure, and Asn274 is close the N terminus of the equivalent helix Cα3 in the other substructure. In the dimer, the N termini of these α helices are adjacent at the central pseudo-twofold axis, and create the active site at the center of the molecule (Figure 3b). This location suggests how the two-step enzyme mechanism might have evolved. Before the duplication event, the original protein might have had a primitive active site at the N terminus of helix α3 activated by the helix dipole. The duplication would then have generated two adjacent active sites that could independently evolve to perform the two-step enzyme reaction. This concept is supported by the observation that

Figure 7



The active-site region of FabH. **(a)** Stereoview showing the essential elements of the active site including the catalytically important residues (Cys112, His244 and Asn274), and the N termini of α helices Nα3 and Cα3. The depiction of the mainchain matches that shown in Figure 3a. The oxyanion hole encompasses the amide groups of Cys112 and Gly306, the cavity created by Gly305 and the aromatic ring of Phe87′ (from the opposite monomer). Phe157 is located in a position to promote decarboxylation of malonate. **(b)** Stereoview of the hydrogen-bonding network that fixes the orientations of the active-site residues His244 and Asn274. The Nδ of His244 is part of a network that includes two structural waters that span the N termini of α helices Nα3 (the carbonyl of Cys112 and the amide of Phe115) and Cα3 (the carbonyl of Ser276 and the amide of Val280). The Oδ of Asn274 forms a hydrogen bond to the amide of Ser276. The two black arrows illustrate that the Nε of His244 and the Nδ of Asn274 point to the same space. **(c)** Proposed mechanism for the condensation reaction. Left: Asn274 and His244 interact with the thioester carbonyl oxygen of malonyl-ACP to facilitate enolization of the substrate and the formation of a carbanion on the C2 of malonate. Right: This carbanion attacks the acetyl-enzyme intermediate attached to the Sγ of Cys112 and the ensuing tetrahedral transition state is stabilized by hydrogen-bond interactions with the amide nitrogens of Cys112 and Gly306 in the oxyanion hole. Panels (a) and (b) were produced using MOLSCRIPT [48].

Cys112 is essential for the transacylation half reaction whereas His244 and Asn274 are required for the decarboxylation half reaction.

Structural comparisons of FabH with related enzymes

The FabH coordinates were submitted to the Dali server [32], and the overall structural organization was found to be very similar to thiolase (β-ketoacyl-CoA thiolase) [33], chalcone synthase [26] and FabF (β-ketoacyl-ACP synthase II) [25]. The monomers all contain the FabH repeated motif in almost identical orientations, and they associate to form dimers in the same fashion. The only difference in the motif occurs in chalcone synthase in which strand Cβ3 is replaced by an α helix. Outside of the conserved motif, the four proteins show considerable differences in the loop regions at the bottom of the molecule. Thiolase has few extended loops apart from residues 155–276 that are equivalent to L7 in FabH. This loop extends around the molecule, including the dimer interface, and contains a number of secondary structure elements. Chalcone synthase resembles FabH in that it has three loops equivalent to L1, L7 and L9, but their conformations are quite different. FabF has four loop regions equivalent to L1, L3, L7 and L14, again with distinct conformations.

Optimal alignment of the four proteins showed that the active sites have similar locations at the N termini of Nα3 and Cα3. In particular, the critical cysteine residues align very closely and are invariably positioned at the N terminus of Nα3. Reaction mechanisms have been proposed for all three of the homologous enzymes but none of the schemes have been experimentally tested or incorporate the α-helix dipole that we believe to be a key contributor to the nucleophilicity of Cys112. This is particularly important for FabF, which performs a very similar reaction to FabH. In the FabF active site His303 aligns with His244 in FabH and was proposed to be involved in decarboxylation as we have suggested (Figure 7c). His340 replaces Asn274, however, and was speculated to be involved in the acyl-enzyme formation by interacting with the Sγ of Cys163 to increase its reactivity. Our scheme predicts that His340 is an active participant in the decarboxylation half reaction. The function of the dual histidine residues in the FabF family of condensing enzymes, therefore, requires experimental verification by biochemical analysis of the appropriate mutants. The active site of chalcone synthase is virtually identical to that of FabH, with the active-site cysteine, histidine and asparagine residues occupying the identical space. Ferrer *et al.* [26] also suggest that Phe215 of chalcone synthase might participate in decarboxylation by promoting the conversion of the negatively charged carboxyl group of malonate to a neutral carbon dioxide molecule. Phe215 of chalcone synthase superimposes on the conserved Phe157 in FabH suggesting that this residue might play a similar role in the decarboxylation reaction of FabH (Figure 7a).

