The EMBO Journal Vol.17 No.5 pp.1183–1191, 1998

Crystal structure of β-ketoacyl-acyl carrier protein synthase II from *E.coli* reveals the molecular architecture of condensing enzymes

Weijun Huang, Jia Jia, Patricia Edwards[1], Katayoon Dehesh[1], Gunter Schneider[2] and Ylva Lindqvist[2]

Division of Structural Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden and [1]Oils Division, Calgene, Inc., 1920 Fifth St., Davis, CA 95616, USA

[2]Corresponding authors
e-mail: gunter@alfa.mbb.ki.se or ylva@alfa.mbb.ki.se

In the biosynthesis of fatty acids, the β-ketoacyl-acyl carrier protein (ACP) synthases catalyze chain elongation by the addition of two-carbon units derived from malonyl-ACP to an acyl group bound to either ACP or CoA. The crystal structure of β-ketoacyl synthase II from *Escherichia coli* has been determined with the multiple isomorphous replacement method and refined at 2.4 Å resolution. The subunit consists of two mixed five-stranded β-sheets surrounded by α-helices. The two sheets are packed against each other in such a way that the fold can be described as consisting of five layers, α–β–α–β–α. The enzyme is a homodimer, and the subunits are related by a crystallographic 2-fold axis. The two active sites are located near the dimer interface but are ~25 Å apart. The proposed nucleophile in the reaction, Cys163, is located at the bottom of a mainly hydrophobic pocket which is also lined with several conserved polar residues. In spite of very low overall sequence homology, the structure of β-ketoacyl synthase is similar to that of thiolase, an enzyme involved in the β-oxidation pathway, indicating that both enzymes might have a common ancestor.

Keywords: condensing enzymes/crystal structure/fatty acid elongation/lipid metabolism/protein crystallography

Introduction

Condensing enzymes comprise a structurally and functionally related family of enzymes that are found in various metabolic pathways, for instance in fatty acid biosynthesis (Somerville and Browse, 1991; Magnuson *et al.*, 1994) and polyketide synthesis, leading in turn to antibiotics, toxins and other secondary metabolites (Hopwood and Sherman, 1990; Katz and Donadio, 1993). These enzymes catalyze carbon–carbon bond-forming reactions by condensing acyl groups, bound to either acyl carrier protein (ACP) or CoA. They can be part of multienzyme complexes, domains of large, multifunctional polypeptide chains such as the mammalian fatty acid synthase, or single enzymes such as the β-ketoacyl synthases in plants and most bacteria (Wakil, 1989; Hopwood and Sherman, 1990; Somerville and Browse, 1991; Katz and Donadio, 1993; Magnuson *et al.*, 1994).

In fatty acid synthesis, the chain elongation step consists of the condensation of acyl groups, derived from acyl-ACP or acyl-CoA with malonyl-ACP. These reactions are catalyzed by a group of enzymes, the β-keto-ACP synthases (KAS, EC 2.3.1.41). The reaction can be described by three distinct steps: (i) the acyl group of acyl-ACP is transferred to a cysteine residue at the active site of the enzyme resulting in a thioester; (ii) generation of a carbanion by the decarboxylation of malonyl-ACP; and (iii) carbon–carbon bond formation by nucleophilic attack of the carbanion onto the carbonyl carbon atom of the thioester (Figure 1). In chemical terms, the overall reaction can be classified as a Claisen condensation.

Several species of β-keto-ACP synthases in plants and bacteria have been identified, distinct in amino acid sequence, chain length specificity for their substrates and sensitivity to cerulenin, an inhibitor of condensing enzymes (Vance *et al.*, 1972; Kauppinen *et al.*, 1988). In *Escherichia coli*, the chain elongation step of fatty acid biosynthesis is carried out by at least three different enzymes, KAS I, II and III (Garwin *et al.*, 1980a; De Mendoza *et al.*, 1982; Kauppinen *et al.*, 1988; Clough *et al.*, 1992; Tsay *et al.*, 1992; Magnuson *et al.*, 1994; Edwards *et al.*, 1997). Both KAS I and KAS II catalyze the condensation of a wide range of saturated acyl-ACPs. KAS I, however, may be responsible for a condensation reaction that cannot be catalyzed by KAS II, namely elongation of C10:1, whereas KAS II is responsible for the elongation of palmoleitic acid C16:1 to *cis*-vaccenic acid C18:1. In *E.coli*, *cis*-vaccenic acid synthesis increases upon a temperature downshift, and KAS II is believed to play a key role in this thermal regulation of fatty acid biosynthesis (Garwin *et al.*, 1980b; Ohashi and Okuyama, 1996). The corresponding genes, *fabF* and *fabB*, have been cloned, sequenced (Kauppinen *et al.*, 1988; Magnuson *et al.*, 1995) and their expressed active recombinant enzymes have been characterized (Edwards *et al.*, 1997). KAS III catalyzes the elongation step of acetyl-CoA and malonyl-ACP to yield acetoacetyl-ACP, the first step in the fatty acid elongation pathway (Clough *et al.*, 1992; Tsay *et al.*, 1992). All these enzymes are active as dimers of identical subunits with an apparent M_r of 42–46 kDa per subunit.

No three-dimensional structure of a β-keto-ACP synthase is as yet available and we have therefore initiated crystallographic studies with the objective of revealing the molecular architecture of the active site as the structural framework for catalysis. In a longer perspective, comparison of the structures of various species of β-keto-ACP synthases with different chain length specificities (Olsen *et al.*, 1995) might provide the basis for redesign of the fatty acid composition of plant oils. As one step towards these objectives, we report the crystal structure of KAS

