

Supplemental Information to:

Biotin Synthesis Begins by Hijacking the Fatty Acid Synthetic Pathway

Steven Lin, Ryan E. Hanson and John E. Cronan

SUPPLEMENTAL METHODS

Construction of plasmids. Bacterial strains, plasmids and primers are listed in Table S2. The *bioC* and *bioH* genes were PCR amplified from MG1655 genomic DNA using oligonucleotide sets A07-A08 and A11-A12, respectively. The *bioC* PCR product was digested with NcoI and XhoI whereas the *bioH* PCR products were digested with BspHI and XhoI. The DNA fragments were ligated into pET28b+ digested with NcoI and XhoI to give plasmids pSTL4 and pSTL6 carrying the *bioC* and *bioH* genes, respectively. Plasmid pSTL13 encodes a mutant version of *bioH* in which codon 82 was changed from serine to alanine by site-directed mutagenesis using oligonucleotides A39 and A40 and the QuikChange procedure (Stratagene). Plasmid pSTL20 carrying the *Bacillus subtilis bioW* gene was constructed by amplification of *bioW* from *B. subtilis* genomic DNA using oligonucleotides A41 and A55. The resulting PCR product was digested with NcoI and XhoI and ligated into pBAD322 digested with NcoI and Sall. Plasmid pCY123 which carries the entire *bio* operon (*bioABFCD*) was constructed by digestion of plasmid pLC25-23¹ with HindIII and EcoRI and ligation of the 6.6 kb *bio* fragment to pBR329 digested with the same enzymes. An in-frame $\Delta bioC$ derivative of plasmid pCY123 was constructed by digestion with BglIII and religation of the plasmid, removing 46 amino acid residues between codon 24 and 70. The resulting plasmid pSTL25 complemented a $\Delta bioD$ strain but not a $\Delta bioC$ strain. The HindIII site lies in the phage λatt region (the original clone was made from a λ lysogen). Plasmid DNAs were extracted by QIAprep Miniprep (Qiagen). The constructs were verified by DNA sequencing conducted by the Core sequencing Facility of the Carver Biotechnology Center of University of Illinois at Urbana-Champaign.

Construction of bacterial strains. The strains used in this study were derivatives of the sequenced *E. coli* K-12 strain MG1655 unless stated otherwise. Chromosomal deletions of *bio* genes in strain BW25113 were constructed by the λ recombination system of Datsenko et al.². The *kan* cassette was amplified from pKD13 using the following oligonucleotides A70-A71 (*bioA* deletion), A15-A16 (*bioC* deletion), A68-A69 (*bioF* deletion) and A17-A18 (*bioH* deletion). Essentially the entire center of the coding sequences of each the genes was deleted in frame leaving <20 residues at the termini. This was essential for function of the genes upstream and downstream of *bioF* and *bioC* because the coding sequences overlap. The *kan* cassette plus the flanking chromosomal segments was transduced into MG1655 using phage P1*vir* and the cassette subsequently eliminated by the FLP helper plasmid². Strains STL23, STL24, STL112 and STL115 carrying in-frame deletions of *bioC*, *bioH*, *bioA* and *bioF*, respectively, were verified as biotin auxotrophs by testing on glucose minimum media plus or minus biotin and the lack of any polarity as demonstrated by complementation with plasmids carrying the gene deleted from the chromosome. Strain STL25 carrying deletions of both the *bioC* and *bioH* genes was constructed by transduction of strain STL23 to kanamycin resistance with a P1 phage stock grown on strain STL29. The *kan* cassette in *bioH* was eliminated by the FLP helper plasmid². The bypass strains STL32, STL33 and STL34 were constructed by transforming strains STL23, STL24 and STL25, respectively, with pYFJ84 which encodes *aasS* under control of a T7 promoter. Strain STL74 was made by transforming STL25 with plasmid pSTL20, which carried the *Bacillus subtilis bioW* gene under control of an *araBAD* promoter.

Esterification of dicarboxylic acid and analysis. The dicarboxylic acids and most of their esters were obtained commercially. Those monoesters that were not commercially available were synthesized by one of three methods. Diesters were partially hydrolyzed under basic conditions that gave hydrolysis and transesterification whereas dicarboxylic acid monoesters were synthesized from the dicarboxylic acid by ion exchange resin-catalyzed esterification in a two-phase solution that gives selective monesterification³ or by a more robust procedure in which acid-catalyzed esterification/transesterification was performed by dissolving the acid to 0.1 M in the anhydrous alcohol plus 0.1 % (v/v) concentrated HCl⁴. These reactions gave mixtures of diester and monoester together with the free dicarboxylic acid with at least half of the products being the monoester as determined by gas chromatography mass spectrometry

(GC/MS) at the Roy J. Carver Metabolomics Center at University of Illinois at Urbana-Champaign. Samples of 1 μ l were injected with a split ratio of 150:1 on the Agilent GC/MS system consisted of a 5973 MSD, a 7683B autosampler and a 6890N gas chromatograph equipped with a 30 m ZB-WAX capillary column, 0.25 mm I.D. and 250 μ m film thickness (Phenomenex). Injection temperature and the MSD transfer line were set to 280 $^{\circ}$ C, the ion source and MS quadrupole were adjusted to 230 $^{\circ}$ C and 150 $^{\circ}$ C respectively. The helium carrier gas was set at a constant flow rate of 2 ml min⁻¹. The temperature program was: 140 $^{\circ}$ C for 5 min, 40 $^{\circ}$ C min⁻¹ to 260 $^{\circ}$ C, 265 $^{\circ}$ C for 10 min. Acquired GC/MS spectra were recorded in the m/z 50-800 scanning range and processed using AMDIS (NIST) and MSD ChemStation D.02.00.275 (Agilent) software and compared with standard mass spectrum libraries NIST08 (NIST), WILEY8n (Palisade), and a custom library. The concentrations of the monoester compounds we synthesized were determined by adding a known concentration of an internal standard monoester having a chain length that differed from the monoester of interest by one methylene group.

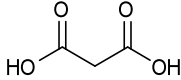
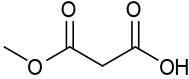
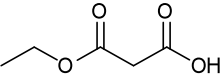
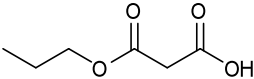
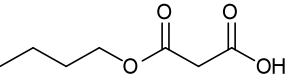
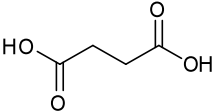
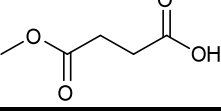
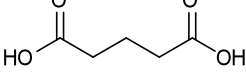
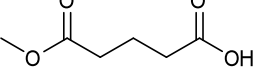
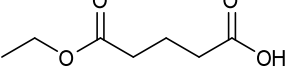
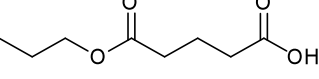
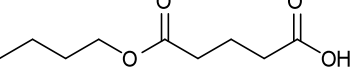
BioH Purification. The C-terminal His-tagged versions of BioH and BioH S82A were over-expressed from strains STL14 and STL47, respectively. The proteins were purified by immobilized-metal affinity chromatography using a nickel-nitrilotriacetic acid column (Ni-NTA) from Qiagen according to the manufacturer's protocol. The cells were resuspended in Loading Buffer (50 mM 3-(N-morpholino)propanesulfonic acid, pH 7.5, 500 mM NaCl, 10 % glycerol and 5 mM 2-mercaptoethanol) containing 10 mM imidazole and lysed by French Press at 17,500 psi. The Ni-NTA column was washed with 20 column volumes of 50 mM imidazole in Loading Buffer. The protein was eluted with 250 mM imidazole in Loading Buffer. The purified proteins were dialyzed in Loading Buffer in Pierce Slide-A-Lyzer dialysis cassette (7000 molecular weight cut-off) and stored at -80 $^{\circ}$ C. The protein concentration was estimated by Bio-Rad protein assay using bovine gamma globulin (Pierce) as standard. The purified proteins were analyzed in sodium dodecyl sulfate 10 % polyacrylamide gel electrophoresis (SDS-PAGE) (**Supplementary Fig. 8**).