It might be significant that all four proteins have extended, but unique, loop structures at the bottom of the molecule, but none at the top. A gene duplication event would have left the active site completely exposed at the bottom of the molecule. The subsequent addition of loops would have sequestered the active site from the exterior, and would have also allowed the evolution of appropriate entrances for different hydrophobic substrates. Thus FabH and chalcone synthase have tunnels, FabF has a closed pocket, and thiolase has an open pocket.

Biological implications

β-Ketoacyl-acyl carrier protein synthase III (FabH) catalyzes the first condensation step in the type II fatty acid synthases found in bacteria and plants. Analogous enzymes are responsible for the condensation reactions in the formation of chalcones and polyketides. FabH is ideally positioned to be a key regulator of the pathway and in bacteria is thought to be regulated by feedback inhibition by the long-chain acyl-ACP end products [11,18]. The FabH structure is similar to that of FabF [25] and chalcone synthase [26] and all three structures define a family of proteins that are closely related to the thiolase structure [33]. The functions of the active-site residues are defined. The histidine and asparagine residues are both required for the decarboxylation step, a helix dipole activates the cysteine, and an oxyanion hole is identified that promotes reactions of the acetyl-enzyme intermediate. The combined structural and site-directed mutagenesis analysis identify the specific locations and functions of the active-site residues in the condensation reaction that will serve as a paradigm for understanding this class of enzymes.

The FabH structure is relevant to the development of new antibacterials. Agents that inhibit fatty acid biosynthesis have proven to be useful therapeutics. Cerulenin inhibits both FabB- and FabF-condensing enzymes by covalent modification of the active-site sulfhydryl [21,22], and all the condensing enzymes are inhibited by the broad-spectrum antibiotic thiolactomycin [23,24]. Isoniazid is an established therapy for tuberculosis that is effective against the enoyl-ACP reductase of the type II system in this organism [3,19]. Recently, the broad-spectrum antibacterial triclosan was also demonstrated to specifically target fatty acid synthesis at the enoyl-ACP reductase step [4,5,20]. The ubiquitous expression indicates an essential function for FabH and makes it an ideal candidate target for the identification of novel antibacterial agents. The high-resolution FabH structure will be essential to this research.

Materials and methods

Protein expression

The wild-type *fabH* gene from *E. coli* was previously cloned into the vector pET-15b (Novagen), forming pET-fabH, for the expression and purification of N-terminally His-tagged FabH from *E. coli* strain BL21(DE3) [11,18]. Lysis and purification of the His-tagged protein was accomplished as described previously [34]. For the preparation of selonomethionyl (SeMet) FabH, the pET-fabH plasmid was transformed into the methionine auxotroph, strain B834(DE3) (Novagen). Growth was performed in a defined medium of M9 salts [35] containing 0.4% glucose, 0.0005% thiamine, 0.1 μg/ml each amino acid and 100 μg/ml ampicillin. Cells were grown to a density of 5×10^8 cells/ml and resuspended in medium lacking methionine. Selenomethionine (0.05 μg/ml; Sigma Chemical Co.) was then added and expression induced. The incubation was continued for 3 h and then cells were harvested, lysed and the SeMet FabH protein purified.

Protein crystallization

Crystals of FabH were grown by the hanging-drop vapor diffusion method using a protein solution containing 15 mg/ml of FabH in 20 mM Tris pH 7.4, 1 mM EDTA and 1 mM DTT. Optimal crystallization conditions were 1.8–2.0 M ammonium sulfate, 2% PEG 400 and 100 mM HEPES pH 7.5. Crystals grew in two forms, both in space group $P2_12_12_1$. One form had cell dimensions a = 64.5 Å, b = 80.7 Å and c = 123.8 Å, and the other had cell dimensions a = 63.1 Å, b = 64.3 Å and c = 165.7 Å. The latter crystals were considerably larger (typically $0.5 \times 0.5 \times 0.2$ mm) diffracted to higher resolution and were used for structure determination. The Matthews coefficient or Vm [36] of 2.5 Å^3/Da (corresponding to an estimated solvent content of 50.5%) indicated one dimer per asymmetric unit. SeMet crystals were grown under the same conditions.