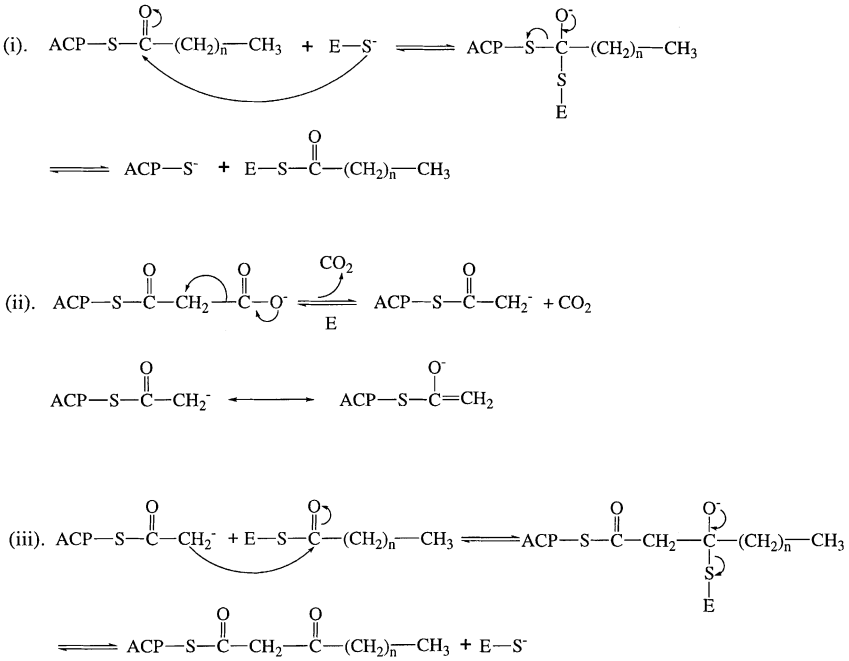

Fig. 1. Scheme of the chemical steps occuring in the fatty acid elongation reaction catalyzed by β-keto acyl synthase.

II from *E.coli* at 2.4 Å resolution and its relationship to the structure of yeast thiolase I (Mathieu *et al.*, 1994, 1997), an enzyme involved in the β-oxidation of fatty acids.

Results and discussion

Structure determination

KAS II from *E.coli* was crystallized in space group $P3_121$ and the crystal structure was determined to 2.4 Å resolution by multiple isomorphous replacement (MIR), including anomalous scattering using mercury and gold derivatives. The final model contains 411 out of the 412 residues of the subunit; no electron density for the N-terminal residue was found. The overall real-space correlation coefficient (Jones *et al.*, 1991) was 0.89. There is well-defined electron density for the polypeptide chain, except in a few loop regions (residues 318 and 374–375) and for some side chains mostly on the molecular surface. Figure 2 shows a representative part of the final refined $2F_o–F_c$ electron density. Residue Leu342 with a 'disallowed' conformation and Ala162 in a 'generously allowed' conformation in the Ramachandran plot are both situated close to the active site residue, Cys163; their electron density is very clear and their unusual conformation might be important for formation of the active site.

The structure of the subunit

The KAS subunit has the shape of a disk with overall dimensions 30×58×60 Å. The subunit can be divided into two halves, each containing a five-stranded mixed β-sheet (Figure 3). The N-terminal half of the structure consists of residues 1–251. The core of this half is the mixed β-sheet, with strand order +3x, +1x, –2x, –1 using the nomenclature defined by Richardson (1977). Two α-helices, α3 and α6, are on one side, and one helix, α7, on the other side of the sheet. The chain makes three long excursions from one end of the β-sheet. The first loop (residues 107–139), connecting β4 with α6, points away from the molecular surface, and part of this loop interacts with the second subunit in the dimer. The other two loop regions, comprising residues 15–53 between β1 and α2 and residues 193–233 between β6 and β7, respectively, interact with each other and with loops at the corresponding end of the second β-sheet in the subunit.

The second half of the subunit is more globular in shape, and also consists of a five-stranded mixed β-sheet with the same strand order as in the N-terminal half of the subunit. Again, helices are packed against the β-sheet, helix α14 on one side and helices α11 and α12 on the other side. A long loop between α14 and β13, comprising residues 358 and 390, packs against the side of the β-sheet and interacts with the excursions from the first β-sheet. The two sheets of the subunit are not only identical in topology, but also similar in structure. Superposition with the program O (Jones *et al.*, 1991) results in 150 equivalent Cα positions with an r.m.s. deviation of 1.6 Å. While this similarity suggests that the KAS subunit might be the result of a gene duplication, there is no strong supporting evidence in the amino acid sequence (13.3% identities, half of which are glycines).

The two sheets pack together facing each other to form the core of the subunit. Two helices, one from each subdomain, form a layer between the two sheets. Thus, the subunit can be described as a five-layered structure, α–β–α–β–α, in the same way as described for yeast thiolase I (Mathieu *et al.*, 1994). The structure starts with a layer of two α-helices, α3 and α6, that pack against the first β-sheet. Two helices, α7 and α14, form the second helical layer sandwiched between the two sheets and, finally, α11 and α12 form the last layer of α-helices (Figure 3B).

The dimer

Consistent with solution studies (Edwards *et al.*, 1997), the quaternary structure of β-keto-ACP synthase in the crystal is a homodimer (Figure 4), and the two subunits

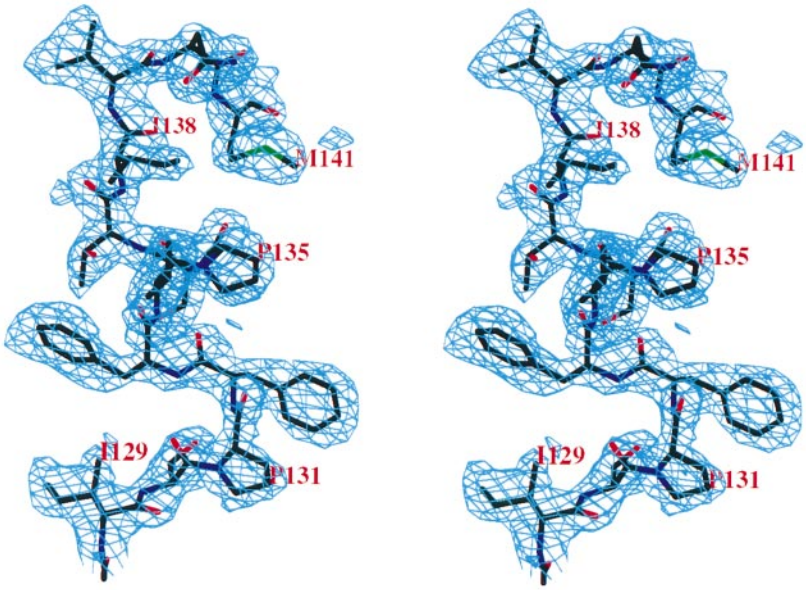

Fig. 2. Stereoview of a part of the final $2F_o–F_c$ electron density map of KAS II at 2.4 Å resolution, contoured at 1.2 σ. Parts of the refined model are included. The figure was generated using O (Jones *et al.*, 1991).

