Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing

Francis K Athappilly[1]* and Wayne A Hendrickson[1,2]

[1]Department of Biochemistry and Molecular Biophysics and [2]Howard Hughes Medical Institute, College of Physicians and Surgeons of Columbia University, New York, NY 10032, USA

Background: Acetyl-coenzyme A carboxylase catalyzes the first committed step of fatty acid biosynthesis. Universally, this reaction involves three functional components all related to a carboxybiotinyl intermediate. A biotinyl domain shuttles its covalently attached biotin prosthetic group between the active sites of a biotin carboxylase and a carboxyl transferase. In *Escherichia coli*, the three components reside in separate subunits; a biotinyl domain is the functional portion of one of these, biotin carboxyl carrier protein (BCCP).

Results: We have expressed natural and selenomethionyl (Se-met) BCCP from *E. coli* as biotinylated recombinant proteins, proteolyzed them with subtilisin Carlsberg to produce the biotinyl domains $BCCP_{sc}$ and Se-met $BCCP_{sc}$, determined the crystal structure of Se-met $BCCP_{sc}$ using a modified version of the multiwavelength anomalous diffraction (MAD) phasing protocol, and refined the structure for the natural $BCCP_{sc}$ at 1.8 Å resolution. The structure may be described as a capped β sandwich with quasi-dyad symmetry. Each half contains a characteristic hammerhead motif. The biotinylated lysine is located at a hairpin β turn which connects the two symmetric halves of the molecule, and its biotinyl group interacts with a non-symmetric protrusion from the core.

Conclusions: This first crystal structure of a biotinyl domain helps to unravel the central role of such domains in reactions catalyzed by biotin-dependent carboxylases. The hammerhead structure observed twice in $BCCP_{sc}$ may be regarded as the basic structural motif of biotinyl and lipoyl domains of a superfamily of enzymes. The new MAD phasing techniques developed in the course of determining this structure enhance the power of the MAD method.

Structure 15 December 1995, **3**:1407–1419

Key words: biotin carboxyl carrier protein, carboxyl transfer, hammerhead structure, selenomethionine

Introduction

All biotin-dependent enzymes [1] share a common feature: a biotinyl domain that transfers an activated carboxyl group to a transcarboxylase center. Acetyl-coenzyme A (CoA) carboxylase is a multicomponent biotin enzyme present in all animals, plants and bacteria. It catalyzes one of the regulated steps (the first committed step) in the biosynthesis of long-chain fatty acids, namely the biotin-dependent carboxylation of acetyl-CoA to form malonyl-CoA [2,3].

Because of its simplicity of organization, acetyl-CoA carboxylase from *Escherichia coli* provides an attractive system to study. Its three functional elements, which reside in separate and easily dissociated subunits, are as follows: biotin carboxylase, which carboxylates biotin in the presence of bicarbonate, ATP and Mg^{2+}; carboxyl transferase, which transfers the carboxylate group from biotin to acetyl-CoA to form malonyl-CoA; and biotin carboxyl carrier protein (BCCP). Between the carboxylation and the transcarboxylation steps, BCCP shuttles its covalently attached biotin prosthetic group, as the carboxybiotinyl intermediate, from the active site of the biotin carboxylase to that of the carboxyl transferase, and then it recycles the empty biotinyl group back for re-carboxylation. The role of BCCP in the carboxyl transfer reaction of acetyl-CoA carboxylase was worked out in a series of early studies by the groups of Vagelos and Lane [3–12].

In 1971, Nervi *et al.* [4] isolated a 9 kDa biotinylated protein from *E. coli*. Fall and Vagelos [8] later showed that this protein was a proteolytic product of a larger intact BCCP, which on digestion with subtilisin Carlsberg yielded a fragment very similar to that isolated by Nervi *et al.* This product of digestion with subtilisin Carlsberg was active as a carboxyl acceptor and donor in the biotin carboxylase and carboxyl transferase reactions of acetyl-CoA carboxylase, and was named $BCCP_{sc}$.

Structural information on individual components, such as BCCP, is important for understanding the structural organization and structure/function relationships in acetyl-CoA carboxylase and other biotin-dependent enzymes. Even though the structure of the biotin carboxylase subunit has been published [13], details of its interaction with the biotinyl domain are still unknown. While preliminary crystallographic investigations on $BCCP_{sc}$ have been reported [8,12], no structure of any biotinyl domain has been published. The family of biotin-dependent enzymes is also related in sequence to the lipoyl domains of lipoic acid dependent enzymes such as pyruvate dehydrogenase [1,14]. There are three reported lipoyl domain structures; two NMR structures [15,16] and the X-ray crystal structure of a more distantly related lipoyl domain from glycine decarboxylase [17].

Biotinylation of apocarboxylases to form holoenzymes is important from a clinical perspective. Defects in either holocarboxylase synthetase or biotinidase resulting in impaired biotinylation are responsible for inherited biotin-responsive multiple carboxylase deficiency in humans [18–20]. The symptoms of this disorder include

*Corresponding author.

 ISSN 0969-2126

skin rash, loss of hair, difficulty in feeding and breathing, developmental delay, seizures and coma. The biotinylation reaction is evolutionarily conserved such that synthetases function with various apocarboxylases across species boundaries. Thus, bacterial holocarboxylase synthetase can biotinylate rat apocarboxylases [21,22]. Only the biotinyl domain of apocarboxylase is necessary for biotinylation [23,24], although synthetases can biotinylate intact apocarboxylases [25–27]. Thus, we expect that an analysis of biotinylation of $BCCP_{sc}$ by the *E. coli* holocarboxylase synthetase (BirA) [25,28], the structure of which is known [29], will enhance our understanding of biotinylation in man.

Here we report the crystal structure of $BCCP_{sc}$ at 1.8 Å resolution. We isolated the gene encoding BCCP by polymerase chain reaction (PCR), expressed and purified BCCP as a recombinant protein, and produced $BCCP_{sc}$ by proteolysis. In order to carry out multiwavelength anomalous diffraction (MAD) phasing, we prepared the selenomethionyl (Se-met) analog of $BCCP_{sc}$, and crystallized it using a novel device. The structure of Se-met $BCCP_{sc}$, as determined at 2.5 Å resolution by a modified version of the MAD phasing technique, was then used to refine the isomorphous crystal structure of natural $BCCP_{sc}$. We also compare this molecule with the known structures of two lipoyl domains [15,17] and, by structure-based sequence alignments, with other biotin and lipoyl enzymes.

Results and discussion

Structure determination

The gene for *E. coli* BCCP was isolated by PCR and cloned into an expression vector producing an N-terminal polyhistidyl tag for affinity purification. The functional biotinylated fragment, $BCCP_{sc}$ (residues 77–156 relative to the intact protein), was obtained by digestion of the recombinant BCCP with subtilisin Carlsberg [8]. N-terminal sequencing, streptavidin binding, and mass spectrometry conclusively showed that the isolated fragment was authentic biotinylated $BCCP_{sc}$. The protein was crystallized in the space group $P2_12_12$ with unit cell parameters of a=65.46 Å, b=37.26 Å, c=35.45 Å. There is one subunit per asymmetric unit and the crystals have a solvent content of 49%.

In order to carry out MAD phasing, we prepared the Se-met analog of $BCCP_{sc}$ and, using anaerobic conditions, grew crystals isomorphous with those of the natural protein using a newly developed dialysis device. Anomalous diffraction data from these crystals were measured at four wavelengths around and at the selenium K-shell edge. Amplitudes of structure factors for normal scattering from the anomalous scatterers (selenium atoms), $|F_A|$, and those from the entire structure, $|F_T|$, and the phase differences, $(\phi_T-\phi_A)$, were calculated using our conventional MAD phasing programs [30–32] (Tables 1,2). Phase estimations were carried out to maximal Bragg spacings of 2.5 Å. The selenium atoms were located by direct methods [33] and their coordinates were refined.

Despite good phasing statistics for 85% of the data, electron-density maps computed using the algebraic [31] or probabilistic [34] estimates of the phases could not be interpreted before or after solvent flattening [35]. At this point, the MAD phasing protocol was modified to incorporate calculated F_A values ($|F_A|$ and φ_A), given the refined parameters of selenium atoms, in an evaluation that then reduces the unknown parameters to $|F_T|$ and φ_T. The new procedure enabled us to estimate phases of significantly more reflections than before. Then, a 2.5 Å resolution electron-density map was computed after solvent flattening [35].

A model for Se-met $BCCP_{sc}$ was built into the electron-density map. This model was then refined at 1.8 Å resolution against data from natural $BCCP_{sc}$ crystals to give

Table 1. Data statistics for Se-met $BCCP_{sc}$ crystals.

	Diffraction measurements*				MAD difference ratios (20.0–2.5 Å)†				Anomalous scattering factors‡	
Wavelength (Å)	Unique reflections§	Redundancy	Completeness (%)	R_{sym}# (%)	λ1	λ2	λ3	λ4	f′ (e)	f″ (e)
0.9871 (Remote λ1)	5559 (2334)	2.0 (2.0)	96.1 (95.8)	5.6 (8.0)	0.029 (0.032)	0.060	0.056	0.048	−4.5 (−4.8)	0.9 (0.6)
0.9794 (Inflection λ2)	5563 (2367)	1.9 (1.9)	96.2 (97.2)	5.7 (8.6)		0.056 (0.032)	0.036	0.065	−11.2 (−10.7)	4.5 (3.3)
0.9792 (Peak λ3)	5506 (2320)	1.8 (1.8)	95.2 (95.2)	5.8 (8.9)			0.067 (0.031)	0.054	−9.1 (−8.9)	5.4 (4.6)
0.9686 (Remote λ4)	5387 (2273)	1.8 (1.8)	93.2 (93.3)	5.7 (8.4)				0.052 (0.031)	−3.9 (−4.0)	3.8 (3.5)

*Values shown are for reflections at 20.0–2.5 Å and 3.0–2.5 Å (in parentheses) spacing intervals. †MAD difference ratios are rms $(\Delta|F|)/\text{rms}(|F|)$, where $\Delta|F|$ is the Bijvöet difference at one wavelength (diagonal elements in the middle panel) or dispersive difference between pairs of wavelengths (off-diagonal elements in the middle panel). The values in parentheses are ratios for centric reflections, which ideally should be zero. They are indicators of noise in the anomalous signals. ‡The refined values of scattering factors from MADLSQ are given. The values in parentheses are obtained after the modified procedure. §Friedel mates, F(h) and F(−h), are included as unique reflections. #$R_{sym}=\Sigma\Sigma|I_i(h)-\langle I(h)\rangle|/\Sigma\Sigma I_i(h)$ where for each reflection h, $I_i(h)$ is the i^{th} observation and $\langle I(h)\rangle$ is the weighted mean of all observations.

