Purification and characterisation of the BIOH protein from the biotin biosynthetic pathway

Nicholas H. Tomczyk, Joanne E. Nettleship[1], Robert L. Baxter, Hilary J. Crichton,
Scott P. Webster, Dominic J. Campopiano*

Department of Chemistry, Joseph Black Building, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, UK

Abstract Conversion of pimeloyl-coenzyme A (CoA) to biotin in *Escherichia coli* requires at least four enzymes encoded by genes in the *bio* operon. One gene, *bioH*, which is not present in the *bio*ABFCD operon, is required for the synthesis of pimeloyl-CoA but its exact role in formation of this intermediate is unknown. To investigate this further, we have overexpressed and purified the *bio*H gene products from both *E. coli* (BIOH EC) and *Neisseria meningitis* (BIOH NM) in *E. coli*. When purified BIOH was incubated with excess CoA and analysed by electrospray mass spectrometry a species of mass corresponding to a BIOH:CoA complex was observed. Mutation of a conserved serine residue to alanine (BIOH EC S82A) did not prevent CoA binding. This is the first report of the purification of BIOH and the observation of a small molecule bound to the protein provides clues to its role in pimeloyl-CoA synthesis. © 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Pimeloyl-coenzyme A; Biotin; Biosynthesis; Coenzyme A; Mass spectrometry

1. Introduction

Biotin (vitamin H) is a sulfur-containing natural product required by a number of enzymes involved in various metabolic pathways where it acts as a carrier of carbon dioxide. It is produced by bacteria, plants and fungi but not by higher eukaryotes. In *Escherichia coli* the *bio* operon encodes genes required for the conversion of pimeloyl-coenzyme A (CoA), the earliest identified intermediate, into biotin [1,2]. Recent studies have revealed the mechanisms and three-dimensional structures of three of these enzymes; 8-amino-7-oxononanoate synthase (AONS, the *bio*F gene product), 7,8-diaminononanoate synthase (DANS, the *bio*A gene product), and dethiobiotin synthetase (DTBS, the *bio*D gene product) [3–7]. The

final sulfur-insertion step has been the subject of intensive research that has revealed dethiobiotin to biotin conversion requires not only the *bio*B gene product (biotin synthase, BS) but a number of enzymes and small molecules, whose exact roles in the mechanism are still unclear [8–11].

In contrast to these later steps, the biosynthetic pathway leading to pimeloyl-CoA, has not yet been elucidated. Based on the results of labelling studies with ^{13}C-labelled acetates it was suggested that pimeloyl-CoA biosynthesis follows modified fatty acid and polyketide synthesis pathways [12,13]. If this is the case, then it raises the perplexing issue that the biotin biosynthetic pathway uses enzymes that are themselves biotin dependent. Both groups of investigators also speculated that a novel acyl carrier protein (ACP) was involved in the synthesis of pimeloyl-CoA, perhaps as a carrier for 3-oxoglutarate. Early genetic studies of biotin biosynthesis identified two genes, *bio*H and *bio*C that were involved in the steps prior to pimeloyl-CoA synthesis in *E. coli*, but their exact biochemical function has remained unclear [1]. A study of the expression of *E. coli bio*H gene and its effect on biotin production was carried out [14]. Somewhat surprisingly, the authors noted that overexpression of the *bio*H gene led to a decrease of total biotin productivity and it was suggested that this was possibly due to BIOH sequestering a biotin precursor. More recently, an analysis of the BIOH sequence revealed that it shares a highly conserved motif with the thioesterase/acyl transferase domains of fatty acid and polyketide synthases (PKSs) suggesting that BIOH could possibly transfer pimeloyl from the active cysteinyl residue of BIOC directly to CoA [15]. It was also observed that the *bio*X gene from *Bacillus sphaericus* contains a phosphopantetheine (PP) attachment site which would suggest that it is an ACP required for pimeloyl-CoA synthesis in this bacteria.

Genes and enzymes known to be involved in bacterial pimeloyl-CoA biosynthesis are summarised in Fig. 1. Enzymes with pimeloyl-CoA synthetase activity have been cloned from *B. sphaericus* (*bio*W gene product) and *Pseudomonas mendocina* (*pau*A gene product), overexpressed in *E. coli* and purified [16,17]. Interestingly, the BioW and pauA (80 kDa) enzymes have no sequence homology either to each other or to BIOH or BIOC, and show significant differences in size, subunit organisation and enzyme kinetics. It has also been shown that the *B. sphaericus bio*W gene and a *bio*Z gene product from *Mesorhizobium* sp. R7A (which has significant homology with β-ketoacyl-ACP synthase (ACPS) III (FabH)), can each complement *E. coli bio*H deletion mutants [18]. Very recently, two separate studies of BIOI protein from *Bacillus subtilis*

*Corresponding author. Fax: (44)-131-650 4743.
E-mail address: dominic.campopiano@ed.ac.uk (D.J. Campopiano).

[1] On secondment from School of Chemistry, Cantock's Close, University of Bristol, Bristol BS8 1TS, UK.

Abbreviations: CoA, coenzyme A; PP, phosphopantetheine; ACP, acyl carrier protein; ACPS, acyl carrier protein synthase; DTT, dithiothreitol; PCR, polymerase chain reaction; LC-MS, liquid chromatography-mass spectrometry; AONS, 8-amino-7-oxononanoate synthase; DANS, 7,8-diaminononanoate synthase; DTBS, dethiobiotin synthetase; BS, biotin synthase

expressed in *E. coli* have shown it to contain a cytochrome P450 cofactor [19,20] but the nature of its substrates, either free or protein bound, are still unclear.