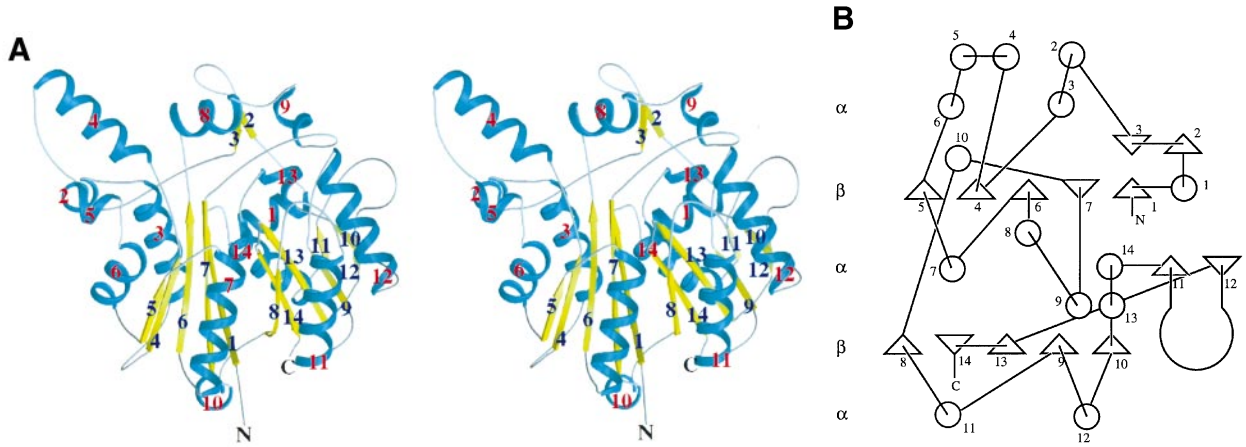

Fig. 3. (**A**) Schematic stereo diagram for the subunit of KAS II. Helices are displayed in light blue with red labels, strands are displayed in yellow with blue labels. The figure was generated with the programs Molscript (Kraulis, 1991) and Raster3D (Merrit and Murphy, 1994). (**B**) Topology diagram of the subunit of KAS II. Triangles represent β-strands, and circles α-helices.

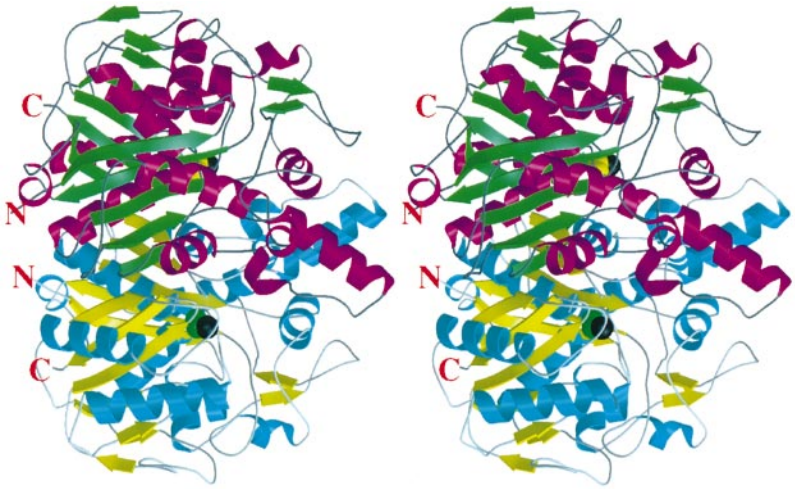

Fig. 4. Stereo view of the dimer of KAS II, the active site cysteine is shown as a space-filling model (black/yellow and black/green spheres, respectively). The figure was generated with the programs Molscript (Kraulis, 1991) and Raster3D (Merrit and Murphy, 1994).

are related by a crystallographic 2-fold axis. The subunits form an extensive interface region as indicated by the large surface/subunit (3030 Å^2) that is buried upon formation of the dimer. The interface area consists of 71% non-polar, 22% polar and 7% charged residues, which is a significantly higher proportion of non-polar residues (at the expense of charged residues) than was found in a survey of subunit–subunit interactions in oligomeric proteins (65% non-polar, 22% polar and 13% charged residues) (Miller *et al.*, 1987).

Several extensive subunit–subunit interaction areas can be defined. Most of the structural elements which participate in dimer formation are located in the N-terminal half of the subunits. A major interaction is through β-strands β5 from both subunits which run antiparallel to each other. In this way, the N-terminal β-sheets in both subunits are extended to form one continuous sheet in the dimer. Another important interaction area are the loops extending from the N-terminal sheet of the subunits. The long loop between β4 and α6 (residues 107–139) interacts with residues from the loop between β6 and β7 (residues 193–233), in particular with residues from helix α8, from the other subunit. Additional interactions are made by residues from helix α6 which packs against the C-terminal end of β-strands β8 and β14.

The fold of condensing enzymes

A detailed comparison of amino acid sequences of condensing enzymes, i.e. β-ketoacyl synthases of fatty acid and polyketide synthesis, revealed that these enzymes form a family, related in amino acid sequence and enzymatic mechanism. The primary sequences of these enzymes are similar, with 18 highly or completely conserved residues when 42 amino acid sequences were compared (Siggaard-Andersen, 1993). Many of these residues are glycine, conserved for structural reasons, or are active site residues which are discussed below. This resemblance in amino acid sequence and the mechanistic similarities strongly suggest that the fold of KAS II is common to the whole family of condensing enzymes.

Relationship to thiolase I

Thiolases are enzymes which can be found in biosynthetic and biodegrading pathways. The biosynthetic enzyme thiolase II catalyzes the formation of acetoacetyl-CoA from acetyl-CoA, whereas thiolase I catalyzes the last step in the β-oxidation pathway, formation of acetyl-CoA from ketoacyl-CoA (Kunau *et al.*, 1994). The three-dimensional structure of the biodegrading thiolase I from yeast is known (Mathieu *et al.*, 1994, 1997), and a comparison revealed that the overall structure of the subunit, the dimer and the active site location of β-ketoacyl synthase are very similar to that of thiolase I.