Data collection

Multiwavelength anomalous dispersion (MAD) measurements were performed at the Cornell High Energy Synchrotron Source (CHESS) on beamline F2 that incorporates an ADSC Quantum 4 CCD detector. Data were collected at 100K from a single SeMet FabH crystal that was flash-frozen by first passing it through stabilization buffers (2.0 M ammonium sulphate, 2% PEG 400 and 100 mM HEPES pH 7.5) containing 5, 10, 15 and 20% glycerol and then plunging it into liquid nitrogen. Four wavelengths were collected around the selenium absorption edge; $\lambda1 = 1.000$ Å (low energy remote), $\lambda2 = 0.9793$ Å (inflection point), $\lambda3 = 0.9789$ Å (absorption peak) and $\lambda4 = 0.9750$ Å (high energy remote). The data at $\lambda1$ were collected in a single sweep of $122 \times 1°$ frames using an exposure time of 30 seconds per degree. The data at $\lambda2$, $\lambda3$ and $\lambda4$ were collected using inversion beam geometry to measure accurately the anomalous signal. To minimize systematic errors, pairs of 110° sweeps separated by 180° were collected in alternating wedges of $10 \times 1°$ frames to produce a total of 220° of data. The exposure time for these measurements was 15 seconds per degree. All the MAD data were collected with a crystal-to-detector distance of 200 mm. Finally, for model refinement purposes, a high-resolution data set was collected from the same crystal. The detector was moved closer, to a distance of 150.0 mm, and $115 \times 1°$ frames were collected using an exposure time of 90 seconds per degree. In all cases, the oscillation data were integrated using MOSFLM [37] and scaled with SCALA [38].

Structure determination

The $\lambda3$ data, which contain the highest anomalous signal, were first used to generate anomalous difference *E*-magnitudes to 3.0 Å resolution using the DREAR program suite [39]. These *E*-values were then used to solve the selenium substructure using the direct methods program Shake-and-Bake (SnB V1.5) [40]. The 240 largest *E*-values generated 1538 triplet invariants. Trial structures starting from 16 randomly positioned selenium atoms (i.e. 2×8 selenomethionine residues) were then subjected to 48 cycles of phase refinement. A total of 137 trials were processed. The resulting bimodal nature of the minimal function [41] clearly identified 13 solutions (with $R_{min} = 0.257$) compared with the rest of the trials (mean $R_{min} = 0.513$). The best six solutions contained an identical set of 12 peaks (usually the top 12). Of the four remaining expected peaks, two corresponded to N-terminal methionines with weak density, and the other two, in retrospect, had been identified by SnB but did not appear consistently in the top 12 peaks. Maximum likelihood heavy-atom refinement and phasing calculations were conducted with MLPHARE [42] using the 12 positions in conjunction with the $\lambda1$ and $\lambda3$ data to 2.5 Å resolution. Data obtained using the $\lambda1$ wavelength (low energy remote) was treated as the 'native'. Inclusion of the data from wavelengths $\lambda2$ and $\lambda4$ had very little effect on the final electron-density maps. Using the program DM [43], phases were improved by solvent flattening and histogram matching.

Model building and refinement

An experimental electron-density map was calculated using CCP4 programs [44] and displayed using the O program [45]. The map was easily interpreted to build both monomers, including sidechains. Using X-PLOR the initial model was refined at 3.0 Å and included several cycles of rigid-body refinement. Thereafter, the model was improved by alternating rounds of refinement against increasingly higher resolution data followed by manual revision using O. In later rounds, water and CoA molecules were included. A final round of refinement was performed using REFMAC [46]. The stereochemistry of the final model was evaluated using PROCHECK [47].

Construction of FabH mutants

Mutations were made in the *fabH* gene in pET-fabH utilizing an overlap polymerase chain reaction (PCR) approach. To mutate the Cys112 to serine, the outside primers were fabH-A (5′-TATCCGCGAACGCCACA-3′) and fabH-D (5′-TTCAGAGGCAGCCAGCA-3′), with the internal primers fabH-B (5′-GCGGAGGCTGCTGCAACG-3′) and fabH-C (5′-GCAGCAGCCTCCGCAGGT-3′). The 397 bp product of this reaction was purified and ligated in the TA cloning vector, pCR2.1 (Invitrogen). A plasmid with the correct mutation-containing sequence was digested with *Mfe*I and *Nsi*I and ligated into pETfabH digested with the same enzymes. The resultant plasmid was sequenced and then transformed into *E. coli* BL21(DE3). The H244A and A274A mutants of FabH were constructed using the outside primers H–Nsi–For (5′-CATTTGACGTTGCAGCAGCC-3′) and H–Bam–Rev (5′-TGCCGGATCCAATTGCGTCATGT-3′). The mutagenic primers were H–H244A–Rev (5′-AGCCTGAGCCGGAACCAGC-3′) and H-H244A-For (5′-GGTTCCGGCTCAGGCTAA-3′); H–N274A–Rev (5′-AGG-TAGCACCGTGGCGATCC-3′) and H–N274A–For (5′-CGCCACGGTGCTACCTCTGC-3′). First round PCRs contained H–Nsi–For plus a mutagenic reverse primer, or H–Bam–Rev plus a mutagenic forward primer. Resultant plasmids were sequenced to ensure the presence of only the desired mutation and then transformed into strain BL21(DE3) as above.