Somewhat surprisingly, there have been no detailed studies on the structure and function of either the *bio*H or *bio*C gene products. To address this issue, we present here, for the first time, the overexpression in *E. coli* of the *bio*H genes from both *E. coli* (BIOH EC) and *Neisseria meningitis* (BIOH NM). Purified BIOH EC and BIOH NM were obtained in high yield and they have been characterised by a combination of substrate binding and mass spectrometry. This work provides a platform for the determination of the BIOH structure and a more thorough investigation into its function in the biotin biosynthetic pathway.

2. Materials and methods

2.1. Materials

All chemicals were purchased from Aldrich or Sigma and were of the highest purity. Ion exchange and metal-affinity columns were purchased from Amersham Pharmacia Biotech and used on an FPLC. Size exclusion chromatography was performed on a Gilson HPLC instrument. Oligonucleotides were from Sigma Genosys. Plasmid and chromosomal DNA preparations were carried out using commercial kits (Promega, Qiagen). Polymerase chain reactions (PCR) were carried out using Ready-to-Go beads from Amersham Pharmacia Biotech. PCR products were cloned into pCR2.1 using the TOPO T/A cloning kit from Invitrogen. All other manipulations were carried out according to standard methods. The biotin auxotrophic strain *Lactobacillus plantarum* (ATCC 8014) was used in the microbial biotin assay [21].

2.2. Cloning, overexpression and purification of the E. coli bioH gene

The *bio*H gene was amplified by PCR from *E. coli* JM101 chromosomal DNA template using a forward primer of 5′-CGC GCG <u>GAA TTC</u> **ATG** A*AC AAC ATC TGG TGG-3′ which contains *Eco*RI (underlined) and *Bsp*HI (bold) restriction sites. The reverse primer was 5′-CGC GCG **CGG ATC C**TA CAC CCT CTG CTT CAA-3′ which contains a *Bam*HI site (bold). The PCR-amplified bioH gene was digested with *Bsp*HI and *Bam*HI, then cloned into pET-16b (Novagen, Madison, WI, USA) cut with *Nco*I and *Bam*HI to yield pBIOH.

HMS174(DE3) cells were transformed with pBIOH and were used to inoculate 500 ml 2×YT and grown overnight at 37°C. These were then decanted into 5 l 2×YT broth containing Amp (100 μg/ml) and then divided into 10×500 ml batches in 1 l shake flasks and grown to an OD_{600} of 1. Isopropyl-β-ᴅ-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce BIOH production and cells were grown for a further 3 h at 37°C. The cells were isolated by centrifugation, washed with 50 mM Tris, pH 7.5 (buffer A) and stored at −20°C. Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) gel analysis revealed a large band at 28 kDa as expected for BIOH (predicted MW 28 505 Da)

The cell paste was resuspended in 4× wet weight of buffer A and sonicated (15 min, 30 s on/off) on ice. The cell extract was centrifuged (2×20 min, 10 000 rpm) and the cell-free extract was then applied to a FPLC fitted with a 26/10 Q-Sepharose column (Pharmacia). BIOH EC was eluted using a linear 0–1 M NaCl gradient (buffer B, 50 mM Tris, 1 M NaCl, pH 7.5) over 20 column volumes (1 l). Analysis of the fractions by SDS–PAGE showed that BIOH EC eluted between 110 and 160 mM NaCl. Fractions containing BIOH EC were pooled and frozen at −20°C. The overall yield of BIOH EC was routinely 25–40 mg from 5 l of cell culture.

2.3. Size exclusion of purified BIOH EC

Purified BIOH EC was subjected to size exclusion chromatography on a previously calibrated Sepharose 200 HR column. A sample of protein (5 mg/ml) was loaded onto the column and eluted with the following buffer 50 mM Tris, 200 mM NaCl, 7 mM β-mercapto-ethanol, pH 7.5. BIOH (28.5 kDa) eluted with a retention time between chymotrypsinogen (25 kDa) and carbonic anhydrase (29 kDa), suggesting that it is monomeric in solution.

2.4. Site-directed mutagenesis of BIOH EC

The megaprimer method was used to generate single mutants of BIOH EC [22]. Mutagenic forward primers: for BIOH S53A (5′-TTC GGG CGT <u>GCT</u> CGG GGA GGA TTT-3′) and BIOH S82A (5′-TGG TTA GGC TGG <u>GCT</u> CTG GGC GGG CTG-3′). The mutagenic residues are underlined. The external forward and reverse primers were as follows: forward (5′-GTT GTG TGG AAT TGT GAG CGG-3′) and reverse (5′-CGC CAG GGT TTT CCC AGT CAC G-3′). Plasmid pUC18/*bio*H was used as a template for the PCR reactions. The BIOH EC S53A and S82A genes were cloned and overexpressed in an identical manner to wild-type BIOH EC.