The comparison of the two structures with the LSQ option in O (Jones *et al.*, 1991) resulted in 208 equivalent Cα atoms with an r.m.s. deviation of 1.7 Å for the subunit and 371 equivalent Cα atoms with an r.m.s. deviation of 1.8 Å for the dimer. Figure 5 shows an overlay of Cα atoms of one subunit of KAS II and the thiolase. The area of structural similarity, comprising about half of the chain, includes the central β-sheets and some of the surrounding α-helices so that the characteristic α–β–α–β–α layers are present in both structures. The loop structures are, however, completely different between the two structures and are also located at different sites in the polypeptide chain, giving rise to large insertions and deletions on alignment of the two sequences. Some of the loops, in particular those at that end of the sheets which participates in the dimer interface of KAS II, do not have their counterparts in the thiolase structure which thus has a 20% smaller dimer interface area, 2400 Å^2. A result of this difference in loop structure is that the active site of thiolase is an open pocket, while in KAS II it is partly covered by these loops. Thus, in spite of the high overall similarity in three-dimensional structure of the core between KAS II and thiolase I, sequence identities are only 18% in the structurally aligned parts, and, outside these regions, no sequence homologies can be detected. Most of the conserved residues are found at the active site of the enzyme and, in fact, some of them are also highly conserved in the whole family of condensing enzymes (Siggaard-Andersen, 1993). In spite of the low sequence homology, the high structural similarity of the protein core suggests that both enzymes might have developed from a common ancestor and thus are evolutionary related.

The active site and the catalytic mechanism

There are two active sites in each dimer, located near the interface region with a distance of 25 Å from each other (Figure 4). Alignment of the amino acid sequences of several β-ketoacyl synthases identified a number of strictly conserved residues (Figure 6). Many of these invariant residues line a predominantly hydrophobic curved pocket, extending from the enzyme surface and leading to the invariant, solvent-accessible residue Cys163 at the floor of this cleft where the pocket turns (Figure 7). This residue is responsible for binding mercury and gold ions used for preparation of the heavy metal derivatives. It is located at the N-terminus of α7 close to the cross-over connections at β5 and β4. The proposed substrate-binding pocket is formed mainly by conserved residues from one subunit in the dimer, although residues from the second subunit form one part of the walls of this cleft (Figure 7). Residues lining this pocket are (invariant residues are underlined, highly conserved residues with a dotted line, see also Figure 6): Cys163; Phe202 and 205–207 at the end of α8; Asp227, Gly 228 and Phe229 just before β7; Thr270–Pro272 in the loop between β8 and α11; Gly276–279 at the entrance to α11; His303–Thr305 and Thr307–Gly310 at the end of β9, the preceding loop and the entrance to α12; His340 in the loop between α13 and α14; Phe398, Gly399, Phe400, Gly401, Gly402 and Asn404 at the end of β13, the turn and start of β14. Residues 131–132 and 134–135 at the entrance of α5 from the second subunit form part of the wall of the active site cleft. The predominatly hydrophobic character of the binding pocket will facilitate binding of the acyl chain of the acyl-ACP substrate.

The assignment of Cys163 as the active site cysteine is based on several lines of evidence. In a sequence comparison of 42 condensing enzymes of fatty acid and polyketide synthesis, Siggaard-Andersen (1993) identified one conserved cysteine residue, which in KAS II corresponds to Cys163. In addition, this cysteine residue superimposes with the active site cysteine in thiolase I, Cys125. Covalent modification studies of β-ketoacyl synthases

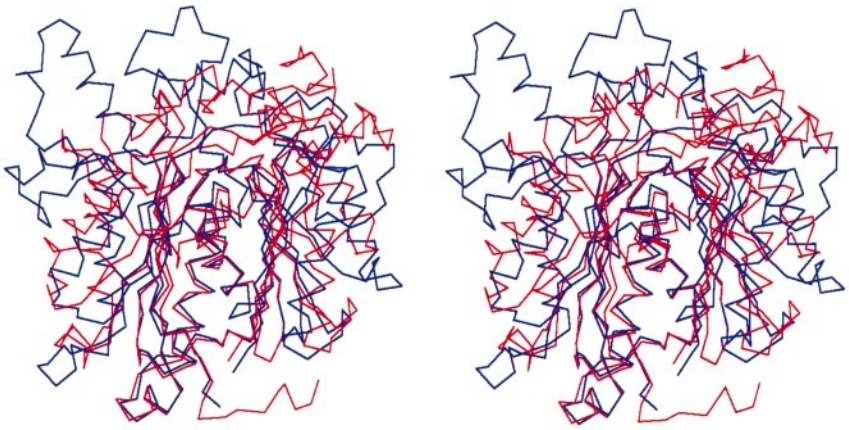

Fig. 5. Superposition of the subunits of KAS II from *E.coli* (blue) and yeast thiolase I (red). The α–β–α–β–α core is conserved, but the remaining parts of the protein structure are very different.

(Kauppinen *et al.*, 1988; Funabashi *et al.*, 1989) and thiolases (Izbicka-Dimitrijevio *et al.*, 1982) as well as mutagenesis of this residue in the β-ketoacyl synthase domain of rat fatty acid synthase (Joshi *et al.*, 1997) and thiolase from *Zooglea ramigera* (Thompson *et al.*, 1989) support the proposed role of this cysteine as the nucleophile in the catalytic reaction.

At the entrance of the active site pocket, a bulky conserved residue, Phe400, is located. This residue points into the active site pocket and in part blocks access to the nucleophilic Cys163. It seems very likely that this side chain has to move in order to allow the acyl substrate to gain full access to the active site. The necessary conformational flexibility might be ensured by the flanking invariant glycine residues, Gly399, Gly401 and Gly402. Interestingly, in thiolase, another bulky residue, Met186, is blocking access to the active site cysteine in a similar fashion (Mathieu *et al.*, 1997).

In close vicinity to the side chain of Cys163 are the side chains of His303 and His340, with distances from the Sγ atom to the Nε atoms of 4.6 and 3.3 Å, respectively. His340 is well situated to act as a base abstracting the proton from the Sγ atom of Cys163, thus enhancing the nucleophilicity of this residue. This histidine is strictly conserved in the sequences of condensing enzymes, which also points to this residue as a major player in catalysis. The Cys–His pair in the active site of condensing enzymes and thiolase I is reminiscent of the catalytic diad found in cysteine proteases (Kamphius *et al.*, 1984).

In the active site of thiolase, one of these two histidine residues, His375 (corresponding to His340 in KAS II), is conserved and might have similar catalytic function (Mathieu *et al*., 1997). A third catalytic residue in thiolase, Cys403, does not have its counterpart in KAS II, indicating that the catalytic mechanism in the two enzymes is not identical.