Table 2. Resolution dependence of MAD differences and phasing statistics.

Shell (Å)	20.0–5.0	5.0–3.0	3.0–2.5	2.5–2.3
Bijvöet difference at the peak wavelength	0.083 (0.031)	0.062 (0.023)	0.078 (0.052)	0.084 (0.064)
Maximum dispersive difference	0.078	0.052	0.067	0.073
Minimum dispersive difference	0.039	0.031	0.047	0.056
$R(\lvert {}^{\circ}F_T \rvert)$*†	0.055 (0.058)	0.066 (0.069)	0.087 (0.102)	0.106 (0.132)
$R(\lvert {}^{\circ}F_A \rvert)$*†	0.253 (0.313)	0.364 (0.493)	0.362 (0.572)	0.498 (0.737)
$\langle\Delta(\Delta\phi)\rangle(^{\circ})$†‡	23.26 (24.20)	33.31 (34.50)	33.28 (34.81)	55.72 (55.27)
Reflections†	366 (380)	1293 (1355)	1107 (1199)	662 (759)

*$R=\Sigma\Sigma\lvert \lvert F_i \rvert - \langle \lvert F \rvert \rangle \rvert / \Sigma\Sigma \langle \lvert F \rvert \rangle$, where for each reflection, F_i is the i^{th} determination from MADLSQ. F_T and F_A are the normal structure factor for the total structure and that for anomalous scatterers alone, respectively. †The choice of maximum cut-off for $\lvert F_A \rvert$ has a pronounced effect on these values. The given values are obtained using a strict cut-off of 204, which is 1.5 times the mean value of $\lvert F_A \rvert$, $\Sigma(f_i^2)^{1/2}$, where the sum is over all anomalous scatterers in the unit cell and f is the atomic scattering factor at scattering angle of zero. The values in parentheses are obtained using a liberal cut-off of 680, which is 1.25 times the value of $F_A(000)$, Σf_i. ‡The average difference between independent determinations of phase differences in MADLSQ.

the structure for the methionyl protein. The final model yielded a crystallographic R value of 18.9% for 14311 reflections with $\lvert F \rvert > 2\sigma_{\lvert F \rvert}$ in the Bragg spacing range 10.0–1.8 Å. The R value for all 14891 measured reflections in that range was 19.5%. The model contains a total of 704 non-hydrogen atoms including 74 water molecules and eight atoms of the C-terminal residue, Glu156, with two alternative conformations. The mean value of individual isotropic B factors for all the non-hydrogen atoms of the model is 17.0 $Å^2$; excluding the solvent, the mean B factor is 14.5 $Å^2$. The model has root mean square deviations (rmsds) from ideal values of bond lengths of 0.007 Å and of bond angles of 1.54°. The restrained rmsds in B factors for bonded atoms are 1.9 $Å^2$ for main-chain atoms and 4.0 $Å^2$ for side-chain atoms. All residues in the model have (ϕ,ψ) values in 'allowed' regions of the Ramachandran plot [36].

Overall folding

The overall folding of $BCCP_{sc}$ is shown in Figure 1. The structure contains two sets of four antiparallel β strands. We designate these sequentially as β1–β8 for the segments comprising residues 81–85, 89–92, 107–109, 113–120, 123–130, 133–138, 144–146 and 150–156, as defined by Kabsch and Sander [37]. The biotinylated lysine (biocytin) is located at a hairpin turn between β4 and β5 near the middle of the polypeptide chain. A 'thumb'-like hairpin protrusion between residues 93 and 102 interacts with the biotinyl group.

The structure of $BCCP_{sc}$ is remarkably symmetric with three levels of internal symmetry. The first level of symmetry is that related to the amino acid sequence. The biocytin residue divides the sequence into N-terminal and C-terminal halves, which are similar. Structurally, the two halves are related by a quasi-dyad axis (Fig. 2). On superposing the two halves, the rms deviation for 28 pairs of Cα atoms from their mean position is 0.52 Å. The second level of symmetry is that observed from a purely structural perspective. The two sequential halves intertwine intimately such that, however one divides

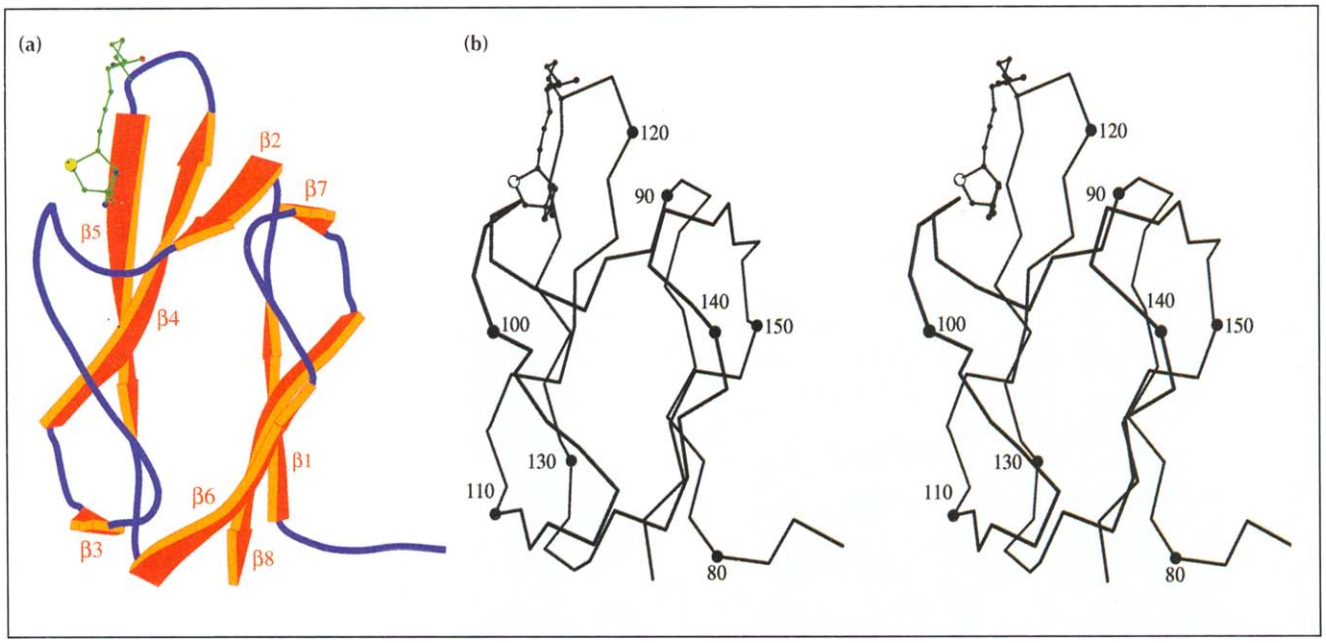

Fig. 1. Overall folding of $BCCP_{sc}$. **(a)** A ribbon diagram of $BCCP_{sc}$. The side chain of biocytin is shown in ball-and-stick representation with carbon, nitrogen, oxygen and sulfur atoms colored green, blue, red, and yellow respectively. The β strands are labeled. The direction of view is along the intramolecular quasi-dyad axis of symmetry. **(b)** A stereo view of the Cα trace of $BCCP_{sc}$ in the same orientation as in (a). Every tenth residue is numbered and is indicated by a filled circle on the Cα atom. (Figure prepared using MOLSCRIPT [66].)

them, quasi-dyad related halves have elements from both the N-terminal and C-terminal segments. Thus, strands β1, β6, β7 and β8 can be thought of as one structural half, but if divided differently, β1, β3, β6 and β8 define a structural half. The third level of symmetry relates to an additional internal element of symmetry within each of the structural halves (explained below).

The biotinyl domain has an unusual β-sandwich structure. Two β sheets, one consisting of β2, β4 and β5, and the other, consisting of β1, β6 and β8, form the sides of a β sandwich. Strand β7 and its counterpart β3 are hydrogen bonded to strands β2 and β6, respectively; but the curvatures of β2 and β6 are such that these interactions do not extend the sheets of the sandwich. Instead, the β7–β2 pair at one end and the β3–β6 pair at the other end form 'caps' to the sandwich structure. Accordingly, the structure may be described as a capped β sandwich with a quasi-dyad axis of symmetry.

The structure of $BCCP_{sc}$ is stabilized by a core of hydrophobic residues (Fig. 3). The hydrophobic character of these residues is conserved in the superfamily of lipoyl and biotin enzymes [38]. These residues are probably important structural determinants. As concluded from biochemical analysis [7], the side chain of the only cystyl residue in $BCCP_{sc}$, Cys116, is buried and is part of the hydrophobic core.