2.5. Cloning, overexpression and characterisation of bioH from N. meningitis

Analysis of the complete genome of *N. meningitis* revealed an open reading frame (ORF) with high amino acid sequence identity to *E. coli* BIOH. Oligonucleotide primers were designed to allow PCR amplification of this gene from the *N. meningitis* strain H44/76 (a B-capsular type). These were as follows: BIOH NM forward (5′-**CCA TGG** G*TC CGG ATG CGG TTA AAA AAG-3′) which contains a *Nco*I site (bold) and BIOH NM reverse (5′-**GGA TCC** TCA ACG CAG ACC ACC TTC AAC AAA GTC-3′) which contains a *Bam*HI site (bold). The mass of wild-type BIOH NM is 26 978 Da, but we made a modification to the coding sequence to facilitate cloning into the pET expression vector, inserting a Gly residue at the N-terminus of BIOH NM (underlined in BIOH NM forward) to give a protein with a predicted mass of 27 035 Da. The PCR-amplified gene was cloned into plasmid pCR2.1 and sequenced to confirm no errors had occurred during the PCR process. The *bio*H NM gene was digested with *Nco*I and *Bam*HI and cloned in pET-16b digested with the same enzymes to give pBIOH NM. This was used to transform HMS174 (DE3) cells and BIOH NM was overexpressed and purified in an identical procedure to that used on BIOH EC using ion-exchange and gel-filtration chromatography to yield ∼30 mg BIOH NM from 5 l broth.

2.6. Mass spectrometry

Mass spectrometry was performed on a Micromass Platform II quadrupole mass spectrometer equipped with an electrospray ion source. The spectrometer cone voltage was ramped from 40 to 70 V and the source temperature set to 140°C. Protein samples were separated with a Waters HPLC 2690 with a Phenomenex C5 reverse phase column directly connected to the spectrometer. The proteins were eluted from the column with a 5–95% acetonitrile (containing 0.01% trifluoroacetic acid) gradient at a flow rate of 0.4 ml/min. The total ion count in the range 500–2000 *m/z* was scanned at 0.1 s intervals. The scans were accumulated and spectra combined and the molecular mass was determined by the MaxEnt and Transform algorithms of the Mass Lynx software (MicroMass).

2.7. Binding of CoA to BIOH

To test BIOH proteins as substrates for CoA binding in vitro the following assays were performed based on the procedure used by Flugel et al. [23]. A typical reaction mixture contained BIOH EC, BIOH NM or BIOH S82A (10 μM) in assay buffer (50 mM Tris, pH 7.5, dithiothreitol (DTT; 1 mM), $MgCl_2$ (10 mM)) in a final volume of 10 ml. Where required CoA (80 μM) was added, and the reaction incubated for 2 h at 37°C. The reaction was quenched by immediate freezing at −20°C. Each reaction was defrosted and diluted two-fold by the addition of 10 ml H_2O. The total sample (20 ml) was applied to a Resource Q ion exchange column. The column was washed with 50 mM Tris, pH 7.5 and the protein was eluted with a linear NaCl (0–1 M) gradient over 20 CVs. The BIOH-containing fractions were identified by SDS–PAGE, concentrated by ultracentrifugation and analysed by liquid chromatography-mass spectrometry (LC-MS).

3. Results and discussion

3.1. Cloning, overexpression and purification of the E. coli and N. meningitis bioH gene products (BIOH EC and BIOH NM)

Primers were designed to allow PCR amplification of the *E. coli bio*H gene from chromosomal DNA template and cloning
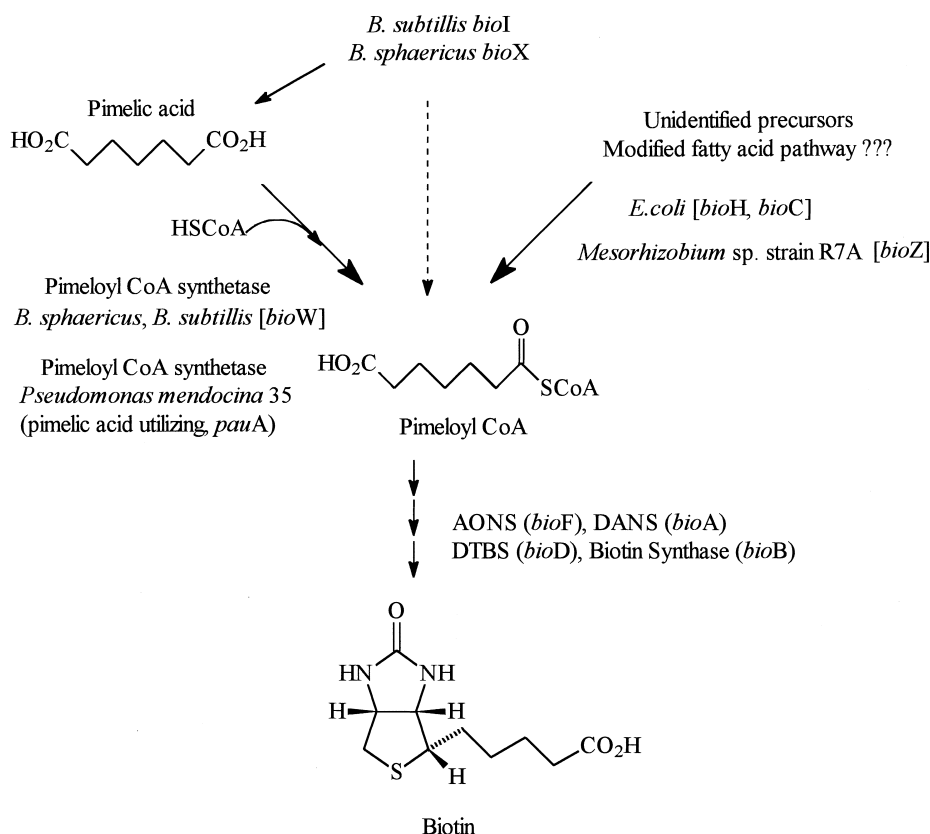
Fig. 1. Routes of pimelate and pimeloyl-CoA synthesis in various Gram-positive and Gram-negative bacteria. Pimelate synthesis has been shown to require the products of the *bio*I gene in *B. subtilis* and the *bio*X gene product in *B. sphaericus*. Pimeloyl-CoA synthesis requires the *bio*C and *bio*H gene products in *E. coli*, and the *bio*Z gene in *Mesorhizobium*. Pimeloyl-CoA synthetase activity has been demonstrated by BioW from *B. sphaericus* and *B. subtilis* and the PauA enzyme from *P. mendocina*. All bacteria studied to date appear to use the same enzymes (AONS, DANS, DTBS, BS) to convert pimeloyl-CoA to biotin.