BioC Purification and Refolding. The C-terminal His-tagged version of BioC was over-expressed as inclusion bodies from strain STL11. The cells were lysed in Lysis Buffer (50 mM

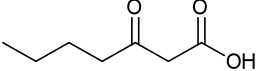
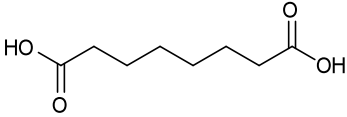
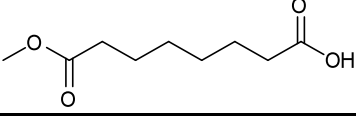
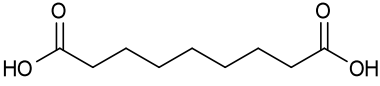
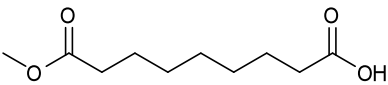
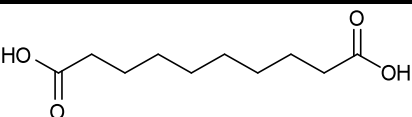
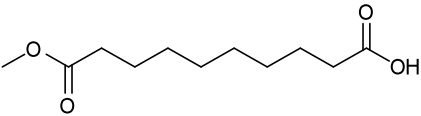
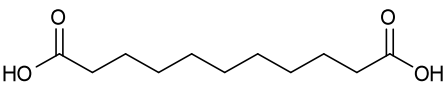
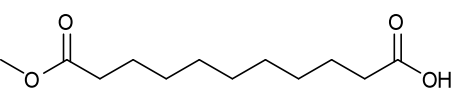
3-(N-morpholino)propanesulfonic acid, pH 7.5, 300 mM NaCl, 10 mM 2-mercaptoethanol, 10 % glycerol, 10 µg/ml of RNase, 10 µg/ml of DNaseI and 10 µg/ml lysozyme) by French Press treatment at 17,500 psi. The insoluble fraction of the cell extract was washed five times in Isolation Buffer (50 mM 3-(N-morpholino)propanesulfonic acid pH 7.5, 1 M NaCl, 0.5 M guanidine-HCl, 10 mM 2-mercaptoethanol, 10 % glycerol, and 5 % Triton X-100) to partially purify BioC inclusion bodies. BioC was solubilized in Solubilization Buffer (50 mM 4-morpholineethanesulfonic acid (pH 6), 1 M NaCl, 4 M guanidine-HCl, 10 % glycerol, 5 mM 2-mercaptoethanol, 5 % sorbitol, 0.05 % Tergitol NP-40, and 10 mM imidazole), purified under denatured condition in Ni-NTA column, and eluted in Solubilization Buffer containing 250 mM imidazole. The denatured BioC was adjusted to 100 µg/ml in 30 ml of Solubilization Buffer and transferred to a 250-ml beaker containing a stir bar. Refolding of BioC was performed by titration by adding 300 ml of Refolding Buffer (Solubilization Buffer lacking guanidine-HCl) to BioC at 25 °C at a rate of approximately 1 drop/sec with constant stirring. The refolded BioC was purified by Ni-NTA column and eluted with Refolding Buffer containing 250 mM imidazole. BioC was dialyzed and stored in Dialysis Buffer (25 mM 4-morpholineethanesulfonic acid, pH 6, 200 mM NaCl, 10 % glycerol and 1 mM 2-mercaptoethanol).

SUPPLEMENTAL RESULTS

Table 1. Compounds tested for bypass of the $\Delta bioC$ mutation in the presence of AaS. The data set we obtained for each of the compounds tested matched that for the compound given in the mass spectral and NMR databases.

Compound	Structure	Source	Analytical	Bypass
Malonic acid		SA ^a	NMR	No
Malonate methyl ester		SA	NMR	Yes
Malonate ethyl ester		This study	GC/MS	Yes
Malonate propyl ester		This study	GC/MS	Yes
Malonate butyl ester		This study	GC/MS	Yes
Succinic acid		SA	NMR	No
Succinate methyl ester		This study	GC/MS	No
Glutaric acid		SA	NMR GC/MS	No
Glutarate methyl ester		SA	NMR GC/MS	Yes
Glutarate ethyl ester		This study	NMR GC/MS	Yes
Glutarate propyl ester		This study	NMR GC/MS	Yes
Glutarate butyl ester		This study	GC/MS	Yes

3-Ketoglutarate methyl ester		Derivative from SA ^b	NMR	Yes
3-Hydroxyglutarate methyl ester		Derivative from SA	NMR	Yes
Glutaconate methyl ester		Derivative from SA	GC/MS	Yes
2-Ketoglutarate methyl ester		Derivative from SA	GC/MS	No
Adipic acid		SA	GC/MS	No
Adipate methyl ester		SA	GC/MS	No
Adipate ethyl ester		This study	GC/MS	No
Hexanoic acid		SA	GC/MS	No
Pimelic acid		SA	NMR GC/MS	No
Pimelate methyl ester		This study	NMR GC/MS	Yes
Pimelate ethyl ester		This study	GC/MS	Yes
Pimelate propyl ester		This study	NMR	Yes
Pimelate butyl ester		This study	GC/MS	Yes
3-Ketopimelate ethyl ester		Derivative from TCI ^c	NMR	Yes
4-Ketopimelate ethyl ester		Derivative from SA	GC/MS	No
Heptanoic acid		SA	NMR	No

3-Ketoheptanoic acid		Derivative from SA ^d	GC/MS	No
Suberic acid		SA	NMR	No
Suberate methyl ester		This study	GC/MS	No
Azelaic acid		SA	NMR	No
Azelate methyl ester		This study	GC/MS	No
Sebacic acid		SA	NMR	No
Sebacate methyl ester		This study	GC/MS	No
Undecanedioic acid		SA	NMR	No
Undecanedioate methyl ester		This study	GC/MS	No

^a Sigma-Aldrich

^b Purchased from Sigma-Aldrich as diester and hydrolyzed to monoester.

^c Purchased from TCI America as diester and hydrolyzed to monoester.

^d Purchased from Sigma-Aldrich as methyl ester and hydrolyzed to acid.

Table 2. Bacterial strains, plasmids and PCR primers.

Strain	Relevant Characteristics	Reference
MG1655	<i>E. coli</i> K-12	CGSC ^a
BM25113	<i>lacI^q rrnB_{T14} ΔlacZ_{WJ16} hsdR514 ΔaraBADAH₃₃ ΔrhaBAD_{LD78}</i>	2
BL21 (DE3)	F ⁻ <i>ompT hsdS_B (r_B m_B) gal dcm</i> (DE3)	Invitrogen
STL11	BL21 (DE3) / pSTL4	This study
STL14	BL21 (DE3) / pSTL6	This study
STL23	MG1655 <i>ΔbioC::FRT</i>	This study
STL24	MG1655 <i>ΔbioH::FRT</i>	This study
STL25	MG1655 <i>ΔbioC::FRT ΔbioH::FRT</i>	This study
STL29	MG1655 <i>ΔbioH::kan</i>	This study
STL32	MG1655 <i>ΔbioC::FRT</i> / pYFJ84	This study
STL33	MG1655 <i>ΔbioH::FRT</i> / pYFJ84	This study
STL34	MG1655 <i>ΔbioC::FRT ΔbioH::FRT</i> / pYFJ84	This study
STL47	BL21 (DE3) / pSTL13	This study
STL49	MG1655 <i>ΔbioH::FRT</i> / pSTL6	This study
STL50	MG1655 <i>ΔbioH::FRT</i> / pSTL13	This study
STL74	MG1655 <i>ΔbioC::FRT ΔbioH::FRT</i> / pSTL20	This study
STL78	MG1655 <i>ΔbioC::FRT</i> / pET16b	This study
STL96	MG1655 <i>ΔbioC::FRT</i> / pSTL25	This study
STL98	MG1655 <i>ΔbioH::FRT</i> / pCY123	This study
STL112	MG1655 <i>ΔbioA::FRT</i>	This study
STL115	MG1655 <i>ΔbioF::FRT</i>	This study
ER90	MG1655 <i>ΔbioF::cat</i>	5,6
YFJ239	BL21 (DE3) / pYFJ84	7
DK574 / pJT93	<i>E. coli</i> holo-ACP over-expression strain	8