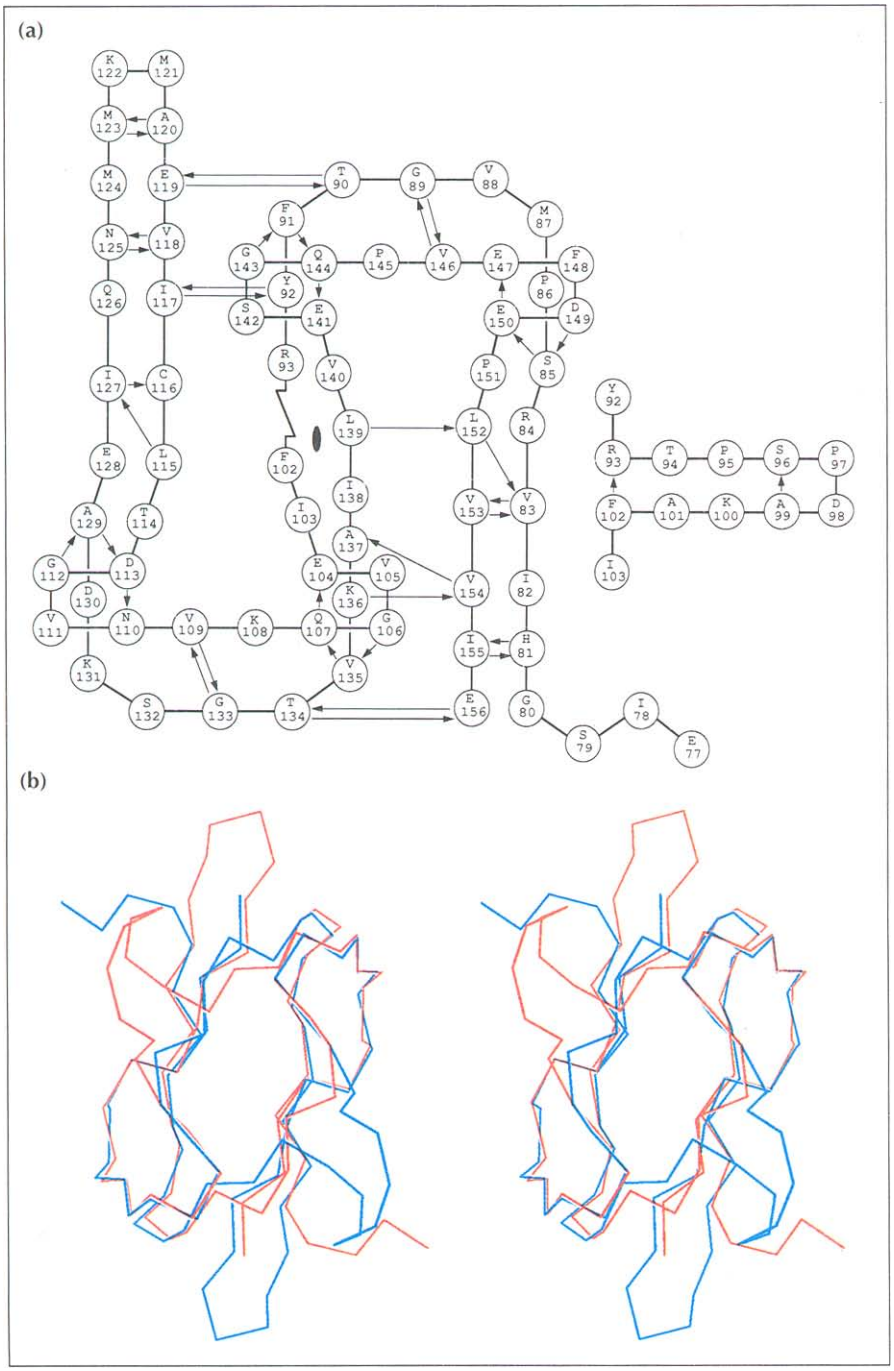

Fig. 2. Main-chain hydrogen bonding and the quasi-dyad symmetry in $BCCP_{sc}$. **(a)** Schematic diagram showing main-chain hydrogen bonding and the intramolecular dyad symmetry. For the sake of clarity, the loop from Arg93 to Phe102, which breaks this symmetry, is shown separately. Three residues at the N terminus and residues Ala120–Met123 do not obey this symmetry. Hydrogen bonds are indicated by arrows pointing from the hydrogen-bond donors (N) to the acceptors (O). **(b)** $BCCP_{sc}$ (red) superimposed on its own image (blue) after rotating by 180° about the quasi-dyad axis. The superposition matrix was calculated (with TOSS [65]) using 28 pairs of Cα atoms with separations ≤1.0 Å. This transformation gave an rmsd of 0.52 Å with a rotation of χ=179.4° and a translation of tχ=0.06 Å. Thus, Cα atoms of residues His81–Phe91 and Ile103–Glu119 in the N-terminal half were superimposed on those of residues Asn125–Val135 and Val140–Glu156 in the C-terminal half. When the cut-off is raised to 2.5 Å, 32 Cα pairs were matched at 0.86 Å by χ=179.8° and tχ=0.00. (Panel (b) was prepared using MOLSCRIPT [66]).

Hammerhead structure

Each symmetry-related half of $BCCP_{sc}$ contains a novel structural element that resembles a hammer (Fig. 4). The first hammer consists of residues 83–93 and 134–156 and its counterpart (subsequently referred to as the second hammer) includes residues 127–137 and 90–119, excluding the thumb-like protrusion (residues 94–101). Thus, the first hammer comprises a portion of β1, together with the whole of β2, β6, β7 and β8, and corresponding segments make up the second hammer. Strands β6 and β8 in the first hammer and β2 and β4 in the second can be thought of as the hammer 'handles'. The hammerhead itself is formed by a turn-strand-turn motif. Each hammerhead (residues 139–152 and 102–115 in the first and second hammers, respectively) contains two type II β turns [39] (Glu141–Gln144 and Glu147–Glu150 in the first and Glu104–Gln107 and Asn110–Asp113 in the second) that are separated by the respective strands β7 and β3. Segments from β1–β2 (83–93) and β5–β6 (127–137) buttress their associated hammerheads with several antiparallel hydrogen bonds. The two hammer-like units are remarkably similar (Fig. 4a).

The third level of symmetry, mentioned above, is the internal symmetry of the hammerhead structure. Segments within a hammerhead, including the buttressing strand, can be superimposed by appropriate transformations (Fig. 4b). The β turns in each of these portions, except for Glu141–Gln144, are stabilized by additional hydrogen bonds between main-chain nitrogen and side-chain oxygen atoms [40]. Thus, 104N, 110N and 147N make hydrogen bonds with Gln107Oε1, Asp113Oδ2 and Glu150Oε1 respectively. In this context, it is surprising that Gln144 is not similarly hydrogen bonded. It may be relevant that Glu141 and Ser142 are involved in lattice contacts in the crystal structure.

Biocytin residue

Biocytin is well defined in the electron-density map (Fig. 5a). The mean isotropic B factor for the 15 non-hydrogen atoms of the biotinyl moiety is 19.0 $Å^2$. Biocytin is located at a type I′ β turn connecting the N-terminal and C-terminal halves of $BCCP_{sc}$. The turn is stabilized by three hydrogen bonds (Fig. 5b). Seven amino acid residues (Tyr92, Thr94, Pro95, Ser96, Pro97, Ile117 and Met124) and four water molecules are located within 4.0 Å of the biotinyl moiety. Side-chain and the main-chain oxygen atoms of Thr94 form hydrogen bonds respectively with the *ureido* carbonyl (O2′) and N1′ atoms of biocytin respectively. Two of the four adjacent water molecules make hydrogen bonds with O2′ and N3′ atoms of biocytin.

Hydrogen bonds formed by the *ureido* ring are important among the interactions involving biocytin, such as those with BirA (holocarboxylase synthetase) and core streptavidin. BirA is a bifunctional protein that biotinylates a specific lysyl residue (Lys122) in BCCP and acts as a repressor of the biotin operon [25,28,29]. Core streptavidin is a protein that binds biotin extremely tightly [41]. The hydrogen bonds formed by the *ureido* ring of biocytin with BirA [29] and core streptavidin (FKA, A Pähler and WAH, unpublished data) in crystal structures of their complexes are compared with those in $BCCP_{sc}$ in Figure 6. In the streptavidin–biocytin complex, three tetrahedrally oriented hydrogen bonds are accepted by the *ureido* oxygen, O2′, to stabilize polarization of the *ureido* ring that results in a partial negative charge on O2′ and partial positive charges on N1′ and N3′. In contrast, O2′ forms only two hydrogen bonds in the structures of $BCCP_{sc}$ and the BirA–biocytin complex.

Comparison with other biotinyl and lipoyl domains

As anticipated by Brocklehurst and Perham [42] for another biotinyl domain, BCCP is indeed similar to the lipoyl domains of known structure. Superpositions with the two available experimental structures of lipoyl domains, namely the H-protein from the glycine decarboxylase complex of pea leaf (pGD) [17] and the lipoyl domain from pyruvate dehydrogenase of *Bacillus stearothermophilus* (bPD) [15], are shown in Figure 7. Both are somewhat less symmetric than BCCP. Nevertheless, all of

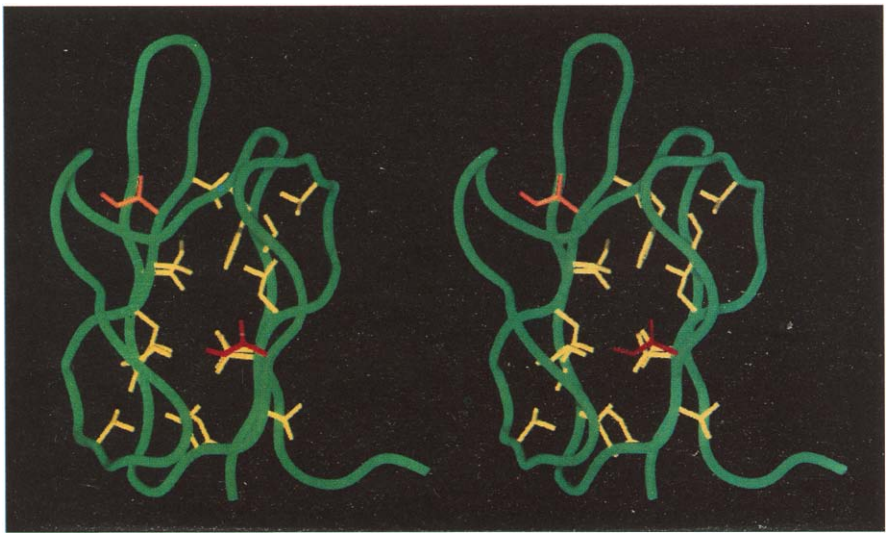

Fig. 3. Hydrophobic core residues in the structure of $BCCP_{sc}$. Side chains of amino acid residues in $BCCP_{sc}$ with fractional molecular surface area [48] of 0.25 or less are shown in yellow. They are Val83, Ser85, Phe91, Ile103, Val109, Leu115, Ile117, Val118, Ile127, Ala129, Val135, Val146, Leu152, Val153, Val154 and Ile155. In the absence of biotin moiety, the fractional surface area of Ile117 (orange) will be 0.30 instead of 0.14 in the present structure. Residue Ile138 (red) corresponds to a highly conserved hydrophobic residue in the biotin and lipoyl enzyme superfamily. In $BCCP_{sc}$, Ile138 has a fractional molecular surface area of 0.40 for the entire residue and 0.29 for its side chain. (Figure prepared using SETOR [67].)