into a suitable expression vector. During our characterisation of BIOH EC the complete genome sequence of *N. meningitis* became available and a gene encoding an ORF (249 amino acids) was identified which had amino acid sequence homology (26% identity, 43% homology) to BIOH EC. Amplification of the *bio*H NM gene from *N. meningitis* genomic DNA template was achieved by PCR. Overexpression of both *bio*H genes using the T7 RNA polymerase from the pET system produced large amounts of soluble BIOH EC and NM (SDS–PAGE, ∼28 kDa, see Fig. 2A,B). Both BIOH EC and BIOH NM were readily purified in a single step using anion exchange chromatography to yield ∼10 mg purified protein/l bacterial extract. Gel permeation chromatography showed that both BIOH proteins behaved as monomers in solution. We also tested for possible BIOH phosphopantetheinylation using the ^{14}C-radiolabelled PP precursor, β-alanine, but BIOH was not labelled (data not shown).

3.2. Site-directed mutagenesis of BIOH EC

Sequence analysis of BIOH EC (see Fig. 4) has revealed two motifs (Gly X Ser X Gly) which are similar to those present in a number of thioesterase/acyl transferase domains involved in fatty acid and polyketide biosynthetic enzymes [15]. A serine-containing motif has been identified as the site of phosphopantetheinylation in the ACP domains of various fatty acid synthases, PKSs and non-ribosomal peptide synthetases [24,25]. Both serine residues within these motifs were con-

verted to alanine to give the BIOH EC mutants S53A and S82A. These mutants were overexpressed using the same conditions as the wild-type BIOH EC gene. The BIOH EC S53A mutant was found to be insoluble and we were unable to study the properties of this protein. The BIOH EC S82A mutant was soluble and behaved essentially as wild type protein during purification.

3.3. Mass spectrometry of purified BIOH EC, BIOH NM and BIOH EC S82A

In order to determine the accurate mass of BIOH EC and to investigate any possible post-translational modifications we analysed the purified protein by LC-MS. The transformed ion

Table 1
LC-MS results of in vitro BIOH:CoA binding assay

Protein	CoA	Observed masses (Da)
BIOH EC	−	28 502.0
BIOH EC	+	28 506.6; 28 849.8, 29 272.0
BIOH NM	−	27 035.0
BIOH NM	+	27 032.0; 27 802.0
BIOH EC S82A	−	28 487.0
BIOH EC S82A	+	28 494.0; 28 840; 29 257

The data was generated using using MaxEnt© software (Micromass) as described in Section 2. Predicted masses are as follows: BIOH EC (28 505 Da), BIOH EC S82A (28 489 Da), BIOH NM (27 035 Da).
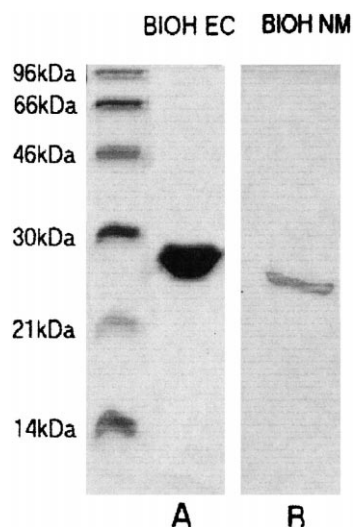
Fig. 2. SDS–PAGE (12%) gels of (A) purified BIOH EC and (B) purified BIOH NM used for LC-MS analysis and in vitro assays.

series gave rise to masses of $28\,502.0 \pm 5.5$ Da which was in agreement with the predicted M_r of 28 505 (see Table 1). The BIOH EC S82A mutant gave a mass of $28\,487.0 \pm 4.9$ Da (predicted 28 489). The recombinant *N. meningitis* BIOH NM gave a mass of $27\,035.0 \pm 7.4$ Da (predicted 27 035).

3.4. Observation of a BIOH:CoA complex

Based on a sequence analysis of BIOH EC, Lemoine et al. [15] suggested that a possible role for the *bio*H gene product could be to transfer pimeloyl from the active cysteinyl residue of BIOC directly to CoA. To test this hypothesis we incubated

purified BIOH EC, BIOH EC (S82A) and BIOH NM each at 10 μM with an excess of CoA (80 μM) and the BIOH-containing fractions were purified by anion exchange before being analysed by LC-MS. The results of these assays are shown in Table 1.