In addition to enhancing the nucleophilicity of Cys163, the enzyme has to stabilize several transition states and intermediates, for instance the negative charge at the carbonyl oxygen in the tetrahedral intermediates during thioester formation and hydrolysis (Figure 1). Charge stabilization might be provided by these histidine residues and, in addition, by a hydrogen bond with the main chain NH of F400. A contribution of the main chain nitrogen of the corresponding residue, G405, in thiolase to stabilization of the negative charge developing at the carbonyl oxygen atom has also been suggested (Mathieu *et al.*, 1997). A positive charge on His303 could aid in guiding the binding of the second substrate into the active site pocket so that the C2 carbon is in position for the condensation. The location of the side chain of His303 suggests that it is involved in the decarboxylation of malonyl-ACP and in the stabilization of the formed carbanion intermediate.

Fig. 6. Sequence of β-ketoacyl synthase II from *E.coli* (Magnuson *et al.*, 1995). The secondary structure elements of KAS II are shown and labeled. Residues strictly conserved in condensing enzymes are highlighted in yellow. The sequences of the following condensing enzymes were used in the alignment (DDBJ/EMBL/GenBank Accession numbers in parentheses): KAS II (P39435), *Arabidopsis thaliana* (P55338), *Mycobacterium tuberculosis* (Q10525), *Synechocystis* sp. strain PCC6803 (g1652360) and *Cuphea wrightii* (g1698687), KASI from *E.coli* (P14926), *Haemophilus influenzae* (P43710), *Arabidopsis thaliana* (P52410), *Hordeum vulgare* (P23902) and *Mycobacterium tuberculosis* (Q10524), KAS from castor bean chloroplast (g294665), KAS domain of fatty acid synthase from human (P49327), rat (P12785) and chicken (P12276).

The sequence alignment of condensing enzymes has revealed a few residues which are highly conserved (at least 90%) in these enzymes (Siggaard-Andersen, 1993). Among those residues are the cysteine acting as a nucleophile and the two histidine residues mentioned above. Two acidic residues, Asp311 and Glu314, and a basic residue, Lys335, are also completely conserved in this

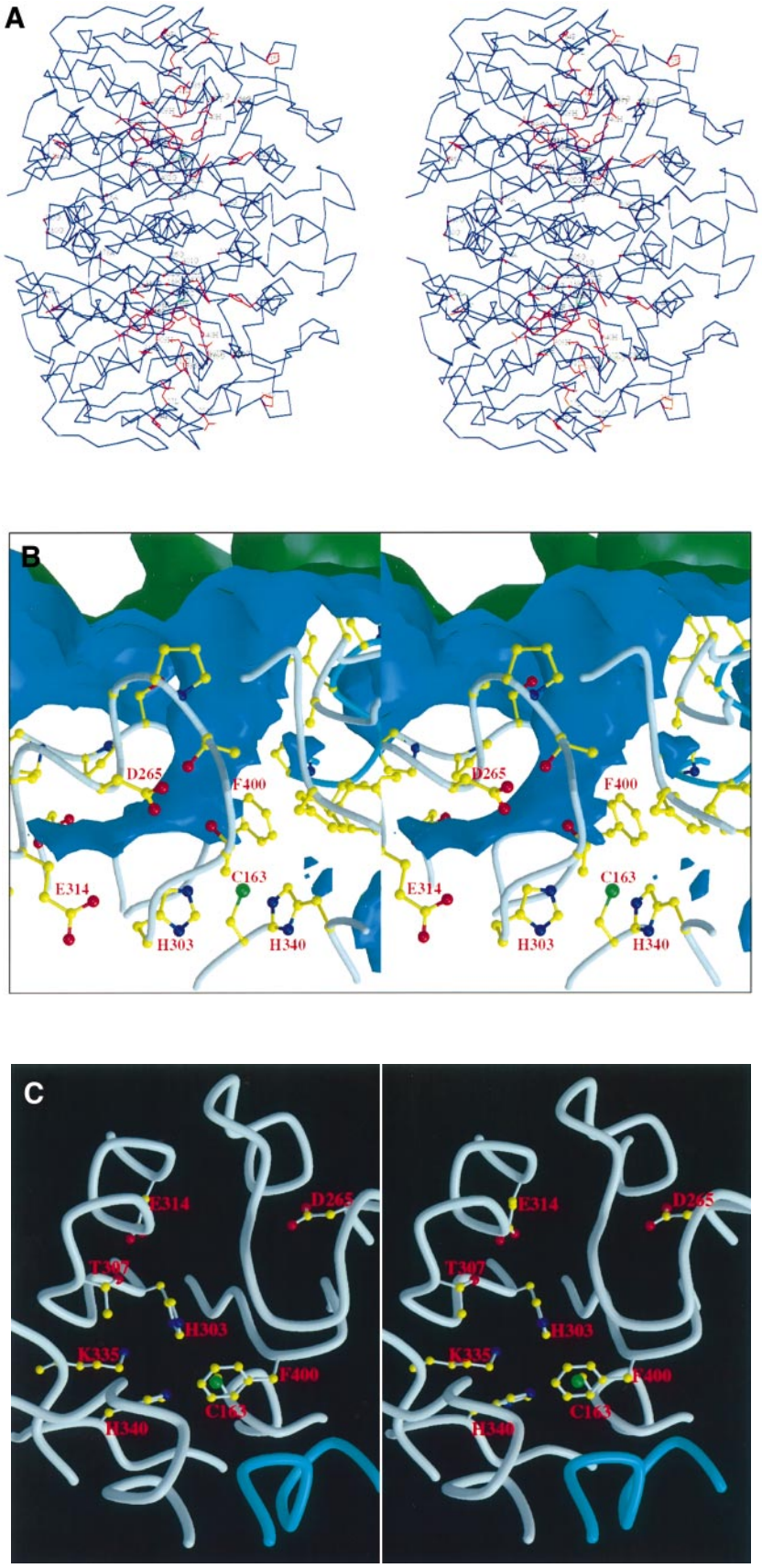

Fig. 7. (**A**) Cα trace of the dimer with all the conserved residues displayed. (**B**) Stereo view illustrating the curved active site pocket of KAS II. The view is perpendicular to the entrance of the active site pocket. The figure was generated with the programs Molscript (Kraulis, 1991) and Raster3D (Merrit and Murphy, 1994). (**C**) Stereo view of the active site as seen from the entrance of the active site pocket. The locations of the conserved active site residues discussed in the text are shown. The polypeptide chains of the two subunits are shown in gray and blue, respectively.