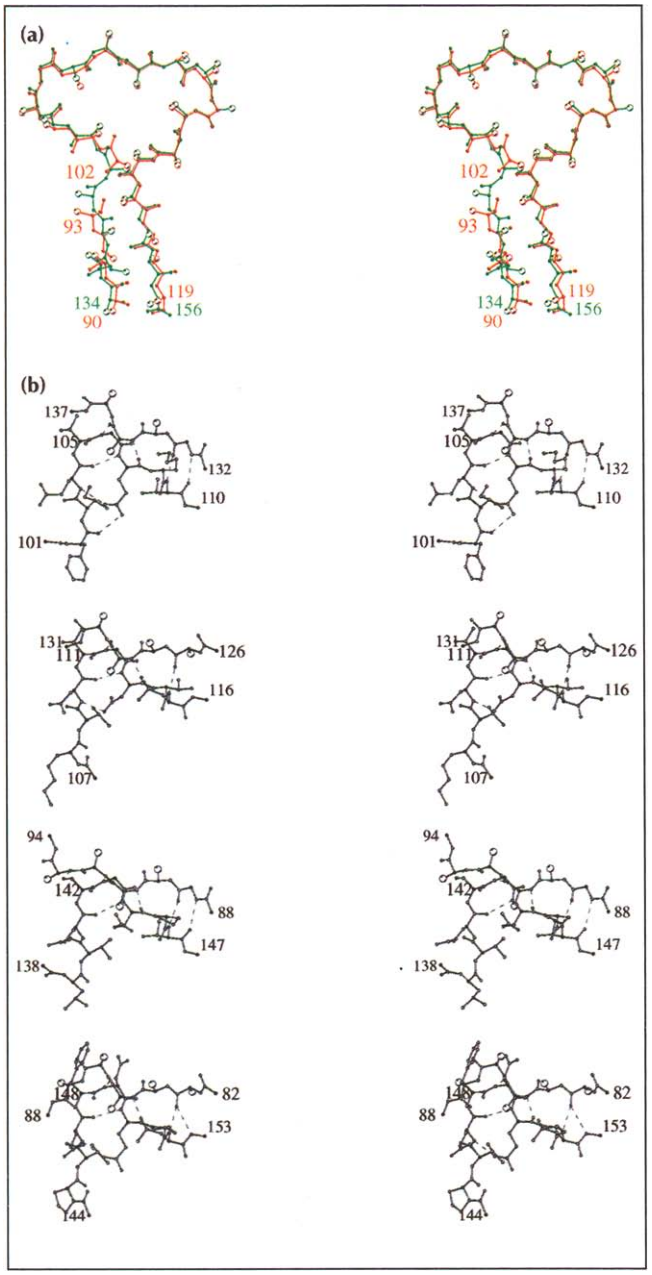

Fig. 4. Symmetry of the hammerhead structures. The two hammerheads in $BCCP_{sc}$ are related by the intramolecular quasi-dyad axis. In turn, each hammerhead has its own internal symmetry. **(a)** Superposition of the actual hammerhead-containing strands (without the buttressing segments) from the two hammerhead structures. Cα atoms of residues Thr134–Glu156 (green) are superimposed on those of residues Thr90–Glu119 (red), excluding the stretch Thr94–Ala101 belonging to the thumb-like protrusion. Side-chain atoms beyond Cβ are not shown. Cβ atoms are indicated by large open circles. **(b)** Elements of hammerhead structures displayed in similar orientations to show their internal symmetry. The region of the second hammerhead, consisting of residues Phe102–Val109 and Gly133–Lys136, is similar to that comprising residues Lys108–Leu115 and Ile127–Asp130. Within the second hammerhead, the Cα positions for the similar regions superimpose with an rmsd of 0.70 Å after the transformation of χ=146.9° and tχ=–1.56 Å. In the first hammerhead, the structure of residues Leu139–Val146 and Gly89–Arg93 are related to that of residues Pro145–Leu152 and Val83–Met87. As in (a), side-chain atoms beyond Cβ are not shown. Hydrogen bonds are indicated by dashed lines. (Figure prepared using MOLSCRIPT [66].)

the structures are clearly similar. $BCCP_{sc}$ has 50 Cα positions within 2.5 Å from corresponding contiguous residues in pGD, yielding an rmsd of 1.08 Å. A similar comparison between $BCCP_{sc}$ and bPD identifies 56 contiguous Cα positions within 3.0 Å separation and yields an rmsd of 1.58 Å (which falls to 48 positions with an rmsd of 1.31 Å if a 2.5 Å cut-off is applied). The lesser similarity of $BCCP_{sc}$ with the bPD model than with the pGD model probably reflects the relative accuracy of NMR *versus* X-ray structures [43] more than real differences. This is in keeping with the fidelity of the quasi-dyad symmetry observed in the two structures; the bPD model gives an rmsd of 1.49 Å for 26 pairs within 2.5 Å, whereas the $BCCP_{sc}$ model has an rmsd of 0.86 Å for 32 pairs.

The importance of the amino acid residues of the hydrophobic core is evident from the fact these residues are at least partially conserved in the superfamily of biotin and lipoyl enzymes [38] (Fig. 8). In addition to the conserved hydrophobic residues, a few conserved glycyl residues occupy structurally important sites in $BCCP_{sc}$ as well as in the two known structures of lipoyl domains [15,17]. Thus, glycyl residues at the corners of the β turns in the hammerheads — corresponding to Gly106, Gly112, Gly143 and Asp149 in $BCCP_{sc}$ — are often conserved. The substitution of aspartic acid for glycine sometimes observed (Fig. 8), is not uncommon in β turns [44]. It may be noted that the internal quasi-symmetry observed in the structure of $BCCP_{sc}$ is also present in related proteins (Figs 7,8).

The hammerhead probably constitutes the basic structural motif of biotinyl and lipoyl domains of a superfamily of enzymes [38]. Figure 7 shows that the H-protein from pGD [17] has one hammerhead motif and the lipoyl domain from bPD [15] has one complete and another partially formed hammerhead. Moreover, the structurally critical glycine and hydrophobic residues corresponding to these hammerhead structures are largely conserved in the superfamily.

In some alignments, *E. coli* BCCP has been considered to lack a proline residue at a position ~28 residues N-terminal to biocytin, which is well conserved in other biotin enzymes. The structure-based alignment in Figure 8 shows that Pro86 in BCCP corresponds to that 'missing' proline. However, along with Browner *et al.* [45], we disagree with the proposition [46] that this residue could be the point of flexibility required for translocation of the biotinyl moiety between the active sites of biotin carboxylase and carboxyl transferase.

Functional interactions of BCCP

It is generally understood that the alternative conformational states of the biocytin residue are important in translocating the biotinyl moiety between the active sites of the biotin carboxylase and carboxyl transferase subunits. The long side chain of biocytin may afford this residue considerable conformational flexibility [47]. In

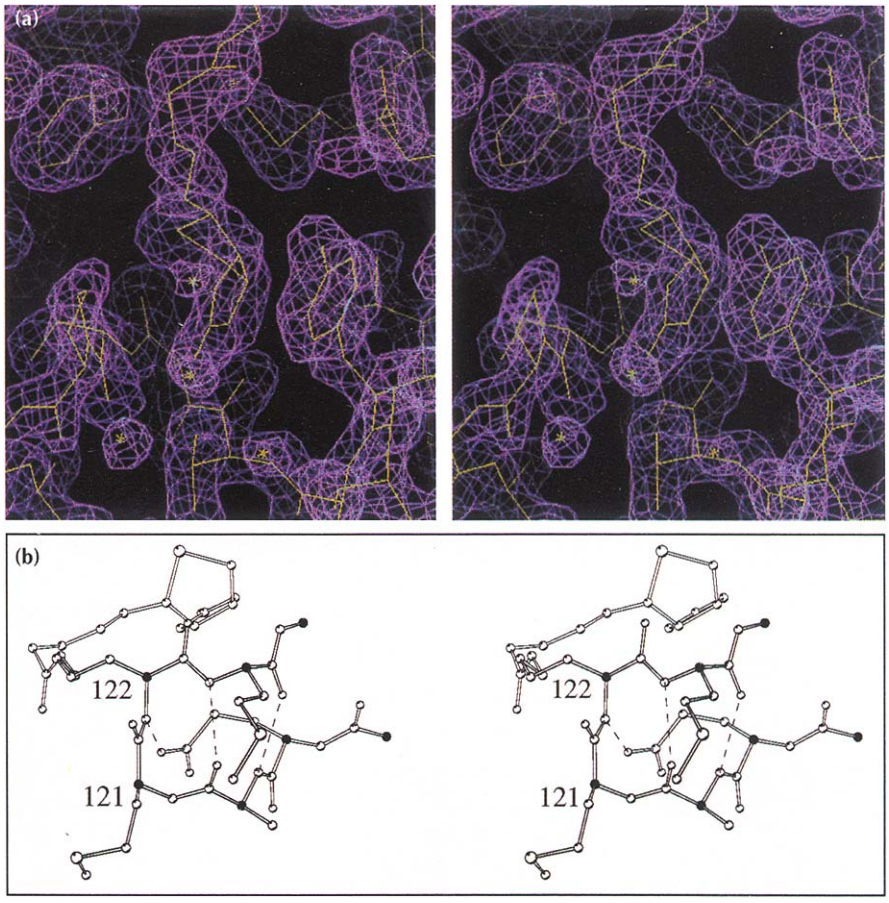

Fig. 5. Interactions of biocytin. **(a)** A portion of the final $2|F_o-F_c|$ electron-density map, contoured at the level of 1.2σ, showing the well-defined biotinyl moiety. Water molecules are indicated by stars. The puckered ring of Pro97 may be seen to the left of the *ureido* ring of biotin. The side chain of Phe148 belonging to a monomer related by crystal symmetry may be seen near the top-left corner. The main-chain and side-chain oxygen atoms of Thr94 (below the *ureido* group) form hydrogen bonds with N1′ and O2′ atoms of biocytin respectively. The angle at the hydrogen atom in the bond between Thr94Oγ1 and O2′ of biocytin is only 111°, however. O2′ and N3′ are each involved in a hydrogen bond with water molecules. (Figure prepared using O [59].) **(b)** A stereoview of the type I′ β turn [39] containing the biocytin residue (122). This turn is stabilized by three hydrogen bonds (shown as dashed lines) including one between the main-chain nitrogen atom of biocytin and Glu119Oε2. (Figure prepared using MOLSCRIPT [66].)