Circular dichroism of FabH mutants

Proteins were diluted into 0.1 M phosphate buffer, pH 7.0, and quantitated by absorbance at 280 nm (the molar absorbance of His-tagged FabH was taken to be 27493 $(M \times cm)^{-1}$ from the Biopolymer calculator at http://paris.chem.yale.edu) immediately prior to measuring the circular dichroism spectrum. Spectra were recorded on an AVIV Instruments Model 62ADS spectrometer scanning from 195–220 nm at 24°C.

Assays of FabH proteins

The wild-type and mutant FabH proteins were assayed for condensation (β-ketobutyryl-ACP synthase) activities according to [11,12]. Briefly, 40 µl reactions were set up in 0.1 M sodium phosphate, pH 7.0, containing 25 µM ACP, 1 mM 2-mercaptoethanol, 65 µM malonyl-CoA, 45 µM acetyl-CoA (specific activity, 59 mCi/mmol; American Radiolabeled Chemicals, Inc.) and 0.2 µg FabD. The final concentration of protein was varied according to the mutation and the particular assay being performed, such that the rate, if any, was in the linear range. Wild-type FabH was assayed in the range of 2–10 ng protein per assay. Reactions were incubated at 37°C for 15 min, and stopped by pipetting 35 µl of the mixture onto a Whatman 3MM filter paper disc. Discs were then washed with successively lower concentrations of ice-cold tricarboxylic acid (10%, 5%, 1%; 20 min per wash with 20 ml per disc), dried and counted. Transacylation (acetyl-CoA:ACP transacylase) assays omitted malonyl-CoA and FabD and reactions were incubated for 30 min. For the decarboxylation assays, acetyl-CoA was omitted, and the cold malonyl-CoA was replaced with 50 µM of [^{14}C]malonyl-CoA (specific activity, 56 mCi/mmol; American Radiolabeled Chemicals, Inc.). Following incubation at 37°C for 30 min, the reactions were placed in an ice slurry, 10 µl gel loading buffer was added, and the entire mixture was then loaded onto a conformationally sensitive 0.5 M urea, 13% polyacylamide gel. Following electrophoresis, the gel was fixed, dried and the radioactivity detected and quantitated by exposure to a phosphor storage screen calibrated with a standard curve of [^{14}C]malonyl-CoA.

Accession numbers

The coordinates of FabH have been deposited in the Brookhaven Protein Data Bank with accession number 1EBL.

Acknowledgements

We are particularly grateful to Ashley Deacon for help in processing the CHESS data and also for assisting in the Shake-and-Bake calculations. We also thank Hee-Won Park and other colleagues for invaluable discussions. R Brent Calder, Magda Kaminska, Amy Sullivan, Hee-Won Park, Allen Price and Charles Ross II provided excellent technical assistance. This work was funded by National Institutes of Heath Grants GM 34496 (COR), GM 44973 (SW), Cancer Center (CORE) Support Grant CA 21765, and the American Lebanese Syrian Associated Charities. The research at CHESS is supported by grant DMR-9311772 from the National Science Foundation and the MacCHESS research resource is funded by grant RR-01646 from the National Institutes of Health.