Appropriate controls were also carried out containing all the assay components in single experiments, i.e. without CoA, MgCl$_2$ and DTT. When BIOH EC had been incubated with CoA, MgCl$_2$ and DTT, the transformed spectra revealed the presence of unmodified BIOH EC at $28\,506.6 \pm 6.52$ Da the appearance of two masses at $28\,849.8 \pm 6.52$ and $29\,272.0 \pm 9.92$ Da (see Fig. 3). The addition of 765.4 Da corresponds with the mass of a CoA molecule (M_r 767 Da) but we are, as yet, unsure of the nature of the 343.2 Da adduct. Each BIOH EC in vitro assay was purified using anion exchange, then reverse-phase chromatography before mass spectrometry analysis and this would suggest that the CoA is relatively tightly bound. Evidence for CoA binding to BIOH EC from our initial experiments was further complemented by studies on BIOH NM and the BIOH EC S82A mutant. When BIOH NM was incubated with CoA, species with masses of $27\,032.0 \pm 9.56$ and $27\,802.0 \pm 8.02$ Da were observed corresponding to BIOH NM with a mass of 770 Da which is within experimental error of the mass of CoA. The same CoA incubation and isolation was carried out on BIOH EC S82A and masses were observed at $28\,494.0 \pm 6.52$, $28\,840.0 \pm 5.40$ and $29\,257 \pm 7.07$ Da. The smaller increase is +343 and the larger species corresponds to BIOH S82A with a mass of 763 Da, which again matches that of CoA.

We observed the BIOH:CoA complex after incubation in a 1:8 ratio but the LC-MS showed free BIOH and BIOH:CoA. We are currently attempting to produce fully loaded BIOH:CoA.
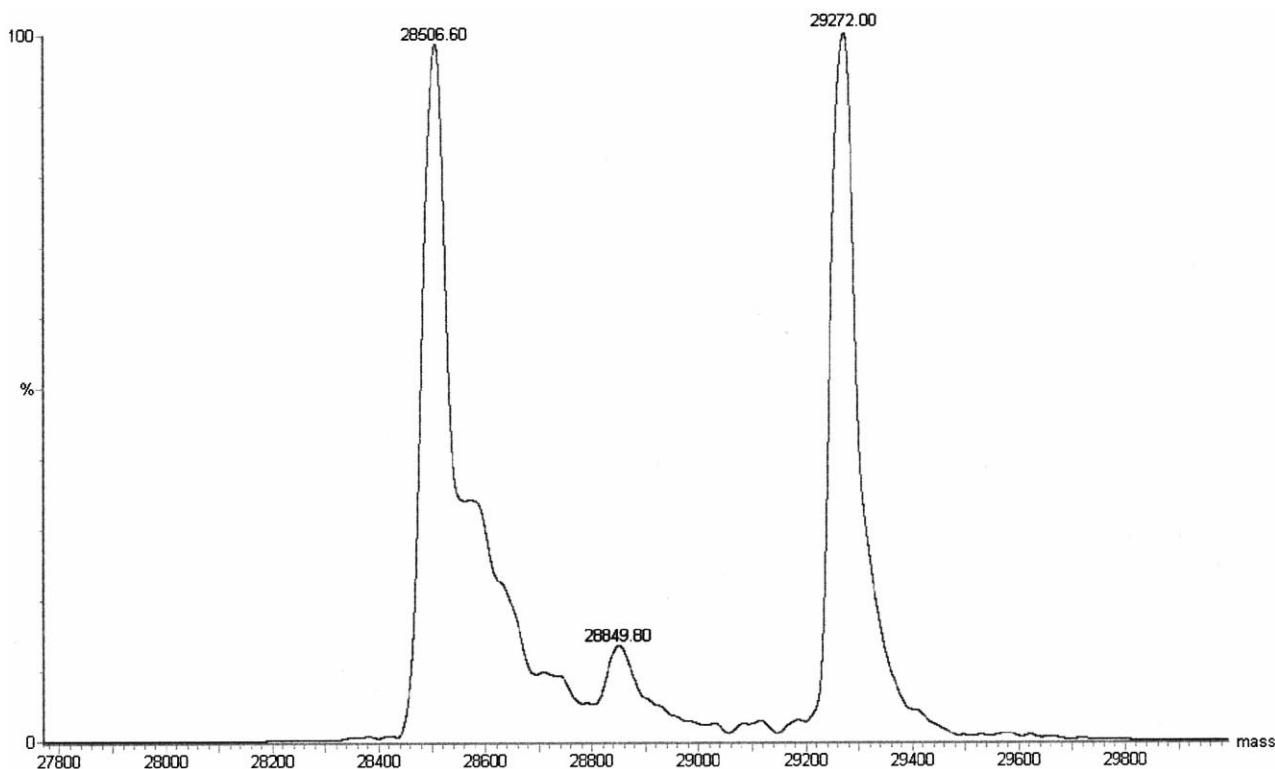