Plasmids	Relevant Characteristics	Primer set	Reference
pET28b+	Kan ^R T7 expression vector		Novagen
pET16b	Amp ^R T7 expression vector		Novagen
pBAD322	Amp ^R arabinose-inducible expression vector		9
pSTL4	pET28b+ encoding C-terminal His-tagged <i>E. coli</i> BioC	A07-A08	This study

pSTL6	pET28b+ encoding C-terminal His-tagged <i>E. coli</i> BioH	A11-A12	This study
pSTL13	pSTL6 derivative carrying the S82A <i>bioH</i> mutation	A39-A40	This study
pSTL20	pBAD322 encoding <i>B. subtilis bioW</i>	A41-A55	This study
pSTL25	<i>bioC</i> in-frame deletion (Δ BglIII) of pCY123		This study
pYFJ84	pET16b encoding N-terminal His-tagged <i>V. harveyi</i> AasS		⁷
pCY123	pBR329 carrying the <i>bioABFCD</i> operon		This study

Primers	Sequence (5' to 3')
A07	GGTGCTCCATGGCAACGGTTAATAAAC
A08	GGTGACAAAATAACTCGAGCTCACGAGCAATC
A11	AGAGAACAATAGCTCATGAATAACATCTGGTGG
A12	ATTTCAAAGCCTCGAGCACCCCTCTGC
A15	GAAATGCAGGATATCGACCGTCTGCTGGAGGTGCTGCATGGCAACGGTT AAGTGTAGGCTGGAGCTGCTTC
A16	TTACTCACGAGCAATCACTCCCAAAAAAAGATGATACGTCAGAGGATAT CGATTCCGGGGATCCGTCGACC
A17	ATGAATAACATCTGGTGGCAGACCAAAGGTCAGGGGAATGTTCATCCTG TGATTCCGGGGATCCGTCGACC
A18	CTACACCCTCTGCTTCAACGCCACCAGCAGGTGACAAAACCTCGGCCGGAT GGTGTAGGCTGGAGCTGCTTC
A39	CCATTTGGTTAGGCTGGGCTCTGGGC
A40	CACCAGCCCGCCAGAGCCAGCC
A41	GAAAGGGCGAGCCATGGTGCAAGAAGAAAC
A55	CTTTTTTCTCTCGAGTCATGCGTCATGATCTTCCTC
A68	GCGCTGATGACCCCGGACACCGACGAATCTTACAACGCGGCAGCATAAT GAGTGTAGGCTGGAGCTGCTTC
A69	ATTAACCGTTGCCATGCAGCACCTCCAGCAGACGGTCGATATCCTGCATT TATTCCGGGGATCCGTCGACC
A70	TTCTTGTTTGCAGAAAGTGTAGCCAGAAACCCTCACGCGGACTTCTCGTT AGTGTAGGCTGGAGCTGCTTC
A71	ATGACAACGGCAGATCTTGCCTTTGACCAACGCCATATCTGGCACCCATA CATTCCGGGGATCCGTCGACC

^aCGSC, Coli Genetic Stock Center, Yale University.

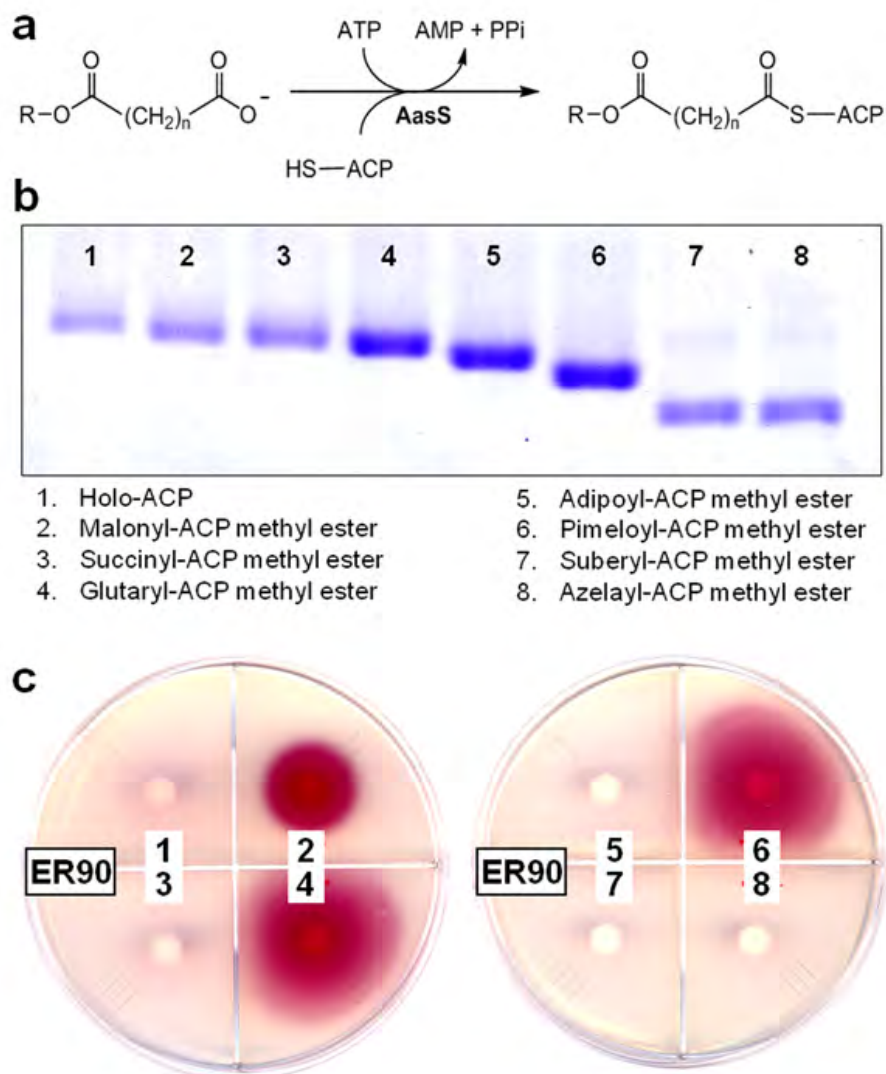


Figure 1. (a) Acylation of ACP with monomethyl esters of various dicarboxylates was carried out using *V. harveyi* AasS. The migration rate of an acyl-ACP in this system is dependent on the length of the acyl chain until a maximum is reached¹⁰. The gel was stained with Coomassie Blue R250 (b). The purified acyl-ACP species of panel B were tested for the ability to support DTB synthesis *in vitro* in the absence of BioC. The acyl-ACPs are numbered as in panel B. (c) Only malonyl-ACP methyl ester (2), glutaryl-ACP methyl ester (4) and pimeloyl-ACP methyl ester (5) supported DTB synthesis.