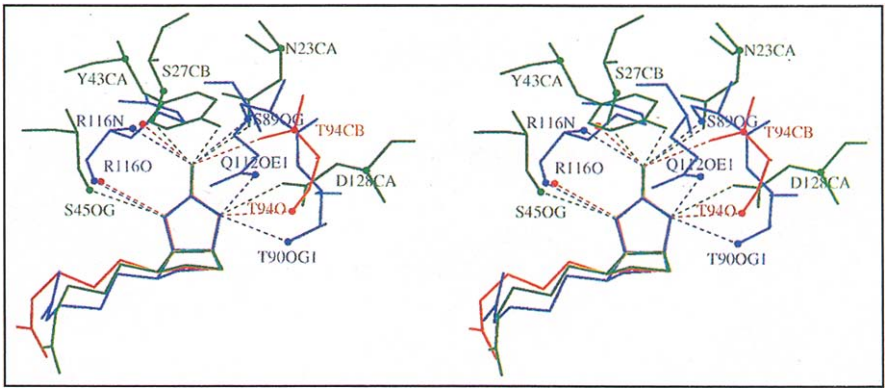

Fig. 6. Hydrogen bonds (dashed lines) formed by the *ureido* ring of biocytin with three different proteins. The *ureido* ring in the structure of $BCCP_{sc}$ (red) was superimposed with those in the structures of the complexes of biocytin with streptavidin (green) and BirA (blue) [29]. The unconnected filled red circles are water molecules in the structure of $BCCP_{sc}$. Selected atoms, highlighted with filled circles, are labeled to help identify the amino acid residues. (Figure prepared using MOLSCRIPT [66].)

this context it is noteworthy that the β turn containing biocytin may assume another conformation, changing from the type I′ [39] conformation observed in this crystal structure to a type I conformation. Indeed, in the structures of the two known lipoyl domains that are related to $BCCP_{sc}$, lipoylated H-protein from pGD [17] and the lipoyl domain from bPD [15], the corresponding β turns have a type I conformation. Conversion between type I′ and type I conformations may be involved in shuttling the biotinyl moiety between the active sites of the biotin carboxylase and carboxyl transferase components of acetyl-CoA carboxylase. A similar mechanism may also pertain for lipoyl domains.

It is now clear that biotinyl and lipoyl domains are quite similar. This poses an interesting question: how do the biotinylating and lipoylating enzymes determine whether a particular lysyl residue will be biotinylated or lipoylated [16]? Although it is known that the amino acid sequence in the immediate vicinity of the particular lysyl residue is not sufficient for recognition by the biotinylating enzyme BirA [25,28], the adjacent methionyl residues are important. In biotinyl domains, the lysyl residue of the conserved sequence Ala-Met-Lys-Met is biotinylated. The corresponding consensus sequence for lipoylation in lipoyl domains is X-Asp-Lys-Ala. Given the specificity of the reactions, the methionyl residues can be expected to be involved in recognition by the biotinylating enzyme. In this context, it is interesting to analyze the accessible surface areas of methionyl residues in the structure. The fractional molecular surface areas [48] for the side chains of the four methionines in $BCCP_{sc}$, Met87, Met121,

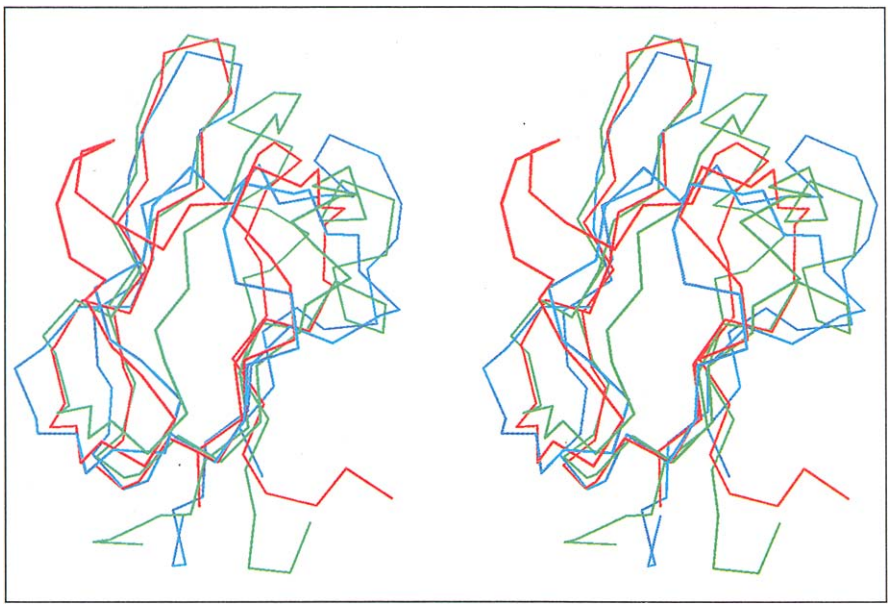

Fig. 7. Superposition of the Cα backbones of $BCCP_{sc}$ (red), lipoylated H-protein from pGD (green) [17] and the lipoyl domain from bPD (blue) [15]. In the case of pGD, only residues 19–108 were used in this superposition. See Figure 8 for the residues of pGD and bPD used to calculate the superposition matrices. (Figure prepared using MOLSCRIPT [66].)

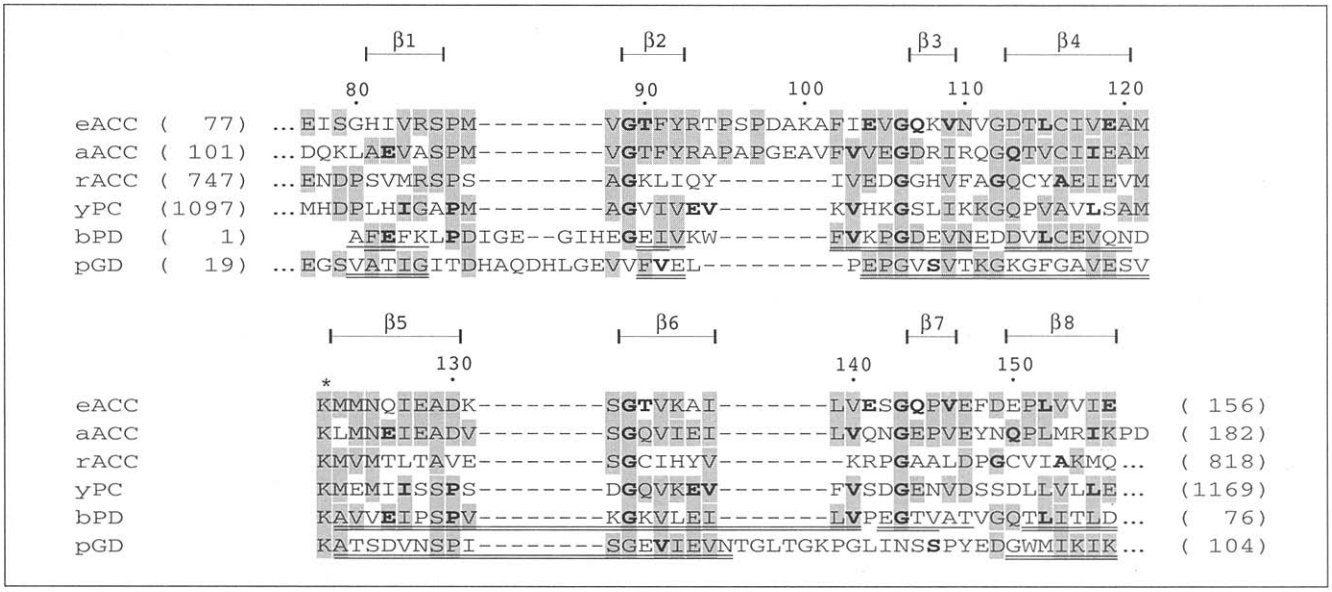

Fig. 8. Structure-based sequence alignment. N-terminal (top panel) and C-terminal (bottom panel) halves of $BCCP_{sc}$ (eACC) are aligned with those of the biotinyl domains from acetyl-CoA carboxylases of *Anabaena sp.* [68] (aACC) and rat [69] (rACC), yeast pyruvate carboxylase [14] (yPC), and lipoyl domains from *B. stearothermophilus* pyruvate dehydrogenase multienzyme complex [15] (bPD) and H-protein of pea leaf glycine cleavage system [17] (pGD). The sequence numbers of the first and last residues in each protein are given in parentheses and every tenth amino acid residue in $BCCP_{sc}$ is numbered above its sequence. An asterisk marks the position of a conserved lysine (Lys122 in $BCCP_{sc}$) that is the site of biotinylation/lipoylation in each of these enzymes. Identical amino acid residues are shaded. Residues that are identical in the N-terminal and C-terminal halves of a particular protein are indicated by bold type. This similarity is notably weak in pGD. The amino acid residues in bPD and pGD whose Cα atoms superpose within 2.5 Å with the corresponding atoms of $BCCP_{sc}$ are underlined twice; those in bPD which superpose within 3.0 Å are underlined once. All underlined residues were used to calculate the superposition matrices used in Figure 7. The β strands in the structure of $BCCP_{sc}$ are indicated.

Met123 and Met124, are 0.45, 0.97, 0.88 and 0.51 respectively. Thus, Met121 and Met123 belonging to the Ala-Met-Lys-Met sequence are more exposed and hence more available to interact with the biotinylating enzyme. It addition, mass spectroscopic analysis showed that two of the four methionyl residues in $BCCP_{sc}$ were susceptible to oxidation. Probably, they are the exposed residues, Met121 and Met123, as Met87 forms part of the hydrophobic core and Met124 contacts the biotinyl moiety.