Table I. Diffraction data and phasing statistics[a]

Compound	Resolution	Total observations	Unique observations	I/σ	Completeness (%)	R_{sym}	R_{deri}	No. of sites	Phasing power	R_{Cullis}
Nati1	2.50 (2.56–2.50)	134 574	17 098	10.3 (1.79)	95.0 (81.6)	7.3 (33.2)				
Nati2	2.40 (2.45–2.40)	141 114	18 580	10.2 (2.26)	95.5 (90.5)	7.0 (27.1)				
Au1	3.00 (3.07–3.00)	35 809	8741	8.2 (1.67)	83.8 (74.2)	8.0 (30.0)	16.2	3	1.41	0.73
Au2	2.80 (2.87–2.80)	77 135	11 441	11.3 (2.46)	89.7 (76.5)	7.0 (24.5)	9.7	3	1.12	0.81
Au3	2.80 (2.87–2.80)	56 911	9240	9.1 (1.93)	72.4 (62.6)	7.8 (28.6)	12.5	3	1.25	0.77
Au4	2.90 (2.97–2.90)	22 790	5106	8.6 (1.71)	44.1 (28.1)	8.1 (25.8)	19.2	2	1.73	0.68
EMTS1	3.00 (3.07–3.00)	82 115	9945	6.1 (1.90)	94.8 (88.4)	13.2 (35.0)	22.2	2	1.40	0.73
EMTS2	2.80 (2.87–2.80)	59 442	7844	10.8 (2.12)	61.1 (55.0)	7.1 (28.3)	14.6	2	1.63	0.71
EMTS3	3.00 (3.07–3.00)	32 117	5366	7.0 (1.66)	51.6 (37.6)	10.5 (28.1)	22.7	2	1.10	0.81
Pip1	3.30 (3.38–3.30)	8848	3788	5.3 (1.54)	48.1 (38.2)	9.5 (22.1)	22.9	2	0.79	0.89
Pip2	2.90 (2.97–2.90)	43 572	8003	7.8 (1.45)	70.3 (46.4)	9.0 (32.1)	20.7	2	0.75	0.89

Overall figure-of-merit = 0.54.
[a]Values in parentheses are those for the highest resolution shell.
R_{deri}: merging *R*-factor between the nati1 and a derivative data set, $\Sigma|F_{PH} - F_P| / \Sigma F_P$
EMTS: ethylmercurithiosalicylate. Au: $KAuI_4$, Pip: di-μ-iodobis(ethylenediamine)diplatinum(II).
$R_{sym} = \Sigma_{hkl}\Sigma_i|I_i - \langle I\rangle| / \Sigma_{hkl}\Sigma_i\langle I\rangle$ where I_i is the intensity of a measurement and $\langle I\rangle$ is the mean intensity for that reflection.
Phasing power = r.m.s. ($|F_H|/E$), where $|F_H|$ = heavy atom structure factor amplitude and E = r.m.s. lack of closure.
R_{Cullis} = lack of closure/isomorphous difference.

enzyme family. These residues are located in the vicinity of the active site, but are most likely not directly involved in catalysis. Asp311 and Glu314 are located on the same face of helix α12, and their side chains point towards β-strand β9. They appear to be primarily of structural importance securing correct geometry in the active site, since they form hydrogen bonds to main chain nitrogens; Asp311 to the peptide nitrogen atoms of Ser306 and Trp222, Glu314 to the NH groups of His303 and Gly304, thus positioning one of the catalytic groups at the active site. The charge of Asp311 is also compensated by an interaction with Arg 249, while Glu314 together with Asp265 will give a negatively charged surface at the very bottom of the binding pocket which might promote the release of CO_2 from malonate-ACP. Lys335, situated between the two histidine residues, is involved in an internal salt bridge to Glu349 and is within hydrogen bond distance to the peptide oxygen of His303. The amino group of Lys335 is pointing away from the Sγ atom of Cys163 (distance 7.8 Å), and the function of this residue might again be primarily structural rather than catalytic.

The side chain of the invariant polar residue Thr307, together with the N-terminus of the preceding α-helix, can form a binding site at the top of the binding cavity for the hydroxyl and the phosphate of the phosphopantetheine group of the acyl-ACP substrates. At some distance from the active site pocket is the invariant residue Asp223. It is located at the surface of the enzyme and could be involved in the interaction with ACP.

The active site architecture as revealed in the structure of the unliganded enzyme thus shows a possible route of approach of the acyl group of the acyl-ACP substrate to the nucleophile at the bottom of the active site cleft, and suggests the role of some enzyme residues in the subsequent steps of the mechanism. However, a more detailed discussion of substrate binding and catalysis will have to await crystallographic studies of enzyme–inhibitor and enzyme–substrate complexes in combination with further biochemical studies.

Conclusions

The crystal structure of KAS II, the first known structure of a condensing enzyme, reveals the fold for this class of enzymes as well as the architecture of the active site. The comparison with thiolases of the β-oxidation pathway shows that both enzymes contain the same fold, which together with mechanistic similarities suggests that these enzymes are evolutionary related.

Lipid biosynthesis recently has attracted considerable interest with the prospect of modifying the composition of these lipids by genetic means with respect to chain length, saturation and chemical modifications such as hydroxylation (Somerville and Browse, 1991; Verwoert *et al.*, 1995). The redesign of the substrate specificities of β-ketoacyl synthases represents one possible approach towards this objective. Structural studies of β-keto-ACP synthases utilizing varying substrates are expected to reveal the molecular basis of substrate specificity and provide the framework for structure-based enzyme design.

Materials and methods

Crystallization and data collection

KAS II was purified as described previously (Edwards *et al.*, 1997). Crystals of the enzyme were grown in solutions of 10 mg of enzyme/ml (in 20 mM HEPES buffer, pH 7.0) by the hanging drop vapor diffusion method at room temperature, using 27% PEG 8000 as precipitant, buffered at pH 7.5 with 0.1 M HEPES. The crystals grow to a size of $0.5\times0.3\times0.2$ mm^3 within 3 days. The lifetime of these crystals is very limited; they will dissolve within 10 days of their appearance. Addition of 0.1% mercaptoethanol to the reservoir solution significantly increased the life time of the crystals.