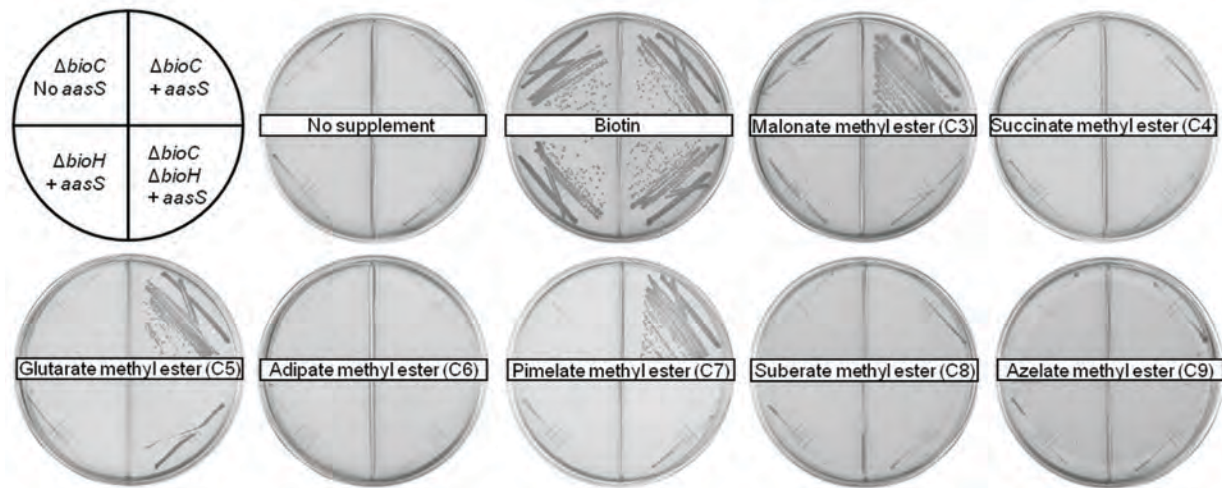


Figure 2. AasS-mediated bypass of the $\Delta bioC$ mutation by supplementation with monomethyl esters of dicarboxylates. Each minimal agar plate contained 0.1 mg/ml sodium ampicillin and 1 mM IPTG. The strain lacking AasS expression carried the empty vector plasmid to provide ampicillin resistance. The methyl esters were tested at nominal concentrations of 0.1 mM except for malonate methyl ester which had a nominal concentration of 1 mM (growth was weak on 0.1 mM). Biotin was added at 4 nM. The plates were incubated overnight at 37 °C. The strains tested were STL78, STL32, STL33 and STL34. Note that the concentrations given are based on the total dicarboxylate of the preparation because the various preparations contain variable levels of the dimethyl esters (which are insoluble in the medium) and the free dicarboxylic acids (which are inactive as AasS substrates). However, the monomethyl ester was the major species in all cases (e.g., **Supplementary Fig. 7**).