Residues Thr94–Ala101 form a protruding thumb in the structure of $BCCP_{sc}$ that breaks the intramolecular dyad symmetry (Fig. 2a). This structural element is divided into two halves by a type I β turn at its center. Four amino acid residues belonging to the N-terminal half of this thumb, Thr94, Pro95, Ser96 and Pro97, are among the seven residues close to the biotinyl moiety. Thr94 is the only amino acid residue that forms hydrogen bonds with the biotinyl moiety. Thus, the thumb, which

appears to be part of an insertion (Fig. 8), tethers the biotinyl moiety to the body of BCCP. This confirms the proposition made by Fall *et al.* [11], on the basis of circular dichroism studies, that the biotin prosthetic group is partially buried in the surface of $BCCP_{sc}$ rather than swinging free at the end of the lysyl side chain. A similar observation was made in the case of the H-protein from pGD in which the lipoate cofactor binds into a cleft at the protein surface [49]. However, it may be noted that N1′ carboxylation would disrupt the hydrogen bond between Thr94O and N1′. It is not clear whether the thumb is involved in interactions of BCCP with the biotinylating enzyme, BirA, or with carboxylase or carboxyl transferase subunits of acetyl-CoA carboxylase.

Assembly of the different subunits of *Propionibacterium shermanii* transcarboxylase does not affect biotinylation of the 1.3S subunit [50]. Biotinylation of BCCP is also thought to be possible both before and after assembly into acetyl-CoA carboxylase. Biotinylation studies on *E. coli*, chicken and human apoacetyl-CoA carboxylases [25–27], and on fusion proteins containing biotinyl domains [23,24], support this conclusion. This suggests that the biocytin β hairpin, notably Lys122, must be near the surface to facilitate biotinylation by BirA as well as the intersubunit translocation of the biotinyl moiety. Interactions of BCCP with biotin carboxylase and carboxyl transferase appear to be independent of its interaction with BirA. Therefore, BCCP probably has three separate regions of interaction.

Biological implications

Biotinyl domains are at the mechanistic heart of all classes of biotin-dependent enzymes. Acetyl-coenzyme A (CoA) carboxylase is a multicomponent biotin enzyme that catalyzes one of the regulated steps (the first committed step) in the biosynthesis of long-chain fatty acids. It has three easily separable functional components: biotin carboxylase, which carboxylates the biotinyl group in the presence of bicarbonate, ATP and Mg^{2+}; carboxyl transferase, which transfers the carboxyl group from biotin to acetyl-CoA to form malonyl CoA; and biotin carboxyl carrier protein (BCCP), which serves to shuttle the biotin prosthetic group between the two active sites. Thus, BCCP plays a central role both in the carboxylation and in the transcarboxylation reaction steps.

The structure of a fragment of *Escherichia coli* BCCP presented here is the first crystal structure of a biotinyl domain. The fold, which can be described as a capped β sandwich, has a quasi-dyad axis of symmetry that divides the domain into two halves. Each half of BCCP has a hammerhead motif that may be regarded as the basic structural motif of both biotinyl and the related lipoyl domains. The biotinylated lysine (biocytin) is located at a hairpin β turn connecting the N-terminal and C-terminal halves. In acetyl-CoA carboxylase, this β turn will probably be near the surface thereby facilitating biotinylation by BirA, the holocarboxylase synthetase of *E. coli* [25,28]. The biotinyl moiety in BCCP must reach both the biotin carboxylase and the carboxyl transferase active sites, and yet BCCP is expected to be biotinylated even after being assembled into the apocarboxylase. Thus, it seems that distinct regions of BCCP must interact with BirA, biotin carboxylase and carboxyl transferase.

Impaired biotinylation is responsible for multiple carboxylase deficiency, which in humans causes symptoms ranging from skin rashes to seizures and coma. Holocarboxylase synthetase catalyzes the attachment of biotin to biotinyl domains in a reaction that is conserved across species boundaries from bacteria to mammals. While the synthetases can biotinylate apocarboxylases, they require only the biotinyl domain for the specific covalent attachment of biotin. The structural biology of the system in bacteria, involving the BirA and BCCP proteins, now seems relevant to the condition in man.

Materials and methods

Cloning and expression

Construction of the expression vector: The gene encoding BCCP was isolated by PCR using oligonucleotide primers designed on the basis of its known DNA sequence [51,52] and *E. coli* chromosomal DNA as the template. It was then introduced into the pRSETB T7 expression vector (Invitrogen Corp., San Diego, CA). This construct provided a metal-binding N-terminal polyhistidyl tag to facilitate purification of the expressed protein. The fidelity of PCR was established by DNA sequence analysis.

Expression and purification of recombinant $BCCP_{sc}$: We transformed BL21 (DE3) pLys S [53] cells with the recombinant plasmid and expressed BCCP in the presence of a large excess of biotin (6 mg L^{-1}). The recombinant BCCP was isolated using a Ni-NTA column (QIAGEN Inc., Chatsworth, CA) under non-denaturing conditions employing an imidazole gradient (0.01–0.5 M).

Expression and purification of Se-met BCCP: The recombinant plasmid was used to transform the methionine auxotroph, DL41 [54], carrying the prophage DE3 (kindly provided by Thomas Peat). We grew the transformed cells in 6 L of a Se-met medium containing a large excess of biotin, where they did not thrive. Unlike natural BCCP, we could not appreciably induce expression of Se-met BCCP by adding isopropyl-thio-β-D-galactoside. Nonetheless, these cells did express minute amounts of Se-met BCCP. Employing a protocol similar to that used to purify natural BCCP, we isolated and purified Se-met BCCP.

Preparation and characterization of $BCCP_{sc}$

Proteolysis and purification: We digested recombinant BCCP and Se-met BCCP with subtilisin Carlsberg [8] to produce $BCCP_{sc}$ and Se-met $BCCP_{sc}$, which were purified using FPLC

MonoQ columns (Pharmacia, Piscataway, NJ). While purifying Se-met $BCCP_{sc}$, we used both dithiothreitol (DTT) and methionine to prevent oxidation of Se-met residues. From 2×6 L preparations, we obtained a total of about 900 μg of Se-met $BCCP_{sc}$ as estimated on the basis of UV absorbance.

Characterization of $BCCP_{sc}$: Authenticity of $BCCP_{sc}$ was established by determining its N-terminal amino acid sequence and molecular mass. Dr David King (HHMI/Tjian Laboratory, University of California, Berkeley, CA) analyzed $BCCP_{sc}$ by mass spectrometry and showed that two of the four methionines in it were susceptible to oxidation.

$BCCP_{sc}$ was analyzed for the presence of an exposed biotinyl moiety by running non-denaturing polyacrylamide gels in the presence and absence of the biotin-binding protein, core streptavidin. The shift in electrophoretic mobility in the presence of streptavidin indicated the presence of an exposed biotinyl moiety in $BCCP_{sc}$.

Crystallization

Crystallization of natural $BCCP_{sc}$: Diffraction quality, plate-like crystals of $BCCP_{sc}$ were obtained at 20°C by vapor diffusion of hanging drops of its solution at a concentration of about 10 mg ml^{-1} against 0.96 M sodium citrate in 7.3 mM borate buffer at pH 8.5 containing approximately 3 mM DTT. In order to prepare a mercury derivative, we treated the protein with 0.9 mM ethyl mercury phosphate. Thereafter, the protein yielded crystals which were thicker than those obtained earlier. These crystals, which were considered to be derivatized, were used for further analysis.

Crystallization of Se-met $BCCP_{sc}$: We tried to crystallize the Se-met analog under conditions similar to those used in the case of natural $BCCP_{sc}$, but with one difference — the crystallization drops were sealed anaerobically [55] to prevent oxidation of Se-met residues. Although a few small crystals were obtained in the initial trials after seeding [56], crystallization trials set up after the protein was two weeks old did not yield any more crystals.

Large, single crystals of Se-met $BCCP_{sc}$ suitable for structural studies were finally obtained by dialysis using a newly developed crystallization device. Droplets of a 10 mg ml^{-1} solution of $BCCP_{sc}$ (2–4 μl) were dialyzed against 2 ml of 0.96 M sodium citrate/2 mM $MgSO_4$/1 mM DTT/7.3 M borate buffer at pH 9.7. As for the vapor diffusion experiments, the crystallization device was sealed within an anaerobic chamber [55]. Several large crystals (0.16×0.86×0.07 mm^3) of Se-met $BCCP_{sc}$ were obtained using this novel device from less than 200 μg of the protein. Details of this device will be published elsewhere.

Data collection and processing

Crystal characterization: We characterized these crystals in a rather unconventional manner using the San Diego multiwire area detector system. Because the commercially available automatic indexing programs failed, we first calculated the goniometer angles in equatorial bisecting geometry corresponding to the measured reflections and then derived the unit cell parameters and the alignment angles for data collection employing methods developed for four-circle diffractometers (FKA and WAH, unpublished data).

Data collection from the $BCCP_{sc}$ crystal: Although cryoprotecting conditions had been found, for technical reasons data were measured at room temperature. We collected a data set complete to 1.8 Å Bragg spacings from a single crystal of approximate dimensions 0.28×0.30×0.13 mm^3 using radiation of wavelength 0.9803 Å at HHMI beamline X4A at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory.

MAD data collection: We collected MAD data from Se-met $BCCP_{sc}$ crystals on imaging plates at beamline X4A at NSLS. The X-ray absorption spectrum near the K-shell edge of the Se-met crystal was measured first. This was used to select monochromator settings for the peak (maximum f″), the inflection point (minimum f′) and two remote energies at 101 eV below and 139 eV above the absorption edge. The crystals were aligned such that reflections related by mirror symmetry could be recorded simultaneously. The wavelength of radiation was changed after each exposure to minimize systematic errors due to radiation damage. Expecting modest radiation sensitivity as for natural $BCCP_{sc}$, the Se-met data were also measured at room temperature. Unfortunately, however, radiation damage was more appreciable in this case. Four different crystals were used to complete the data collection.

Data Processing: Oscillation data from the Fuji phosphorous imaging plates were read out with a Fuji scanner. These images were indexed and profile-fitted integrated intensities were obtained by using DENZO [57]. The data were reduced and scaled with programs of CCP4 [58] and MADSYS [31] packages (Tables 1–3).

Processing of the data from natural $BCCP_{sc}$ crystals showed that the data did not have much anomalous signal, indicating that the crystal was not a mercury derivative. This 1.8 Å resolution data set was used for the final refinement of the structure of natural $BCCP_{sc}$ (Table 3).