References

1. Rock, C.O. & Cronan, J.E., Jr. (1996). *Escherichia coli* as a model for

the regulation of dissociable (type II) fatty acid biosynthesis. *Biochim. Biophys. Acta* **1302**, 1-16.
2. Ohlrogge, J.B. (1982). Fatty acid synthetase: plants and bacteria have similar organization. *TIBS* **7**, 386-387.
3. Banerjee, A., *et al.* & Jacobs, W.R., Jr. (1994). *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science* **263**, 227-230.
4. McMurray, L.M., Oethinger, M. & Levy, S. (1998). Triclosan targets lipid synthesis. *Nature* **394**, 531-532.
5. Heath, R.J., Yu, Y.-T., Shapiro, M.A., Olson, E. & Rock, C.O. (1998). Broad spectrum antimicrobial biocides target the FabI component of fatty acid synthesis. *J. Biol. Chem.* **273**, 30316-30321.
6. Verwoert, I.I.G.S., van der Linden, K.H., Walsh, M.C., Nijkamp, H.J.J. & Stuitje, A.R. (1995). Modification of *Brassica napus* seed oil by expression of the *Escherichia coli fabH* gene, encoding 3-ketoacyl-acyl carrier protein synthase III. *Plant Mol. Biol.* **27**, 875-886.
7. Waller, R.F., *et al.* & McFadden, G.I. (1998). Nuclear-endcoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. *Proc. Natl Acad. Sci. USA* **95**, 12352-12357.
8. Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J. & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675-679.
9. Hopwood, D.A., Chater, K.F. & Bibb, M.J. (1995). Genetics of antibiotic production in *Streptomyces coelicolor* A3(2), a model streptomycete. *Biotechnology* **28** , 65-102.
10. Martin, C.R. (1993). Structure, function and regulation of chalcone synthase. *Int. Rev. Cytol.* **147**, 233-284.
11. Heath, R.J. & Rock, C.O. (1996). Inhibition of β-ketoacyl-acyl carrier protein synthase III (FabH) by acyl–acyl carrier protein in *Escherichia coli*. *J. Biol. Chem.* **271**, 10996-11000.
12. Tsay, J.-T., Oh, W., Larson, T.J., Jackowski, S. & Rock, C.O. (1992). Isolation and characterization of the β-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. *J. Biol. Chem.* **267**, 6807-6814.
13. Rock, C.O. & Jackowski, S. (1982). Regulation of phospholipid synthesis in *Escherichia coli*. Composition of the acyl–acyl carrier protein pool *in vivo*. *J. Biol. Chem.* **257**, 10759-10765.
14. Voelker, T.A. & Davies, H.M. (1994). Alteration of the specificity and regulation of fatty acid synthesis of *Escherichia coli* by expression of a plant medium-chain acyl-acyl carrier protein thioesterase. *J. Bacteriol.* **176**, 7320-7327.
15. Jiang, P. & Cronan, J.E., Jr. (1994). Inhibition of fatty acid synthesis in *Escherichia coli* in the absence of phospholipid synthesis and release of inhibition by thioesterase action. *J. Bacteriol.* **176**, 2814-2821.
16. Cho, H. & Cronan, J.E., Jr. (1995). Defective export of a periplasmic enzyme disrupts regulation of bacterial fatty acid synthesis. *J. Biol. Chem.* **270**, 4216-4219.
17. Heath, R.J., Jackowski, S. & Rock, C.O. (1994). Guanosine tetraphosphate inhibition of fatty acid and phospholipid synthesis in *Escherichia coli* is relieved by overexpression of glycerol-3-phosphate acyltransferase (plsB). *J. Biol. Chem.* **269**, 26584-26590.
18. Heath, R.J. & Rock, C.O. (1996). Regulation of fatty acid elongation and initiation by acyl–acyl carrier protein in *Escherichia coli*. *J. Biol. Chem.* **271**, 1833-1836.
19. Rozwarski, D., Grant, G., Barton, D., Jacobs, W. & Sacchettini, J.C. (1998). Modification of NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. *Science* **279**, 98-102.
20. Heath, R.J., Rubin, J.R., Holland, D.R., Zhang, E., Snow, M.E. & Rock, C.O. (1999). Mechanism of triclosan inhibition of bacterial fatty acid synthesis. *J. Biol. Chem.* **274**, 11110-11114.
21. Kauppinen, S., Siggaard-Anderson, M. & van Wettstein-Knowles, P. (1988). β-Ketoacyl-ACP synthase I of *Escherichia coli*: nucleotide sequence of the *fabB* gene and identification of the cerulenin binding residue. *Carlsberg. Res. Commun.* **53**, 357-370.
22. Moche, M., Schneider, G., Edwards, P., Dehesh, K. & Lindqvist, Y. (1999). Structure of the complex between the antibiotic cerulenin and its target, β-ketoacyl-acyl carrier protein synthase. *J. Biol. Chem.* **274**, 6031-6034.
23. Jackowski, S., Murphy, C.M., Cronan, J.E., Jr. & Rock, C.O. (1989). Acetoacetyl-acyl carrier protein synthase: a target for the antibiotic thiolactomycin. *J. Biol. Chem.* **264**, 7624-7629.