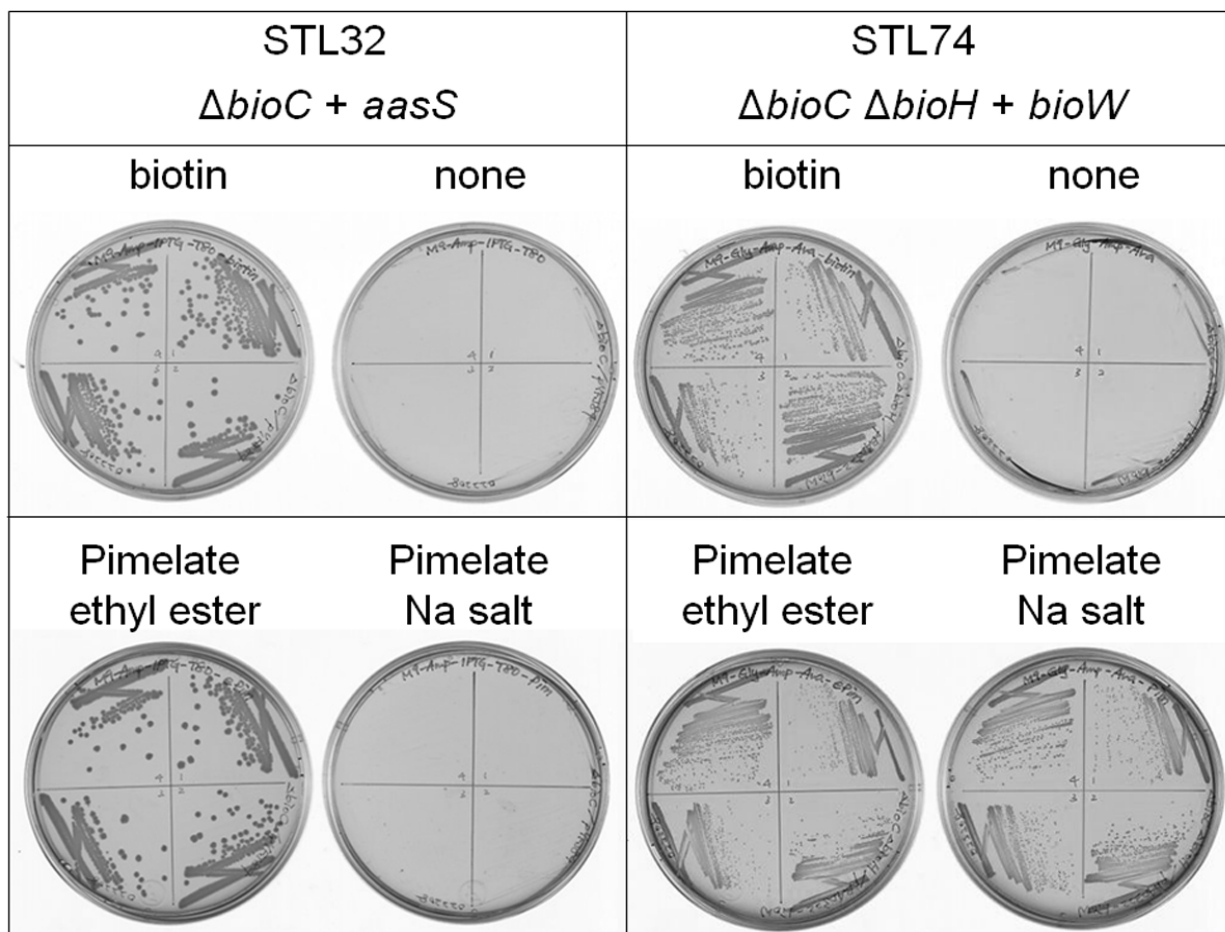
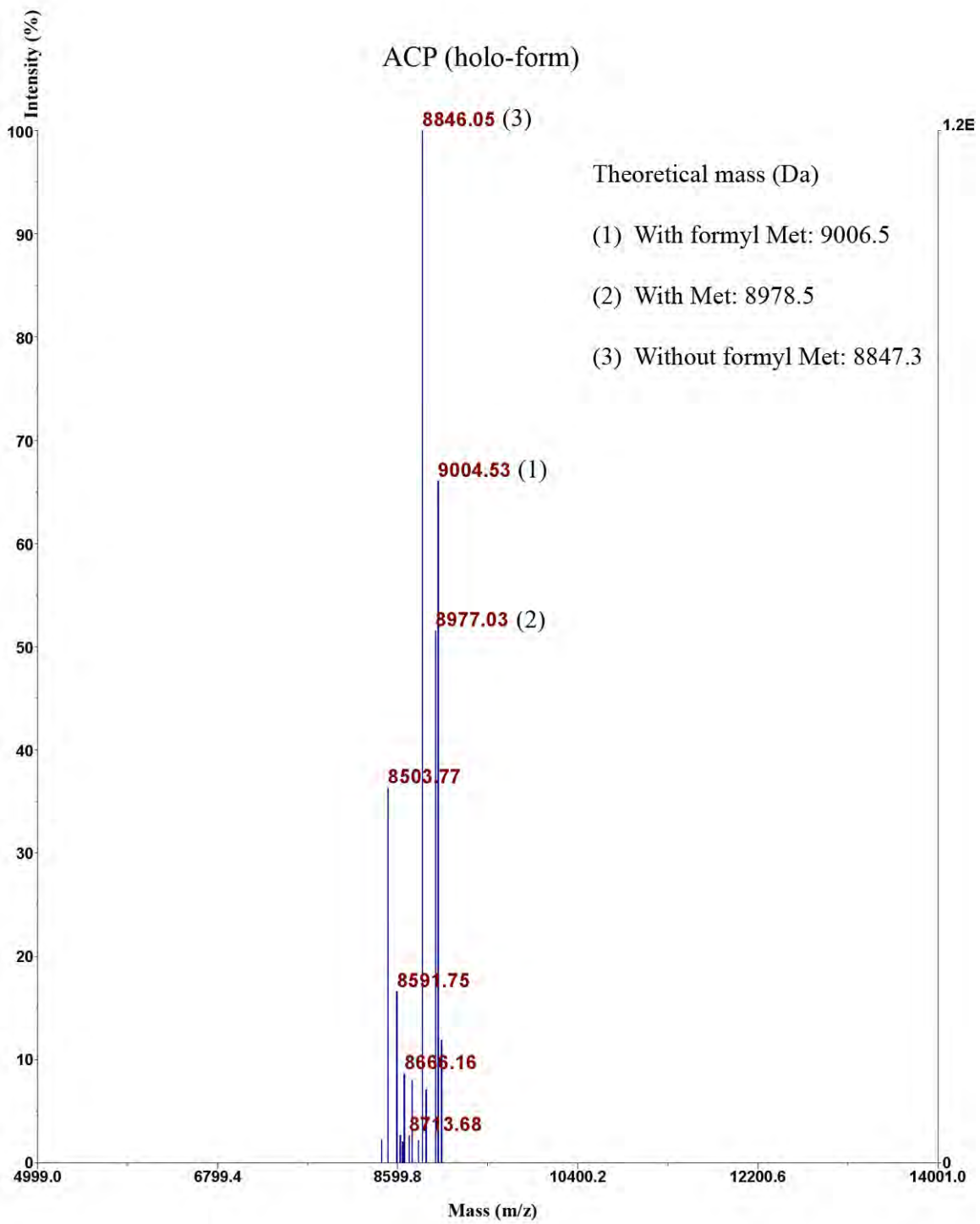
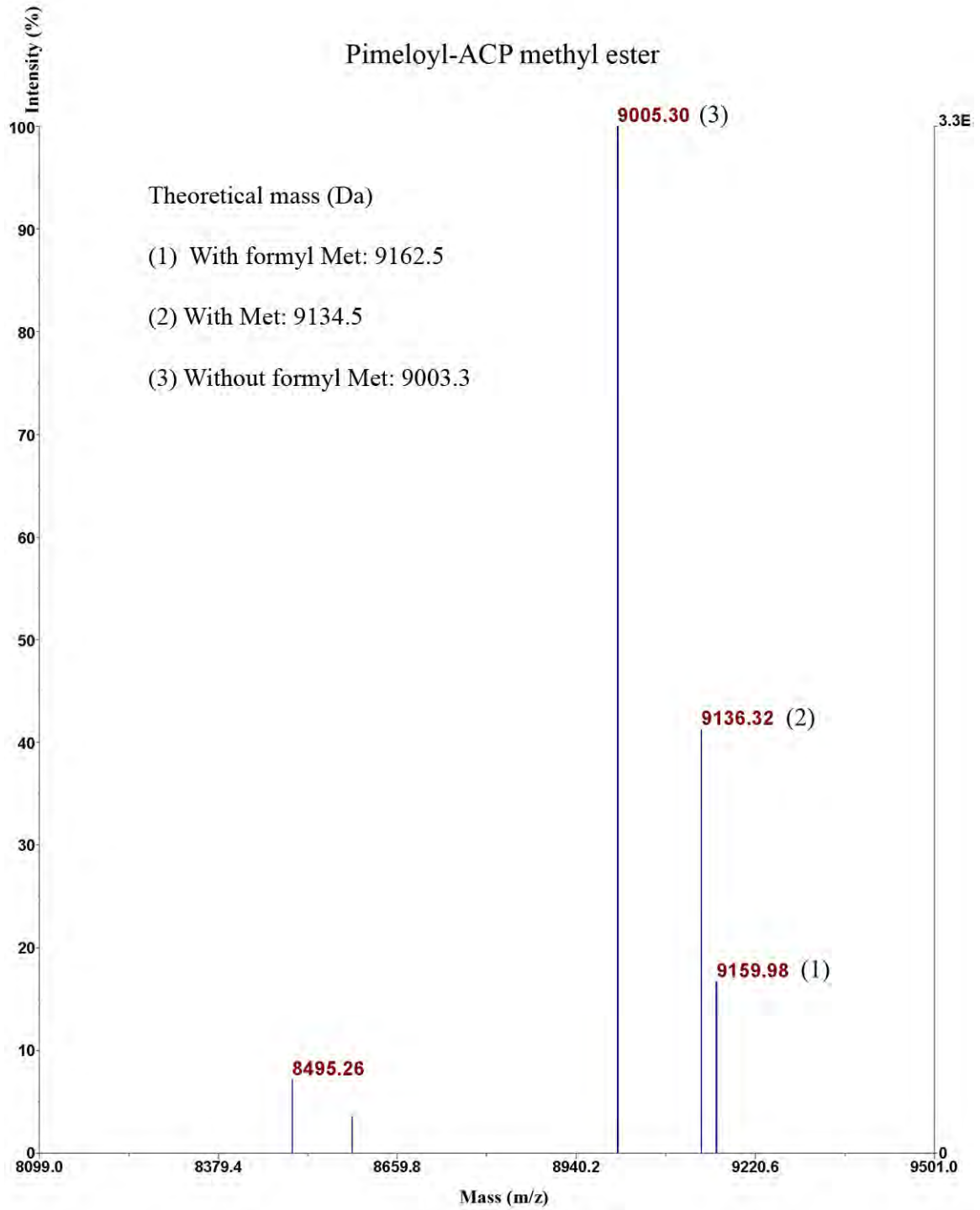
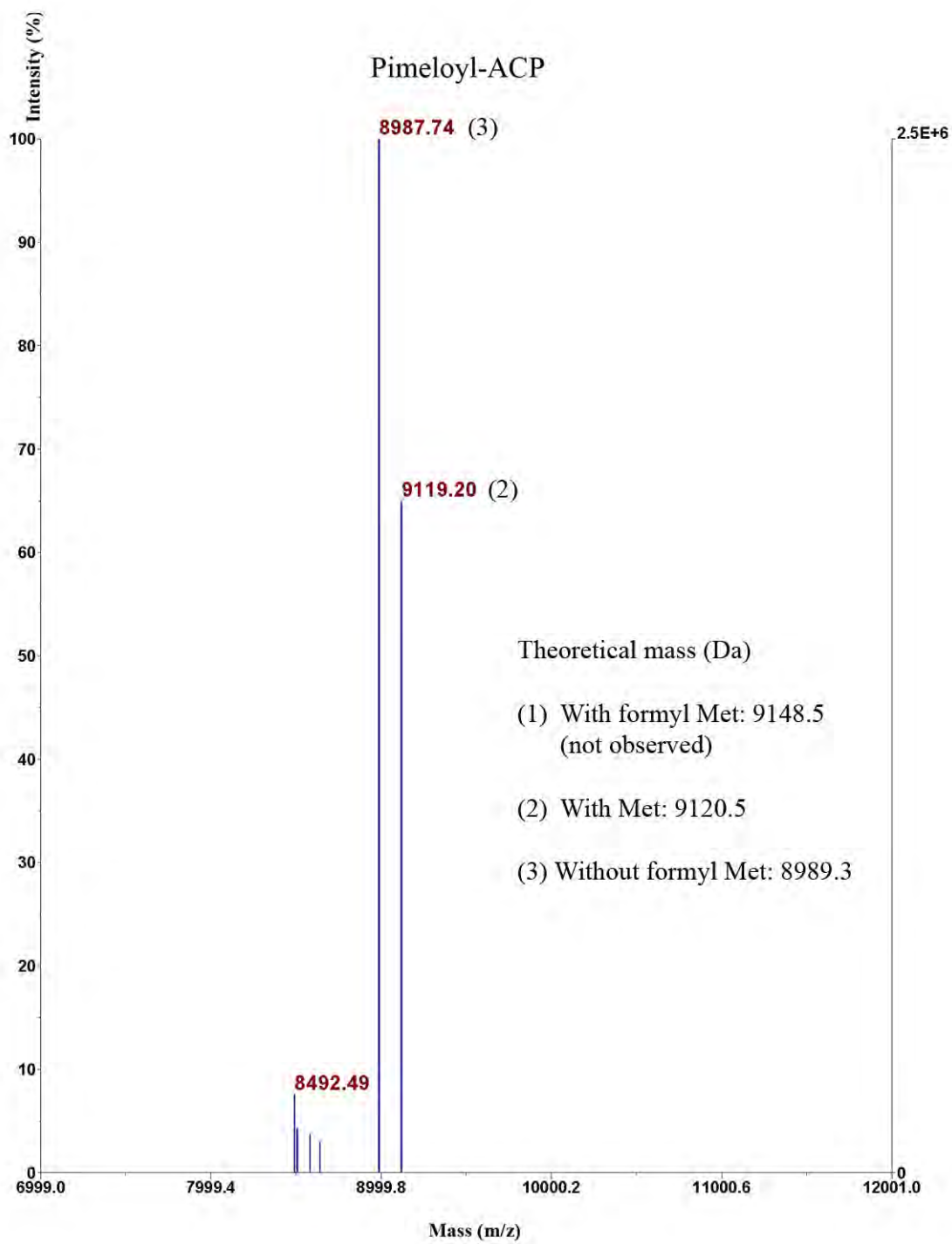


Figure 3. The permeability of derivatives *E. coli* strains STL32 and STL74 carrying plasmids encoding either AasS or the *B. subtilis* BioW were tested. Four random colonies of the each strain were streaked on minimal agar supplemented with 0.1 mM pimelate sodium salt and 0.2 % arabinose for induction. The expression of BioW, a pimeloyl-CoA synthetase, allows bypass of both *bioC* and *bioH* in the presence of pimelic acid whereas pimelic acid was unable to support growth of strain STL32 because pimelic acid is not a substrate for AasS.