Structure determination using MAD phasing

Analysis of MAD data: Using the conventional programs MADLSQ and MERGIT, we calculated the moduli of normal structure factors, $|F_A|$ and $|F_T|$ (corresponding to the anomalous scatterers and the entire structure, respectively) and the phase difference, $(\phi_T - \phi_A)$. The statistics of phasing indicated that phasing could be reliable up to a resolution of 2.5 Å (Table 2). As we could not interpret the Patterson map computed from $|F_A|^2$ coefficients, selenium positions were obtained using the direct methods program MULTAN [33]. The positional and thermal parameters of these atoms were then refined to an R value of 43.0% against 1647 reflections in the Bragg spacing range 20.0–2.9 Å using the program ASLSQ (WAH, unpublished program).

We could not interpret the electron-density maps computed at different resolutions using the correct coordinates of selenium atoms. Automatic solvent flattening [35], performed using the CCP4 [58] density modification program DM, and estimation of phase probabilities using the program MADABCD [34], did not significantly improve the situation.

Modified MAD phasing technique: At this point, we analyzed the agreement of the $|F_A|$s calculated by MADLSQ and those calculated on the basis of the refined coordinates of selenium atoms. A number of estimates of $|F_A|$s from MADLSQ, and particularly those with amplitudes greater than three times the mean value of $|F_A|$, had very poor agreements. Table 2 demonstrates that lower maximum cut-off values for $|F_A|$ significantly improved the phasing statistics. Use of minimum and maximum cut-off for $|F_A|$, and deletion of 56 outlier

Table 3. Data statistics for the natural $BCCP_{sc}$ crystal.

Shell (Å)	10.0–5.0	5.0–2.8	2.8–2.3	2.3–2.0	2.0–1.8	10.0–1.8
Data collection						
Measurements*	1116	6497	6393	7324	7355	28655
Unique reflections	591	3338	3228	3746	3988	14891
Friedel pairs†	225	1445	1453	1691	1798	6612
Redundancy‡	1.9	1.9	2.0	2.0	1.8	1.9
Completeness (%)§	94.0	97.6	97.3	96.5	94.7	96.4
R_{sym} (%)	1.5	2.1	3.9	5.4	9.2	3.2
Refinement						
Unique reflections	590	3315	3148	3616	3642	14311
Friedel pairs†	225	1440	1426	1646	1656	6393
Completeness (%)§	93.8	96.9	94.9	93.2	86.5	92.6
R_{ref} (%)#	25.3	15.8	18.7	18.3	21.0	18.9

*Reflections with intensities less than or equal to zero were ignored. †The indicated number of Friedel pairs are included among the unique reflections. ‡Redundancy is the average number of measurements per unique reflection. §Completeness of data is the ratio of the number of unique reflections measured to the total number theoretically possible. Data completeness after selecting reflections with amplitudes $F>2\sigma_{|F|}$. #$R_{ref}=\Sigma|F_{obs}(h)-F_{calc}(h)|/\Sigma F_{obs}(h)$.

reflections improved the agreement appreciably to 2.3 Å resolution. Thus, by deleting outlier estimates of $|F_A|$ it was possible to refine the positional and thermal parameters of selenium atoms to an R value of 33.8% against 2522 reflections in the Bragg spacing between 20.0 Å and 2.5 Å using the program ASLSQ. However, electron-density maps, computed using these selected reflections or those obtained by using an $|F_A|$ cut-off of 400 (approximately three times the theoretical mean value of $|F_A|$), could not be interpreted.

Then, we modified the MAD phasing procedure. Structure factors for normal scattering from the entire protein structure including their phases ($|F_T|$ and ϕ_T) were determined afresh, making use of ϕ_As and scaled values of $|F_A|$s computed on the basis of the refined parameters of the anomalous scatterers. The reasoning behind this modification was as follows. The parameters of selenium atoms were known reliably. Therefore, F_As calculated based on them were dependable. If these F_As did not agree with those from MADLSQ, the results of MADLSQ could be improved by incorporating the better estimates of F_As. Hence, the MAD equations were reformulated in such a way that there were only two unknowns, $|F_T|$ and ϕ_T, instead of three, $|F_A|$, $|F_T|$ and $(\phi_T-\phi_A)$ as in the original MAD formulation [30,31]. A program along the lines of MADLSQ was developed to determine the structure factors corresponding to the total structure ($|F_T|$s and ϕ_Ts). This enabled us to increase the completeness level from 85% to 99.7% (3249 reflections) in the Bragg spacing range 20.0–2.5 Å. Details of this method will be published elsewhere.

After the refinement of the structure, we compared the phases calculated by the conventional and the modified MAD procedures against those computed based on the final model, wherein the methionyl sulfur atoms were replaced by selenium atoms. The mean phase error for 2747 reflections phased by the conventional MAD method with the strict $|F_A|$ cut-off was 49.5°, while the mean phase error for these reflections phased by the modified procedure was 50.2°. For 2918 reflections phased with the liberal $|F_A|$ cut-off, the mean phase errors for the conventional and modified procedures were 51.9° and 51.4° respectively. Thus, the main advantage of the modification in this case seems to have been to increase the level of completeness.

Determination of the structure of Se-met $BCCP_{sc}$: Automatic solvent flattening was repeated after the modified MAD phasing to give a free R value drop from 59.0% to 42.6% in 80 cycles. An electron-density map was then computed at 2.5 Å resolution in which a model for Se-met $BCCP_{sc}$ was traced using the computer graphics program O [59]. Recognition of similarity to the lipoyl domain of *Bacillus stearothermophilus* pyruvate dehydrogenase [15] helped to confirm the chain trace.

Towards the end of model building, a further improved electron-density map was computed. This was accomplished by combining 2.3 Å resolution phases calculated using the modified MAD procedure with SIRAS estimates of phases obtained by using the CCP4 [58] program MLPHARE, considering the Se-met analog as a derivative of natural $BCCP_{sc}$. As earlier, the electron-density map was computed after solvent flattening. Although this map was better than the previous one, it was not much used as model building was already nearly complete.

Refinement and analysis of the structure of natural $BCCP_{sc}$

The model of Se-met $BCCP_{sc}$ was refined against the 1.8 Å data from natural $BCCP_{sc}$ crystals using X-PLOR [60,61]. Despite the relatively small anomalous signal from the six sulfur atoms in the structure, considering the quality of data (R_{sym}=3.2% to 1.8 Å resolution) Bijvöet-related reflections were not merged. The program FRODO [62] was used for refitting of the model during refinement.

Hydrogen atoms were located using HAFFIX (WAH, unpublished program). The positions of the rotatable hydrogen atoms of hydroxyl groups were fixed at positions closest to potential hydrogen-bond acceptors so as to enhance hydrogen bonding. Generally, interactions between potential acceptor and donor atoms at distances less than the sum of their van der Waals radii with an angle at hydrogen of 120° or more were accepted as hydrogen bonds [40]. For hydrogen bonds involving water molecules, only a distance criterion was used.

PROCHECK [63] was used to estimate the quality of the refined model. All of its 'quality indicators' are within one standard deviation from the mean of well refined structures, all but one being better than the mean. Molecular surface areas

were calculated with the MS package [64]. Molecular superpositions were carried out with the program TOSS [65].

Acknowledgements: We thank C Ogata, X Zhu, H Wu, Y Liu, C Bingman and SR Hubbard for assistance in data collection; BW Low, M Eisenberg, DJ Leahy, J Ngai, PD Kwong and L Shapiro for useful discussions and A Athappilly for help in preparing this manuscript. This work was supported in part by NIH grant GM34102. Beamline X4A at NSLS, a DOE facility, is supported by the Howard Hughes Medical Institute. The atomic coordinates are being deposited in the Protein Data Bank.