X-ray data sets were collected at 4°C using a MAR imaging plate, mounted on a Rigaku rotating anode operating at 50 kV and 90 mA. Data set NATI2 was collected with a MAR image plate at beamline BW7B at the EMBL outstation, DESY, Hamburg. Data frames were evaluated with DENZO and SCALEPACK (Otwinowski, 1993). Statistics of the data collection are given in Table I. The determination of the space group and cell dimensions was carried out using the auto-indexing option in DENZO and pseudo precession images calculated with the program PATTERN (G.Lu, unpublished). This analysis showed that the crystals are of space group $P3_121$ (or its enantiomorph $P3_221$) with cell

Table II. Refinement statistics

Resolution (Å)	8.0–2.4
No. of reflections in working set	16 745
No. of reflections in test set	1810
No. of non-hydrogen protein atoms	3004
No. of waters	31
R-factor	0.234
R-free	0.279
R.m.s. deviations from ideals	
Bond distance (Å)	0.004
Angle distance (Å)	0.018
Average *B*-value (Å^2)	37.4
r.m.s. *B* (Å^2)	0.8
Ramachandran plot	
Percentage of non-glycine residues in the	
most favorable regions	88.3
additional allowed regions	11.1
generously allowed regions	0.3
disallowed regions	0.3

dimensions a = 76.4 Å and c = 146.8 Å, γ = 120°. There is one molecule in the asymmetric unit, giving a V_M value of 2.76 Å^3/Da.

MIR phasing and model building

Three independent heavy metal derivatives were prepared by soaking KAS II crystals in a 5 mM solution of the heavy atom compound in mother liquor for several hours. The difference Patterson maps calculated for these derivatives had clear peaks on the Harker sections. One major binding site was located by manual inspection of the mercury derivative. Phases calculated from this derivative were used to identify the positions of the other heavy metal sites in the remaining derivatives. Finally, the parameters for all the heavy metal sites were refined with MLPHARE (Otwinowski, 1991; CCP4, 1994). Inclusion of the anomalous signal from the mercury and gold derivatives identified the space group as $P3_121$.

The MIR map was improved by solvent flattening using the program DM (Cowtan and Main, 1996). In these electron density maps, a number of secondary structure elements were clearly recognizable, and the space group was confirmed to be $P3_121$ by the appearance of right-handed helices. Polyalanine model building with the program O (Jones *et al.*, 1991), crystallographic refinement with XPLOR (Brünger *et al.*, 1987) and phase combination with SIGMAA (Read, 1986) were performed iteratively before a complete trace of the main chain and the assignment of the complete biological amino acid sequence, derived from the DNA sequence of the *fabF* gene (Magnuson *et al.*, 1995), were achieved.

Crystallographic refinement and model validation

The model was refined with XPLOR (Brünger *et al.*, 1987) and REFMAC (Murshudov *et al.*, 1997). A portion of 10% of the reflections was set apart for monitoring *R*-free (Brünger, 1992) throughout the refinement process. Bond and angle parameters as described by Engh and Huber (1991) were used. Iterative cycles consisting of refinement, phase combination and model building eventually led to a complete protein model with an *R*-free of 37% (resolution interval 8–2.5 Å). At this stage, a new data set to 2.4 Å resolution collected at beamline X11 at the EMBL outstation, DESY, Hamburg, became available, and refinement continued with this data set. After a few rounds of remodeling and positional refinement, anisotropic overall *B*-factor correction was performed before individual *B*-factors were refined using XPLOR, resulting in an *R*-free of 30.8%. Refinement continued with REFMAC, resulting in a significant improvement of the electron density map. Two more rounds of refinement were performed, yielding an *R*-free of 27.9% and an *R*-cryst of 23.3%. The model was analyzed with PROCHECK (Laskowski *et al.*, 1993). The Ramachandran plot showed one residue in the disallowed regions and one residue in the generously allowed regions; both residues are well defined in electron density. Details of the refinement procedure and the statistics for the final model are given in Table II. The positions of the heavy metal sites were confirmed by difference Fourier analysis, using phases calculated from the refined protein model. The mercury and gold ions bind to residue Cys163 and the platinum compound binds to Met149 in a cleft on the molecular surface.

Initial alignment of KAS and thiolase was achieved using TOP (G.Lu, unpublished). Structural comparisons were carried out using the LSQ option in the program O (Jones *et al.*, 1991), with residues considered to be equivalent if their Cα atoms were within 3.8 Å and in a stretch of more than three consecutive equivalenced atoms. Molecular surface calculations were made with GRASP (Nichols *et al.*, 1991) and MAPROP (G.Kleijwegt, unpublished). The coordinates for thiolase I (Mathieu *et al.*, 1997) were retrieved from the Protein Data Bank, Brookhaven (entry number 1afw). The atomic coordinates and structure factors for KAS II have been deposited with the Protein Data Bank, Brookhaven, accession codes 1kas and r1kassf, respectively.

Acknowledgements

We thank Dr Rik Wierenga for communicating results prior to publication. We gratefully acknowledge access to synchrotron radiation at the EMBL outstation, DESY, Hamburg, and thank the staff, in particular Dr Victor Lamzin, for help at the beamline. This work was supported by a grant from the Swedish Natural Science Research Council, the Foundation for Strategic Research and the Knut and Alice Wallenberg foundation.

References

Brünger,A.T. (1992) Free *R*-value: a novel statistical quantity for assessing the accuracy of crystal structure. *Nature*, **355**, 472–475.

Brünger,A.T., Kuriyan,J. and Karplus,M. (1987) Crystallographic *R* factor refinement by molecular dynamics. *Science*, **235**, 458–460.

Collaborative Computational Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.

Clough,R.C., Mathis,A.L., Barnum,S.R. and Jaworsky,J.G. (1992) Purification and characterization of 3-ketoacyl-acyl carrier protein synthase III from spinach. A condensing enzyme utilizing acetyl-coenzyme A to initiate fatty acid biosynthesis. *J. Biol. Chem.*, **267**, 20992–20998.

Cowtan,K.D. and Main,P. (1996) Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallogr.*, **D52**, 43–48.

De Mendoza,D., Garwin,J.L. and Cronan,J.E.,Jr (1982) Overproduction of *cis*-vaccenic acid and altered temperature control of fatty acid synthesis in a mutant of *Escherichia coli*. *J. Bacteriol.*, **151**, 1608–1611.

Edwards,P., Nelsen,J.S., Metz,J.G. and Dehesh,K. (1997) Cloning of the *fabF* gene in an expression vector and *in vitro* characterization of recombinant *fabF* and *fabB* encoded enzymes from *Escherichia coli*. *FEBS Lett.*, **402**, 62–66.

Engh,R.A. and Huber,R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr.*, **A47**, 392–400.

Funabashi,H., Kawaguchi,A., Tomoda,H., Omura,S., Okuda,S. and Iwasaki,S. (1989) Binding site of cerulenin in fatty acid synthase. *J. Biochem.*, **105**, 751–755.