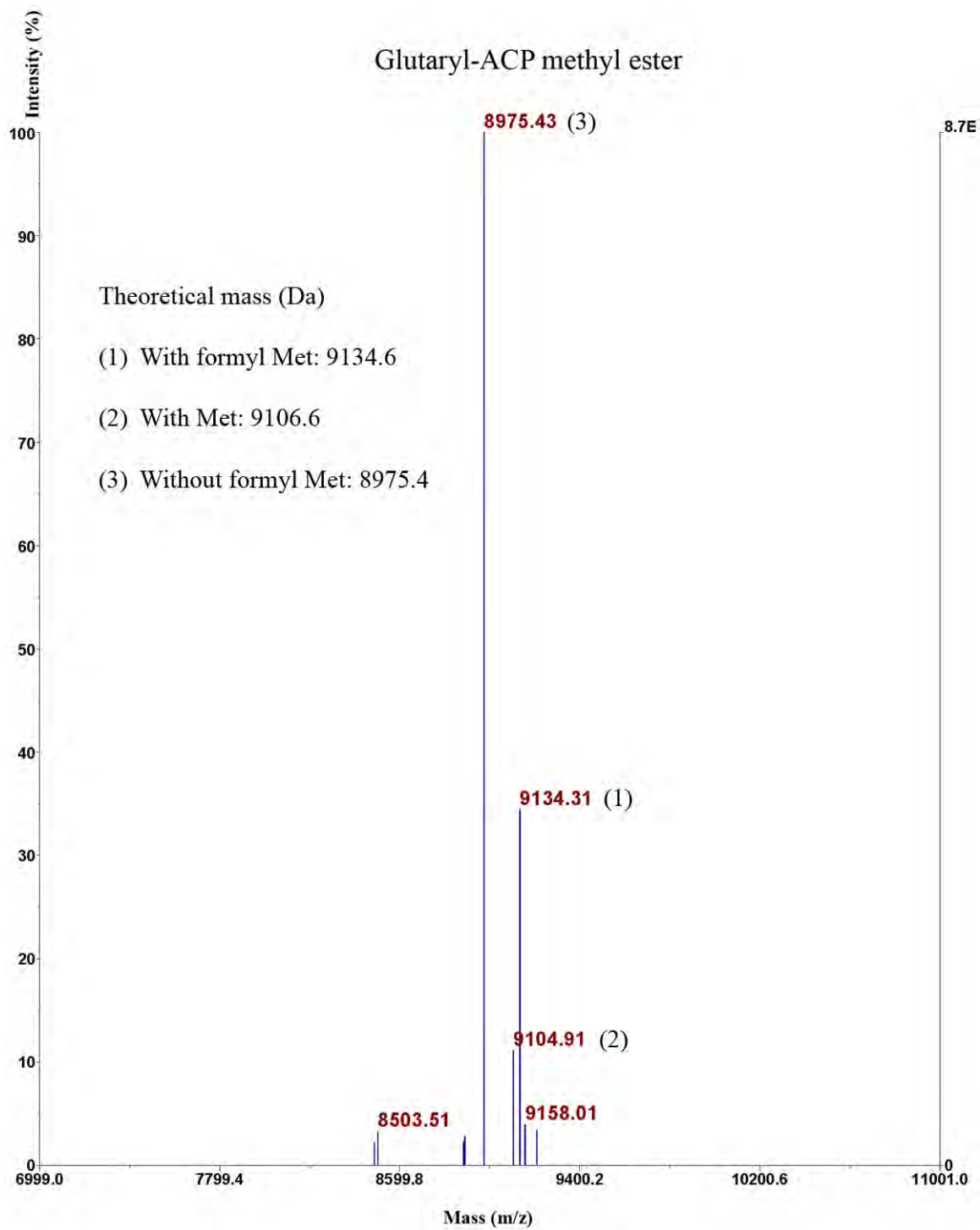


Pimeloyl-ACP methyl ester





Glutaryl-ACP methyl ester



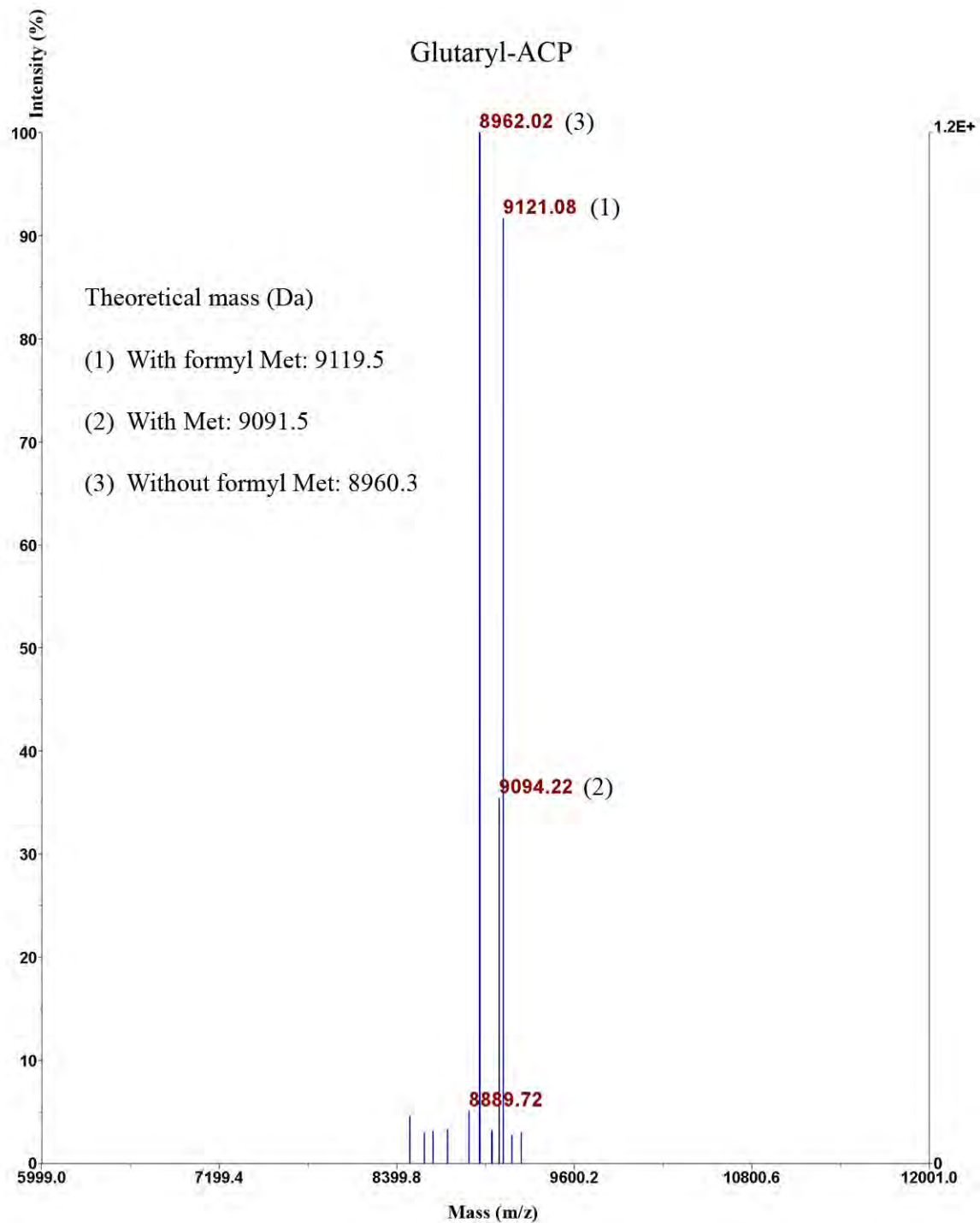


Figure 4. Analysis of pimeloyl-ACP and glutaryl-ACP and their methyl esters by MALDI/MS. A portion of the ACP retains the N-terminal N-formylmethionine or methionine of the primary translation product presumably due to titration of peptide deformylase or methionine aminopeptidase upon ACP overexpression.