References

1. Knowles, J.R. (1989). The mechanism of biotin-dependent enzymes. *Annu. Rev. Biochem.* **58**, 195–221.
2. Wakil, S.J., Stoops, J.K. & Joshi, V.C. (1983). Fatty acid synthesis and its regulation. *Annu. Rev. Biochem.* **52**, 537–579.
3. Lane, M.D., Moss, J. & Polakis, S.E. (1974). Acetyl-CoA carboxylase. *Curr. Top. Cell. Regul.* **8**, 139–195.
4. Nervi, A.M., Alberts, A.W. & Vagelos, P.R. (1971). Acetyl-CoA carboxylase. 3. Purification and properties of a biotin carboxyl carrier protein. *Arch. Biochem. Biophys.* **143**, 401–411.
5. Fall, R.R., Nervi, A.M., Alberts, A.W. & Vagelos, P.R. (1971). Acetyl-CoA carboxylase: isolation and characterization of native biotin carboxyl carrier protein. *Proc. Natl. Acad. Sci. USA* **68**, 1512–1515.
6. Guchhait, R.B., Moss, J., Sokolski, W. & Lane, M.D. (1971). The carboxyl transferase component of acetyl-CoA carboxylase: structural evidence for intersubunit translocation of the biotin prosthetic group. *Proc. Natl. Acad. Sci. USA* **68**, 653–657.
7. Fall, R.R. & Vagelos, P.R. (1972). Acetyl coenzyme A carboxylase. Molecular forms and subunit composition of biotin carboxyl carrier protein. *J. Biol. Chem.* **247**, 8005–8015.
8. Fall, R.R. & Vagelos, P.R. (1973). Acetyl coenzyme A carboxylase. Proteolytic modification of biotin carboxyl carrier protein. *J. Biol. Chem.* **248**, 2078–2088.
9. Guchhait, R.B., Polakis, S.E., Dimroth, P., Stoll, E., Moss, J. & Lane, M.D. (1974). Acetyl coenzyme A carboxylase system of *Escherichia coli*. Purification and properties of the biotin carboxylase, carboxyltransferase, and carboxyl carrier protein components. *J. Biol. Chem.* **249**, 6633–6645.
10. Fall, R.R. & Vagelos, P.R. (1975). Biotin carboxyl carrier protein from *Escherichia coli*. *Methods Enzymol.* **35**, 17–25.
11. Fall, R.R., Glaser, M. & Vagelos P.R. (1976). Acetyl coenzyme A carboxylase. Circular dichroism studies of *Escherichia coli* biotin carboxyl carrier protein. *J. Biol. Chem.* **251**, 2063–2069.
12. Sutton, M.R., Fall, R.R., Nervi, A.M., Alberts, A.W., Vagelos, P.R. & Bradshaw, R.A. (1977). Amino acid sequence of *Escherichia coli* biotin carboxyl carrier protein (9100). *J. Biol. Chem.* **252**, 3934–3940.
13. Waldro, G.L., Rayment, I. & Holden, H.M. (1994). Three dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. *Biochemistry* **33**, 10249–10256.
14. Lim, F., Morris, C.P., Occhiodoro, F. & Wallace, J.C. (1988). Sequence and domain structure of yeast pyruvate carboxylase. *J. Biol. Chem.* **263**, 11493–11497.
15. Dardel, F., Davis, A.L., Laue, E.D. & Perham, R.N. (1993). Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037–1048.
16. Green, J.D.F, Laue, E.D., Perham, R.N., Ali, S.T. & Guest, J.R. (1995). Three-dimensional structure of a lipoyl domain from the dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328–343.
17. Pares, S., Cohen-Addad, C., Seiker, L., Neuburger, M. & Douce, R. (1994). X-ray structure determination at 2.6 Å resolution of a lipoate-containing protein: the H-protein of the glycine decarboxylase complex from pea leaves. *Proc. Natl. Acad. Sci. USA* **91**, 4850–4853.
18. Burri, B.J., Sweetman, L. & Nyhan, W.L. (1981). Mutant holocarboxylase synthetase: evidence for the enzyme defect in early infantile biotin-responsive multiple carboxylase deficiency. *J. Clin. Invest.* **68**, 1491–1495.
19. Suzuki, Y., *et al.*, & Narisawa, K. (1994). Isolation and characterization of mutations in the human holocarboxylase synthetase cDNA. *Nature Genetics* **8**, 122–128.
20. Wolf, B. & Heard, G.S. (1989). Disorders of biotin metabolism. In *The Metabolic Basis of Inherited Disease*. (Scriver, C.R., Beaudet, A.L., Sly, W.S. & Valle, D. eds), pp. 2083–2103, McGraw-Hill, NY.
21. McAllister, H.C. & Coon, M.J. (1966). Further studies on the properties of liver propionyl coenzyme A holocarboxylase synthetase and the specificity of holocarboxylase formation. *J. Biol. Chem.* **241**, 2855–2861.
22. Bai, D.-H., Moon, T.-W., López-Casillas, F., Andrews, P.C. & Kim, K.-H. (1989). Analysis of the biotin-binding site on acetyl-CoA carboxylase from rat. *Eur. J. Biochem.* **182**, 239–245.
23. Li, S.-J. & Cronan, J.E., Jr (1992). The gene encoding the biotin carboxylase subunit of *Escherichia coli* acetyl-CoA carboxylase. *J. Biol. Chem.* **267**, 855–863.
24. Cronan, J.E., Jr (1990). Biotination of proteins *in vivo*: a post-translational modification to label, purify and study proteins. *J. Biol. Chem.* **265**, 10327–10333.
25. Eisenberg, M.A., Prakash, O. & Hsiung, S.-C. (1982). Purification and properties of the biotin repressor: a bifunctional protein. *J. Biol. Chem.* **264**, 15167–15173.
26. Murthy, P.N.A. & Mistry, S.P. (1974). Synthesis of acetyl coenzyme A holocarboxylase *in vitro* by a cytosolic preparation from chicken liver. *Proc. Soc. Exp. Biol. Med.* **147**, 114–117.
27. Packman, S., *et al.*, & McKay, C. (1984). Acetyl CoA carboxylase in cultured fibroblasts: differential biotin dependence in the two types of biotin-responsive multiple carboxylase deficiency. *Am. J. Hum. Genet.* **36**, 80–92.
28. Barker, D.F. & Campbell, A.M. (1981). The *birA* gene of *Escherichia coli* encodes a biotin holoenzyme synthetase. *J. Mol. Biol.* **146**, 451–467.
29. Wilson, K.P., Shewchuk, L.M., Brennan, R.G., Otsuka, A.J. & Matthews, B.W. (1992). *Escherichia coli* biotin holoenzyme-synthetase/*bio* repressor crystal structure delineates the biotin- and DNA- binding domains. *Proc. Natl. Acad. Sci. USA* **89**, 9257–9261.
30. Hendrickson, W.A. (1985). Analysis of protein structure from diffraction measurement at multiple wavelengths. *Trans. Am. Cryst. Assoc.* **21**, 11–21.
31. Hendrickson, W.A., Smith, J.L., Phizackerley, R.P. & Merritt, E.A. (1988). Crystallographic structure analysis of lamprey hemoglobin from anomalous dispersion of synchrotron radiation. *Proteins* **4**, 77–88.
32. Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51–58.
33. Main, P. (1985). MULTAN — a program for the determination of crystal structures. In *Crystallographic Computing 3: Data Collection, Structure Determination, Proteins and Databases*. (Sheldrick, G.M., Krüger, C. & Goddard, R. eds), pp. 206–215, Clarendon Press, Oxford, UK.
34. Pahler, A., Smith, J.L. & Hendrickson, W.A. (1990). A probability representation for phase determination from multiwavelength anomalous dispersion. *Acta Cryst. A* **46**, 537–540.
35. Wang, B.C. (1985). Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115**, 90–112.
36. Ramachandran, G.N. & Sasisekharan, V. (1968). Conformations of polypeptides and proteins. *Adv. Protein Chem.* **23**, 283–438.
37. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
38. Toh, H., Kondo, H. & Tanabe, T. (1993). Molecular evolution of biotin-dependent carboxylases. *Eur. J. Biochem.* **215**, 687–696.
39. Chou, P.Y. & Fasman, G.D. (1977). Beta-turns in proteins. *J. Mol. Biol.* **115**, 135–175.
40. Baker, E.N. & Hubbard, R.E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **44**, 97–179.
41. Pähler, A. Hendrickson, W.A., Gawinowics Kolks, M.A., Argaraña, C.E. & Cantor, C.R. (1987). Characterization and crystallizations of core streptavidin. *J. Biol. Chem.* **262**, 13933–13937.
42. Brocklehurst, S.M. & Perham, R.N. (1993). Prediction of the three-dimensional structures of the biotinylated domain from yeast pyruvate carboxylase and of the lipoylated H-protein from the pea leaf glycine cleavage system: a new automated method for the prediction of protein tertiary structure. *Protein Sci.* **2**, 626–639.
43. Smith, L.J., *et al.*, & Wlodawer, A. (1994). Comparison of four independently determined structures of human recombinant interleukin-4. *Nat. Struct. Biol.* **1**, 301–309.
44. Wilmot, C.M. & Thornton, J.M. (1990). Beta-turns and their distortions: a proposed nomenclature. *Protein Eng.* **3**, 479–493.
45. Browner, M.F., Taroni, F., Sztul, E. & Rosenberg, L.E. (1989). Sequence analysis, biogenesis, and mitochondrial import of the α-subunit of rat liver propionyl-CoA carboxylase. *J. Biol. Chem.* **264**, 12680–12685.
46. Samols, D., Thornton, C.G., Murtif, V.L., Kumar, G.K., Haase, F.C. & Wood, H.G. (1988). Evolutionary conservation among biotin enzymes. *J. Biol. Chem.* **263**, 6461–6464.

47. Stallings, W. & DeTitta, G.T. (1985). Crystallographic investigations of biotin and carboxybiotin derivatives. In *Biotin*. (Dakshinamurti, K. & Bhagavan, H.N., eds), pp. 152–168, The New York Academy of Sciences, NY.
48. Shrake, A. & Rupley, J.A. (1973). Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* **79**, 351–371.
49. Cohen-Addad, C., Pares, S., Seiker, L., Neuburger, M. & Douce, R (1995). The lipoamide arm in the glycine decarboxylase complex is not freely swinging. *Nat. Struct. Biol.* **2**, 63–68.
50. Goss, N.H. & Wood, H.G. (1984). Formation of N^{ϵ}-(biotinyl)lysine in biotin enzymes. *Methods Enzymol.* **107**, 261–278.
51. Alix, J.-H. (1989). Laboratory methods. A rapid procedure for cloning genes from λ libraries by complementation of *E. coli* defective mutants: application to the *fabE* region of the *E. coli* chromosome. *DNA* **8**, 779–789.
52. Muramatsu, S. & Mizuno, T. (1989). Nucleotide sequence of the *fabE* gene and flanking regions containing a bent DNA sequence of *Escherichia coli*. *Nucleic Acids Res.* **17**, 3982.
53. Studier, F.W., Rosenberg, A.H., Dunn, J.J. & Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
54. Hendrickson, W.A., Horton, J.R. & LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9**, 1665–1672.
55. Wu, H., Lustbader, J.W., Liu, Y., Canfield, R.E. & Hendrickson, W.A. (1994). Structure of human chorionic gonadotropin at 2.6 Å resolution from MAD analysis of the selenomethionyl protein. *Structure* **2**, 545–558.
56. Stura, E.A. & Wilson, I.A. (1990). Analytical and production seeding techniques. *Methods* **1**, 38–49.
57. Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*. (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
58. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.
59. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
60. Brünger, A.T. (1992). *X-PLOR Manual, Version 3.1*. Yale University Press, New Haven, CT.
61. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein refinement. *Acta Cryst. A* **47**, 392–400.
62. Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Cryst.* **11**, 268–272.
63. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.
64. Connolly, M.L. (1983). Solvent accessible surfaces of proteins and nucleic acids. *Science* **221**, 709–713.
65. Hendrickson, W.A. (1979). Transformations to optimize the superposition of similar structures. *Acta Cryst. A* **35**, 158–163.
66. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.
67. Evans, S.V. (1993). SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graphics* **11**, 134–138.
68. Gornicki, P., Scappino, L.A. & Haselkorn, R. (1993). Genes for two subunits of acetyl coenzyme A carboxylase of *Anabaena sp.* strain PCC 7120: biotin carboxylase and biotin carboxyl carrier. *J. Bacteriol.* **175**, 5268–5272.
69. Lopez-Casillas, F., Bai, D.H., Luo, X.C., Kong, I.S., Hermodson, M.A. & Kim, K.H. (1988). Structure of the coding sequence and primary amino acid sequence of acetyl-coenzyme A carboxylase. *Proc. Natl. Acad. Sci. USA* **85**, 5784–5788.

Received: **4 Sep 1995**; revisions requested: **29 Sep 1995**; revisions received: **26 Oct 1995**. Accepted: **27 Oct 1995**.