Fig. 3. The transformed LC-MS data of the BIOH EC incubated with CoA in vitro as described in Section 2.

```
             1                                                        #                                    *100
BIOH_EC   (1) --------------MNNIWWQTKGQGNVHLVLLHGWGLNAEVWRCIDEELS-SHFTLHLVDLPGFGRSRGFGALSLADMAEAVLQQAPDKAIWLGWSLGG
BIOH_KS   (1) ----------------MKQPNLVMLPGWGMEKDAFQPLIKPLS-EVFHLSFIEWRDMKTLN-----DFEERVIDTIASIDGPVFLLGWSLGS
BIOH_NM   (1) ----------------MPDAVKKVYLIHGWGANRHMFDDLMPRLP-ATWPVSAVDLPGHGDAPFVRPFDIAAAADGIAAQIDAPADILGWSLGG
BIOH_PA   (1) ------------------MRDHLILLPGWGLGSAPLEPLRDALHEREPHLNVLIEPLPSLD--D----AADWLDELDDNLPRDSWLAGWSLGG
BIOH_VC   (1) MAEWMKMEIDGETVTMALYWQVSGQG-QDLVLVHGWGMNGAVWQQTAQALS-DHFRVHVVDLPGYGHSAEQHAASLEEIAQALLEHAPRNAIWVGWSLGG
BIOH_XF   (1) ----------------MYIEVTGYG-PALVLIHGWAMHSGVFAPLVEQLR-AHHTLYLVDLPGHGYNHTTLTPLALPHVVHAIAAATPPAVWLGWSLGG
Consensus (1)                                 GW              L                                             GWSLG

             101                                                                                          200
BIOH_EC  (86) LVASQIALTHPERVQALVTVASSPCFSARDEWP-GIKPDVLAGFQQQLSDDFQRTVERFLALQTMGTE-TARQDARALKKTVLALPMPEVDVLNGGLEIL
BIOH_KS  (71) LLSLELVSSYREKIKGFILIGATSRFTTGDNYSFGWDPRMVERMKKQLQRNKEKTLTSFYEAMFSEAEKEEGFYHQFITTIQSEFHGDDVFSLLIGLDYL
BIOH_NM  (78) LVALYLAARHPDKVRSLCLTASFARLTADEDYPEGLAAPALGKMVGAFRSDYAKHIKQFLQLQLLHTP-DADGIIGRILPDLARCGTP--QALQEALDAA
BIOH_PA  (70) MLAGELAARRGDDCRGLLTLASNPCFRVREDWPNAMPAETFEDFFEAFLLEPHLTRKRFTLLVSQG----ARDPRTLARQLQVALPQLEREALVAGLQLL
BIOH_VC  (99) LVATHMALHHSDYVSKLVTVASSPKFAAQGSWR-GIQPDVLTAFTDQLVADFQLTIERFMALQAMGSP-SARQDVKVLKQAVLSRPMPNPQSLLAGLTML
BIOH_XF  (82) LFALHAAATLPQ-VRGLIMLAATPCFVRREDWPHAVEVSIFTQFAEDLKQNYTETINRFLALDTLGST-YAQSELRQLRQILNARHTPNTATLQAGLELL
Consensus (101)                                                         F                                     L  L

             201
BIOH_EC (184) KTVDLRQPLQNVSMPFLRLYGYLDGLVPRKVVPMLDKLWP--HSESYIFAKAAHAPFISHPAEFCHLLVALKQRV-----256
BIOH_KS (171) LQKDVRVKLDQIETPILLIHGREDKICPLEASSFIKENLG-GKAEVHIIEGAGHIPFFTKPQECVQLIKTFIQKEYIHD-248
BIOH_NM (175) ERADARHLLDKIDVPVLLVFGGKDAITPPRMGEYLHRRLK--GSRLVVMEKAAHAPFLSHAEAFAALYRDFVEGGLR---249
BIOH_PA (166) GQLDTRAALENFRGPQLHLFAEADALVPLAAAEALLEWLP---DVEVSTLAASHGLPLECPDEVAGAILRFLREGDDA--240
BIOH_VC (197) AEVDLRDELQHISVPMLRLYGRLDGLVPAKVARDLNHLAP--YSEAFMFDQSSHAPFMTEAEAFCQQLIEFATR------268
BIOH_XF (180) AHTDLRRAVIDLTPPSLWIAGQRDRLVPAASIHAATALAPSGQTELLTITGGGHAPFLSHANQMTAALQHFIATLP----255
Consensus (201)    D R        P L        D   P                   H
```

Fig. 4. Sequence alignment of BIOH proteins in the SwissProt and GenBank databases, generated using CLUSTAL W. The serine residues mutated in this study are marked Ser53 (#) and Ser82 (*). The accession numbers of the BIOH sequences used are BIOH EC (*E. coli*, K12, AAC76437), BIOH NM (*N. meningitis* MC58, AAF40724), BIOH KS (*Kurthia* sp. 538-KA26, BioH-II, BAB39462), BIOH VC (*V. cholerae*, AAF95858), BIOH PA (*P. aeruginosa*, AAG03891) and BIOH XF (*X. fastidiosa* 9a5c, AAF84165).

These results provide evidence that BIOH EC and BIOH NM can bind CoA in vitro and that the S82A mutation present in the conserved Gly X Ser X Gly sequence motif in BIOH EC did not effect this binding. DTT was added to the assay mixture to ensure CoA was in its reduced state and the fact that BIOH EC was not modified by DTNB suggests (data not shown) that there are no free thiols on the surface of BIOH which could be modified by the thiol of reduced CoA.