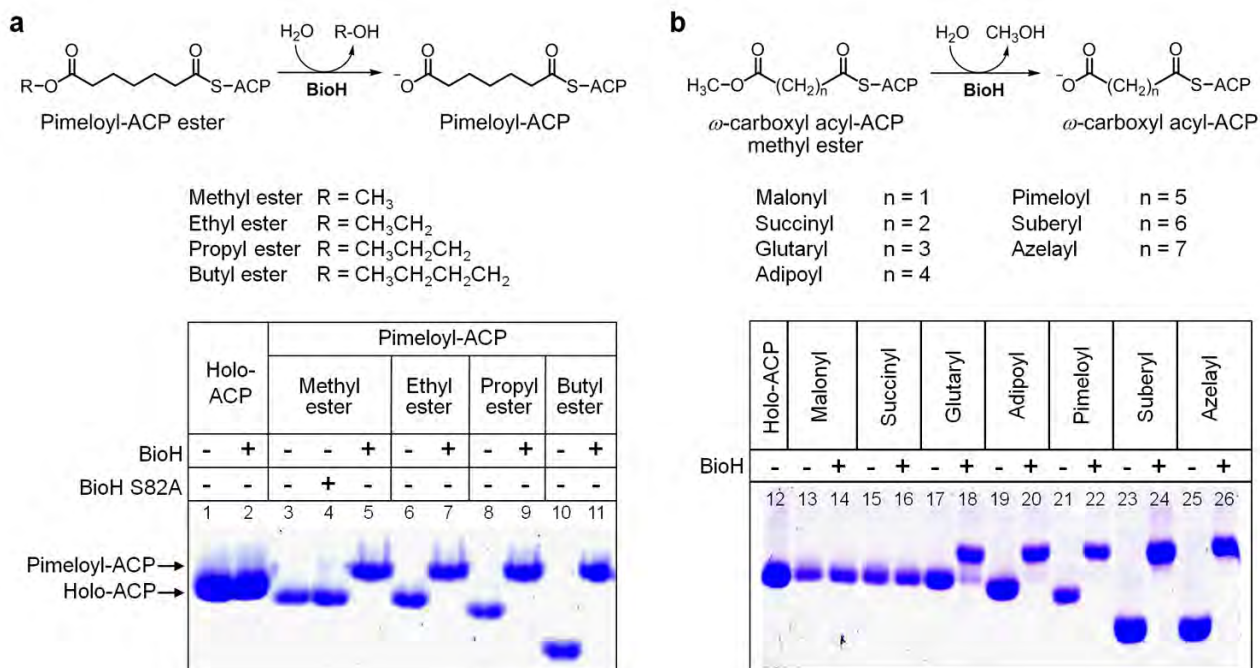


Figure 5. (a) BioH catalyzed hydrolysis of the ester bonds of pimeloyl-ACP esters. The reactions contained 50 mM Tris-HCl (pH 7.0), 5 % glycerol, 5 μ g/ml BioH and 2 mM pimeloyl-ACP esters and were incubated at 37 $^{\circ}$ C for 1 h. Hydrolysis results in the slower migrating pimeloyl-ACP as demonstrated in a conformationally-sensitive electrophoretic mobility assay as described in Experimental Procedures. Mutant BioH S82A lacked the catalytic nucleophile (see text) and was unable to hydrolyze the ester bonds (lane 4). (b) BioH cleavage of the methyl ester moieties of other ω -carboxyl acyl-ACP methyl esters of acyl chain length of 3 to 9 carbon atoms. Hydrolysis of the methyl ester bond results in the slower migrating ω -carboxyl acyl-ACPs in all cases.

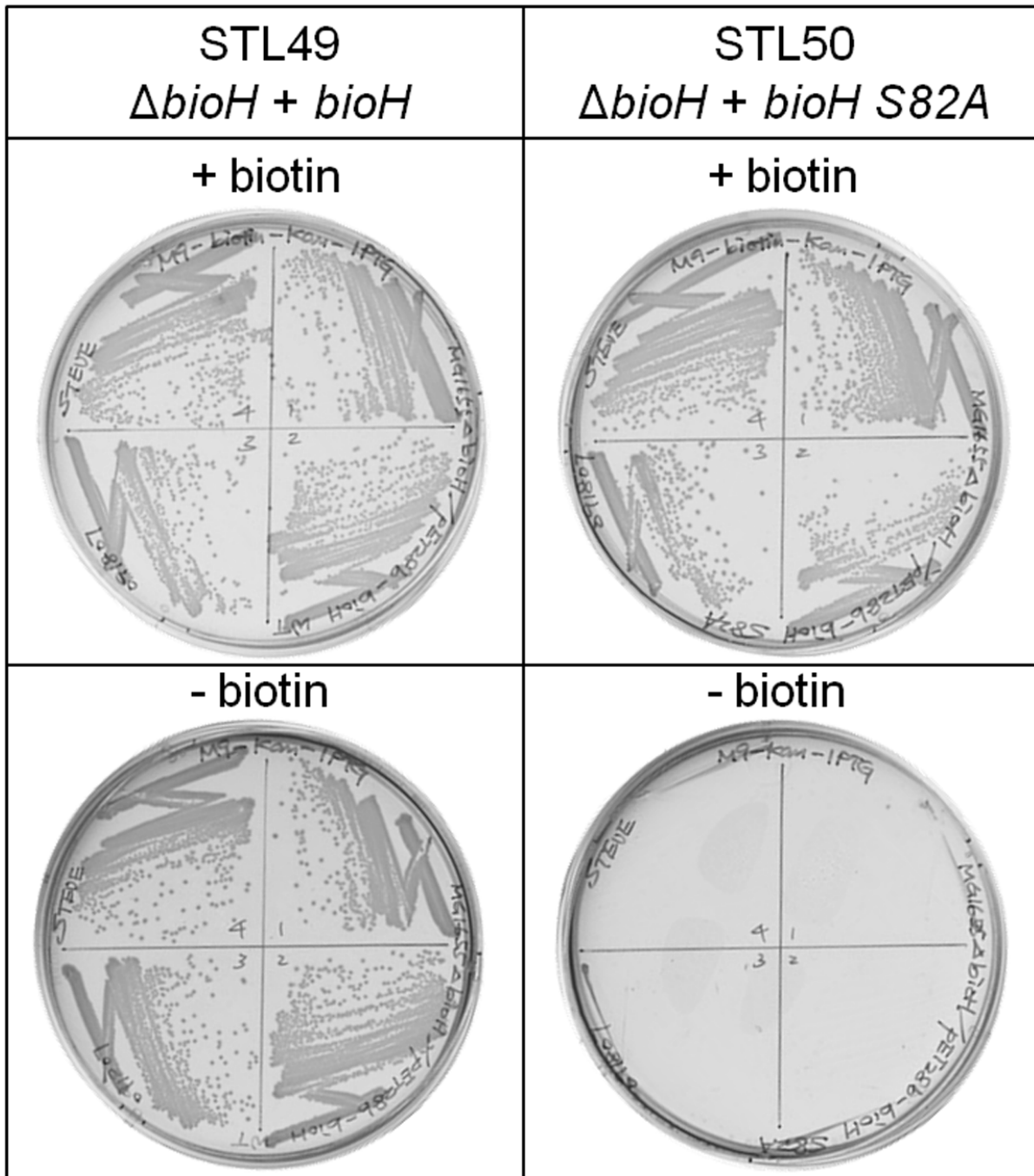


Figure 6. Complementation by expression of BioH or BioH S82A from multi-copy plasmids in a *ΔbioH* strain. Four random colonies of the each strain were streaked on minimal agar supplemented with 0.1 mM IPTG for induction. The expression of the wild type BioH allowed growth on biotin-free medium whereas the expression of BioH S82A failed to restore growth.