Garwin,J.L., Klages,A.L. and Cronan,J.E.,Jr (1980a) Structural, enzymatic, and genetic studies of β-ketoacyl-acyl carrier protein synthases I and II of *Escherichia coli*. *J. Biol. Chem.*, **255**, 11949–11956.

Garwin,J.L., Klages,A.L. and Cronan,J.E.,Jr (1980b) β-Ketoacyl-acyl carrier protein synthase II of *Escherichia coli*. Evidence for function in the thermal regulation of fatty acid synthesis. *J. Biol. Chem.*, **255**, 3263–3265.

Hopwood,D.A. and Sherman,D.H. (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.*, **24**, 37–66.

Izbicka-Dimitrijevio,E. and Gilbert,H.F. (1982) Two sulfhydryl groups near the active site of thiolase I from porcine heart: modification of thiolase with the fluorescent thiol reagent *S*-mercurio-*N*-dansyl-L-cysteine. *Biochemistry*, **21**, 6112–6118.

Jones,T.A., Zou,J.-Y., Cowan,S. and Kjeldgaard,M. (1991) Improved methods for building protein models into electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 100–119.

Joshi,A.K., Witkowski,A. and Smith,S. (1997) Mapping of functional interactions between domains of animal fatty acid synthase by mutant complementation *in vitro*. *Biochemistry*, **36**, 2316–2322.

Kamphius,I.G., Kalk,K.H., Swarte,M.B.A. and Drenth,J. (1984) Structure of papain refined at 1.65 Å resolution. *J. Mol. Biol.*, **179**, 233–256.

Kauppinen,S., Siggaard-Andersen,M. and von 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.

Katz,L. and Donadio,S. (1993) Polyketide synthesis: prospects for hybrid antibiotics. *Annu. Rev. Microbiol.*, **47**, 875–912.

Kraulis,P. (1991) Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Kunau,W.-H., Dommes,Y. and Schulz,H. (1994) β-Oxidation of fatty acids in mitochondria, peroxisomes and bacteria. A century of continued progress. *Prog. Lipid Res.*, **34**, 267–342.

Laskowski,R.A., McArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the quality of protein structures. *J. Appl. Crystallogr.*, **26**, 282–291.

Magnuson,K., Carey,M.R. and Cronan,J.E.,Jr (1995) The putative *fabJ* gene of *Escherichia coli* fatty acid synthesis is the *fabF* gene. *J. Bacteriol.*, **177**, 3593–3595.

Magnuson,K., Jackowski,S., Rock,C.O. and Cronan,J.E. (1994) Regulation of fatty acid biosynthesis in *Escherichia coli*. *Microbiol. Rev.*, **57**, 522–542.

Mathieu,M., Zeelen,J., Pauptit,R.A., Erdmann,R., Kunau,W.-H. and Wierenga,R.K. (1994) The 2.8 Å crystal structure of peroxisomal 2-ketoacyl-CoA thiolase of *Sacharomyces cerevisiae*: a five-layered αβαβα structure constructed from two core domains of identical topology. *Structure*, **2**, 797–808.

Mathieu,M., Modis,Y., Zeelen,J.P., Engel,C.K., Abagyan,R.A., Ahlberg,A., Rasmussen,B., Lamzin,V.S., Kunau,W.H. and 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.

Merrit,E.A. and Murphy,M.E.P. (1994) Raster 3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.*, **D50**, 869–873.

Miller,S., Lesk,A.M., Janin,J. and Chothia,C. (1987) The accessible surface area and stability of oligomeric proteins. *Nature*, **328**, 834–836.

Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.*, **D53**, 240–255.

Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.

Ohashi,Y. and Okuyama,H. (1986) Effects of the temperature range and the lack of beta-ketoacyl acyl-carrier protein synthase II on fatty acid synthesis in *Escherichia coli* K12 after shifts in temperature. *Biochim. Biophys. Acta*, **876**, 146–153.

Olsen,J.G., Kadziola,A., Sigaard-Andersen,M., Chuck,J.-A., Larsen,S. and von Wettstein-Knowles,P.V. (1995) Preliminary X-ray diffraction analysis of β-ketoacyl-[acyl carrier protein] synthase I from *Escherichia coli*. *Protein Peptide Lett.*, **1**, 246–251.

Otwinowski,Z. (1991) Maximum likelyhood refinement of heavy-atom parameters. In Wolf,W., Evans,P.R. and Leslie,A.G.W. (eds), *Proceedings of the CCP4 Study Weekend: Isomorphous Replacement and Anormalous Scattering*. SERC Daresbury Laboratory, Warrington, UK, pp. 87–95.

Otwinowski,Z. (1993) Oscillation data reduction program. In Sawyer,L., Isaacs,N. and Bailey,S.S. (eds), *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*. SERC Daresbury Laboratory, Warrington, UK, pp. 56–62.

Read,R.J. (1986) Improved coefficients for map calculation using partial structures with errors. *Acta Crystallogr.*, **A42**, 140–149.

Richardson,J.S. (1977) β-Sheet topology and the relatedness of proteins. *Nature*, **268**, 495–500.

Siggaard-Andersen,M. (1993) Conserved residues in condensing enzyme domains of fatty acid synthases and related sequences. *Protein Sequence Data Anal.*, **5**, 325–335.

Somerville,C. and Browse,J. (1991) Plant lipids: metabolism, mutants, and membranes. *Science*, **252**, 80–87.

Thompson,S., Mayerl,F., Peoples,O.P., Masamune,S., Sinskey,A.J. and Walsh,C. (1989) Mechanistic studies on β-ketoacyl-thiolase from *Zoogloea ramigera*: identification of the active-site nucleophile as Cys89, its mutation to Ser89, and kinetic and thermodynamic characterization of wild-type and mutant enzymes. *Biochemistry*, **28**, 5735–5742.

Tsay,J.T., Oh,W., Larson,T.J., Jackowski,S. and 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.

Vance,D.E., Goldberg,I., Mitsuhashi,O. and Bloch,K. (1972) Inhibition of fatty acid synthetases by the antibiotic cerulenin. *Biochem. Biophys. Res. Commun.*, **48**, 649–656.

Verwoert,I.I., van der Linden,K.H., Walsh,M.C., Nijkamp,H.J. and 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.

Wakil,S.J. (1989) Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry*, **28**, 4523–4530.

Received November 27, 1997; revised December 23, 1997; accepted December 29, 1997