We have yet to measure the K_d for CoA (the determination of the CoA binding constant of the various BIOH proteins requires the use of radiolabelled CoA and equilibrium dialysis which we did not have access to for these initial studies). The intracellular *E. coli* CoA concentration is 50 μM (grown in glucose medium), and can be increased to >500 μM by the addition of 25 mM pantothenate [23]. The fact that BIOH is isolated in the CoA-free state suggests that the CoA binding constant may be >100 μM and it is fortunate that the BIO-H:CoA complex prepared in vitro is stable to electrospray ionisation. For comparison, *E. coli* acyl CoA synthase is purified in an unbound form and has a K_m (CoA) of 50 μM in a steady-state apo-ACP phosphopantetheinyl transferase assay.

3.5. Amino-acid sequence analysis of BIOH

3.5.1. BIOH homologs. Previous sequence analysis of the BIOH EC had highlighted certain motifs in the primary sequence that suggested that the protein belonged to the thioesterase family of proteins [15]. This Gly X Ser X Gly sequence motif was found in many thioesterase, lipase and serine protease domains [26] and is present in two regions within BIOH EC. Until recently, the only BIOH sequence available from the GenBank and SwissProt databases was the *E. coli* protein. During the course of this work another five sequences appeared in the databases with high sequence similarity to *E. coli* BIOH. These putative BIOH proteins are present in the genomes of pathogenic bacteria such as *N. meningitis*, *Pseudomonas aeruginosa*, *Vibrio cholera* and *Xylella fastidiosa* and the aerobic bacterium *Kurthia*. We carried

out a sequence alignment using CLUSTAL W to identify conserved residues (shown in Fig. 4) and it shows that 18 residues are absolutely conserved over the length of the six sequences. It is interesting to note that Ser53 is not conserved but the BIOH S53A mutant is insoluble and that Ser82 of BIOH EC is one of the residues within the conserved block of Gly Trp Ser Leu Gly. We observed the BIOH EC S82A:CoA complex by LC-MS but a measurement of the K_d of this mutant will reveal if it has an effect on CoA binding. This mutant will also be used in future work to determine the biochemical function of the BIOH:CoA complex.

3.5.2. Similarity with thioesterases. The highly conserved Gly X Ser X Gly sequence motif of BIOH highlighted in Fig. 4 has recently been found to play a crucial role in various thioesterase enzymes [15]. Although thioesterases catalyse a number of physiologically important reactions, only five crystal structures have been reported to date from bacteria and eukaryotic organisms. These include the *E. coli* medium chain acyl-CoA thioesterase II [27] and the human acyl protein thioesterase I [28]. The eukaryotic thioesterases belong to the α/β hydrolase superfamily whereas the prokaryotic thioesterases appear to be distant homologs. A significant fact is that although these enzymes use either a protein-bound or a free small molecule thioester as a substrate they each have a catalytic site made up a triad of Ser, His and Asp with the Ser residue proposed to act as the nucleophile for attack on the carbonyl of the thioester bond. The possibility that BIOH could belong to the thioesterase family is further strengthened by the observation that the sequence alignment reveals three conserved residues which could act as a potential catalytic triad at Ser82, His235 and Asp187 or Asp207 (see Fig. 4). It may well be that BIOH acts upon a pimeloyl-ACP substrate and current efforts are aimed at detecting such a protein-bound substrate.

3.6. Biotin production assay

A previous report had found that overexpression of BIOH

in *E. coli* had resulted in a decrease in the production of (total) biotin [14]. The authors suggested that the overexpression of BIOH reduces the level of some biotin precursor, possibly synthesised through a modified fatty acid pathway. Increased levels of BIOH may have disrupted the normal metabolic rate, resulting in a reduction of the specific intermediate(s) for pimeloyl-CoA synthesis. We tested both pBIOH EC/HMS174 (DE3) and HMS174 (DE3) cells and found both produced the same amount of biotin (data not shown).

3.7. Possible roles of BIOH in biotin biosynthesis

Genetic complementation experiments coupled with enzymatic assays have provided clues to the functions of various proteins involved in pimeloyl-CoA synthesis in different bacteria outlined in Fig. 1. Deletion strains of *E. coli bioH* can be complemented by genes encoding cytochrome P-450 proteins (*bio*I), pimeloyl-CoA synthetase (*bio*W) and β-ketoacyl-ACPS (*bio*Z) in the presence of exogenous pimelic acid [16,18,29]. These results, combined with our observation of a BIO-H:CoA complex by LC-MS, lead us to propose a possible role for BIOH as the CoA donor to a pimeloyl-ACP (or condensing enzyme), releasing pimeloyl-CoA. We have also overexpressed *E. coli* BIOC but found it to be insoluble, so our current studies are focused on producing soluble BIOC and determining both the extent of any post-translational modification and its possible interaction with BIOH (Tomczyk and Campopiano, unpublished results).

Since pimeloyl-CoA synthesis appears to require protein-bound components it also explains why a simple pimeloyl-CoA synthetase activity (i.e. an enzyme which catalyses pimeloyl-CoA formation from pimelate, CoA and MgATP) has not been detected in *E. coli* extracts. It is interesting to note that a recent study of the function of LipA from the lipoate biosynthetic pathway has shown that an octanoyl-ACP is the immediate precursor to lipoate and that free octanoate is not [30]. Purified BIOH and labelled CoA will be valuable tools to study pimeloyl-CoA biosynthesis in vitro. The determination of the structure of BIOH alone and in complex with CoA will also increase our understanding of its function in the biotin biosynthetic pathway.