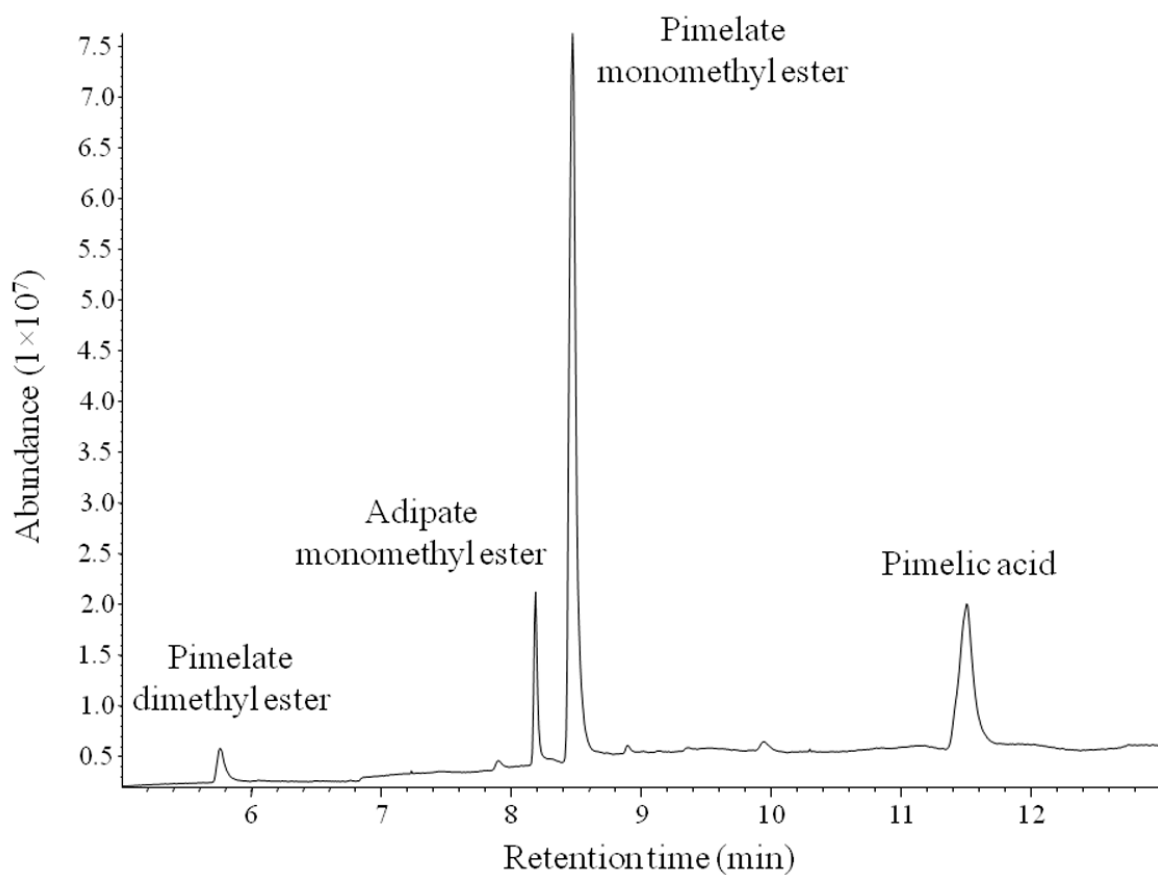


Figure 7. Products of the esterification of pimelate with methanol as determined by GC/MS. Monomethyl adipate was added as internal standard at 14 mM to the synthesized monomethyl pimelate mixture. Monomethyl pimelate was 68.9 % of the total pimelate as calculated from the peak areas of the monomethyl esters of adipate and pimelate.

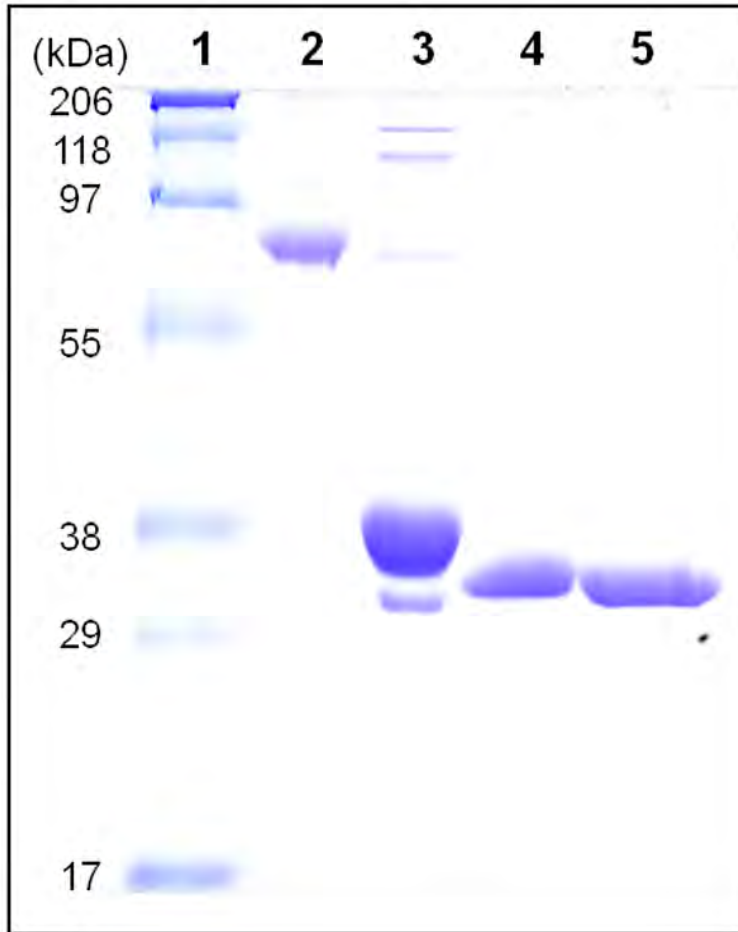


Figure 8. Analysis of purified proteins by sodium dodecyl sulfate 10 % polyacrylamide gel electrophoresis. Lane 1, protein standards (Bio-Rad); lane 2, AaS; lane 3, BioC; lane 4, BioH; and Lane 5, BioH S82A. The proteins were visualized by staining the gel with Coomassie Blue R250.

References:

- 1 Clarke, L. & Carbon, J. A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E. coli* genome. *Cell* **9**, 91-99 (1976).
- 2 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**, 6640-6645 (2000).
- 3 Saitoh, M., Fujisaki, S., Ishii, Y. & Nishiguchi, T. Convenient selective monoesterification of alpha, omega-dicarboxylic acids catalyzed by ion-exchange resins. *Tetrahedron Lett* **37**, 6733-6736 (1996).
- 4 Furniss, B. S., Hannaford, A. J., Smith, P. W. G. & Tatchell, A. R. *Vogel's textbook of practical organic chemistry*. 5th edn, (Longman Scientific & Technical, 1989).
- 5 Choi-Rhee, E. & Cronan, J. E. A nucleosidase required for in vivo function of the S-adenosyl-L-methionine radical enzyme, biotin synthase. *Chem Biol* **12**, 589-593 (2005).
- 6 Choi-Rhee, E. & Cronan, J. E. Biotin synthase is catalytic in vivo, but catalysis engenders destruction of the protein. *Chem Biol* **12**, 461-468 (2005).
- 7 Jiang, Y., Chan, C. & Cronan, J. E. The soluble acyl-acyl carrier protein synthetase of *Vibrio harveyi* B392 is a member of the medium chain Acyl-CoA synthetase family. *Biochemistry* **45**, 10008-10019 (2006).
- 8 Cronan, J. E. & Thomas, J. Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. *Methods Enzymol* **459**, 395-433 (2009).
- 9 Cronan, J. E. A family of arabinose-inducible *Escherichia coli* expression vectors having pBR322 copy control. *Plasmid* **55**, 152-157 (2006).
- 10 Jiang, Y., Morgan-Kiss, R. M., Campbell, J. W., Chan, C. H. & Cronan, J. E. Expression of *Vibrio harveyi* acyl-ACP synthetase allows efficient entry of exogenous fatty acids into the *Escherichia coli* fatty acid and lipid A synthetic pathways. *Biochemistry* **49**, 718-726 (2010).