24. Tsay, J.-T., Rock, C.O. & Jackowski, S. (1992). Overproduction of β-ketoacyl-acyl carrier protein synthase I imparts thiolactomycin resistance to *Escherichia coli* K-12. *J. Bacteriol.* **174**, 508-513.
25. Huang, W., Jia, J., Edwards, P., Dehesh, K., Schneider, G. & Lindqvist, Y. (1998). Crystal structure of β-ketoacyl-acyl carrier protein synthase II from *E. coli* reveals the molecular architecture of condensing enzymes. *EMBO J.* **17**, 1183-1191.
26. Ferrer, J.-L., Jez, J.M., Bowman, M.E., Dixon, R.A. & Noel, J.P. (1999). Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nat. Struct. Biol.* **6**, 775-784.
27. Kamphuis, I.G., Kalk, K.H., Swarte, M.B.A. & Drenth, J. (1984). Structure of papain refined at 1·65 Å resolution. *J. Mol. Biol.* **179**, 233-256.
28. Hol, W.G.J., van Duijnen, P.T. & Berendsen, H.J.C. (1978). The α-helix dipole and the properties of proteins. *Nature* **273**, 443-446.
29. Kortemme, T. & Creighton, T.E. (1995). Ionisation of cysteine residues at the termini of model α-helical peptides. Relevance to unusual thiol pK_a values in proteins of the thioredoxin family. *J. Mol. Biol.* **253**, 799-812.
30. Kraut, J. (1977). Serine proteases: structure and mechanism of catalysis. *Annu. Rev. Biochem.* **46**, 331-358.
31. Ordentlich, A., *et al.* & Shafferman, A. (1999). Functional characteristics of the oxyanion hole in human acetylcholinesterase. *J. Biol. Chem.* **273**, 19509-19517.
32. Holm, L. & Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478-480.
33. Mathieu, M., *et al.* & Wierenga, R.K. (1997). The 1.8 Å crystal structure of the dimeric peroxisomal 3-ketoacyl-CoA thiolase of *Saccharomyces cerevisiae*: implications for substrate binding and reaction mechanism. *J. Mol. Biol.* **273**, 714-728.
34. Petty, K.J. (1994). Metal-chelate affinity chromatography. In *Current Protocols in Molecular Biology* (Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Strugh, K., eds), pp. 10.11.8-10.11.22. John Wiley & Sons, Inc., New York.
35. Miller, J.H. (1972). *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
36. Matthews, B.W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.
37. Leslie, A.G.W., Brick, P., & Wonacott, A.T. (1986). In *An Improved Package for the Measurement of Oscillation Photographs*, pp. 33-39. SERC Daresbury Laboratory, Warrington, UK.
38. Evans, P.R. (1993). Data reduction. Proceeding of CCP4 Study Weekend: Data Collection & Processing 114-122.
39. Blessing, R.H., Guo, D.Y. & Langs, D.A. (1996). Statistical expression value of the Debye-Waller factor and *E*(hkl) values for macromolecular crystals. *Acta Crystallogr. D* **52**, 257-266.
40. Miller, R., Gallo, H.G., Khalak, H.G. & Weeks, C.M. (1994). SnB: Crystal structure determination via Shake-and-Bake. *J. Appl. Crystallogr.* **27**, 613-621.
41. Hauptman, H.A. (1991). A minimal principle in the phase problem. In *Crystallographic Computing 5: From Chemistry to Biology* (Moras, D., Podjarny, A.D. & Thierry, J.C., eds), pp. 324-332. IUCr and Oxford University Press, Oxford, UK.
42. Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In *Isomorphous Replacement and Anomalous Scattering* (Wolf, W., Evans, P.R. & Leslie, A.G.W., eds), pp. 80-88. SERC Proceedings, Daresbury Laboratories, Warrington, UK.
43. Cowtan, K. (1994). DM: An automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 31, 34-38.
44. Collaborative Computation Project, N. 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763.
45. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110-119.
46. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240-255.
47. Laskowski, R.A., McArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the quality of protein structures. *J. Appl. Crystallogr.* **26**, 282-291.
48. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.
49. Esnouf, R. M. (1999). Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta*

Because *Structure with Folding & Design* operates a 'Continuous Publication System' for Research Papers, this paper has been published on the internet before being printed (accessed from http://biomednet.com/cbiology/str). For further information, see the explanation on the contents page.