References

[1] Eisenberg, M.A. (1987) in: *Escherichia coli* and *Salmonella typh-imurium* - Cellular and Molecular Biology (Neidhardt, F., Ed.), pp. 544–550, American Society for Microbiology, New York.

[2] Otsuka, A.J., Buonocristiani, M.R., Howard, P.K., Flamm, J., Johnson, C., Yamamoto, R., Uchida, K., Cook, C., Ruppert, J. and Matsuzaki, J. (1988) J. Biol. Chem. 263, 19577–19585.

[3] Schneider, G. and Lindqvist, Y. (2001) FEBS Lett. 495, 7–11.

[4] Webster, S.P., Alexeev, D., Campopiano, D.J., Watt, R.M., Alexeeva, M., Sawyer, L. and Baxter, R.L. (2000) Biochemistry 39, 516–528.

[5] Kack, H., Sandmark, J., Gibson, K., Schneider, G. and Lindqvist, Y. (1999) J. Mol. Biol. 291, 857–876.

[6] Gibson, K.J., Lorimer, G.H., Rendina, A.R., Taylor, W.S., Cohen, G., Gatenby, A.A., Payne, W.G., Roe, D.C., Lockett, B.A., Nudelman, A., Marcovici, D., Nachun, A., Wexler, B.A., Marsili, E.L., Turner, I.M., Howe, L.D., Kalbach, C.E. and Chi, H. (1995) Biochemistry 34, 10976–10984.

[7] Alexeev, D., Baxter, R.L. and Sawyer, L. (1994) Structure 2, 1061–1072.

[8] Ifuku, O., Kishimoto, J., Haze, S., Yanagi, M. and Fukushima, S. (1992) Biosci. Biotech. Biochem. 56, 1780–1785.

[9] Birch, O.M., Hewitson, K.S., Fuhrmann, M., Burgdorf, K., Baldwin, J.E., Roach, P.L. and Shaw, N.M. (2000) J. Biol. Chem. 275, 32277–32280.

[10] Bui, B.T.S., Florentin, D., Fournier, F., Ploux, O., Mejean, A. and Marquet, A. (1998) FEBS Lett. 440, 226–230.

[11] McIver, L., Baxter, R.L. and Campopiano, D.J. (2000) J. Biol. Chem. 275, 13888–13894.

[12] Ifuku, O., Miyaoka, H., Koga, N., Jiro, S., Haze, S., Wachi, Y. and Kajiwa, M. (1994) Eur. J. Biochem. 220, 585–591.

[13] Sanyal, I., Lee, S. and Flint, D. (1994) J. Am. Chem. Soc. 116, 2637–2638.

[14] Koga, N., Kishimoto, J., Haze, S. and Ifuku, O. (1996) J. Ferm. Bioeng. 81, 482–487.

[15] Lemoine, Y., Wach, A. and Jeltsch, J. (1996) Mol. Microbiol. 19, 639–647.

[16] Ploux, O., Soularue, P., Marquet, A., Gloeckler, R. and Lemoine, Y. (1992) Biochem. J. 287, 685–690.

[17] Binieda, A., Fuhrmann, M., Lehner, B., Rey-Berthod, C., Frutiger-Hughes, S., Hughes, G. and Shaw, N. (1999) Biochem. J. 340, 793–801.

[18] Sullivan, J.T., Brown, S.D., Yocum, R.R. and Ronson, C.W. (2001) Microbiology 147, 1315–1322.

[19] Green, A.J., Rivers, S.L., Cheesman, M., Reid, G.A., Quaroni, L.G., Macdonald, I.D.G., Chapman, S.K. and Munro, A.W. (2001) J. Biol. Inorg. Chem. 6, 523–533.

[20] Stok, J.E. and De Voss, J.J. (2000) Arch. Biochem. Biophys. 384, 351–360.

[21] DeMoll, E. and Shive, W. (1986) Anal. Biochem. 158, 55–58.

[22] Sarker, G. and Sommers, S.S. (1990) BioTechniques 8, 404–407.

[23] Flugel, R.S., Hwangbo, Y., Lambalot, R.H., Cronan, J.E. and Walsh, C.T. (2000) J. Biol. Chem. 275, 959–968.

[24] Marahiel, M.A., Stachelhaus, T. and Mootz, H.D. (1997) Chem. Rev. 97, 2651–2673.

[25] Cane, D.F. and Walsh, C.T. (1999) Chem. Biol. 6, R319–R325.

[26] Brenner, S. (1988) Nature 227, 680–685.

[27] Li, J., Derewenda, U., Dauter, Z., Smith, S. and Derewenda, Z.S. (2000) Nat. Struct. Biol. 7, 555–559.

[28] Devedjiev, Y., Dauter, Z., Kuznetsov, S.R., Jones, T.L.Z. and Derewenda, Z. (2000) Structure 8, 1137–1146.

[29] Bower, S., Pero, J., Perkins, J.B., Rogers Yocum, R., Howitt, C.L. and Rahaim, R. (1996) J. Bacteriol. 178, 4122–4130.

[30] Miller, J.R., Busby, R.W., Jordan, S.W., Cheek, J., Henshaw, T.F., Ashley, G.W., Broderick, J.B., Cronan Jr., J.E. and Marletta, M.A. (2000) Biochemistry 39, 15166–15178.