A new enzyme superfamily – the phosphopantetheinyl transferases

Ralph H Lambalot[1], Amy M Gehring[1], Roger S Flugel[1,2], Peter Zuber[3], Michael LaCelle[3], Mohamed A Marahiel[4], Ralph Reid[5], Chaitan Khosla[6] and Christopher T Walsh[1]

Background: All polyketide synthases, fatty acid synthases, and non-ribosomal peptide synthetases require posttranslational modification of their constituent acyl carrier protein domain(s) to become catalytically active. The inactive apo-proteins are converted to their active holo-forms by posttranslational transfer of the 4′-phosphopantetheinyl (P-pant) moiety of coenzyme A to the sidechain hydroxyl of a conserved serine residue in each acyl carrier protein domain. The first P-pant transferase to be cloned and characterized was the recently reported *Escherichia coli* enzyme ACPS, responsible for apo to holo conversion of fatty acid synthase. Surprisingly, initial searches of sequence databases did not reveal any proteins with significant peptide sequence similarity with ACPS.

Results: Through refinement of sequence alignments that indicated low level similarity with the ACPS peptide sequence, we identified two consensus motifs shared among several potential ACPS homologs. This has led to the identification of a large family of proteins having 12–22 % similarity with ACPS, which are putative P-pant transferases. Three of these proteins, *E. coli* EntD and o195, and *B. subtilis* Sfp, have been overproduced, purified and found to have P-pant transferase activity, confirming that the observed low level of sequence homology correctly predicted catalytic function. Three P-pant transferases are now known to be present in *E. coli* (ACPS, EntD and o195); ACPS and EntD are specific for the activation of fatty acid synthase and enterobactin synthetase, respectively. The apo-protein substrate for o195 has not yet been identified. Sfp is responsible for the activation of the surfactin synthetase.

Conclusions: The specificity of ACPS and EntD for distinct P-pant-requiring enzymes suggests that each P-pant-requiring synthase has its own partner enzyme responsible for apo to holo activation of its acyl carrier domains. This is the first direct evidence that in organisms containing multiple P-pant-requiring pathways, each pathway has its own posttranslational modifying activity.

Addresses: [1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA, [2]Committee on Higher Degrees in Biophysics, Harvard University, Cambridge, MA 02138, USA, [3]Department of Biochemistry and Molecular Biology, Louisiana State University Medical Center, Shreveport, LA 71130, USA, [4]Biochemie, Fb Chemie, Philipps Universitat Marburg, Hans-Meerwein-Strasse, D35032 Marburg Germany, [5]Biomolecular Resource Center, University of California, San Francisco, Surge 104-Box 0541, San Francisco, CA 94143, USA and [6]Department of Chemical Engineering, Stanford University, Stanford, CA 94305, USA.

Correspondence: Christopher T Walsh
E-mail: walsh@walsh.med.harvard.edu

Key words: ACP, acyl carrier protein, biosynthesis, non-ribosomal peptide synthetase, phosphopantetheine

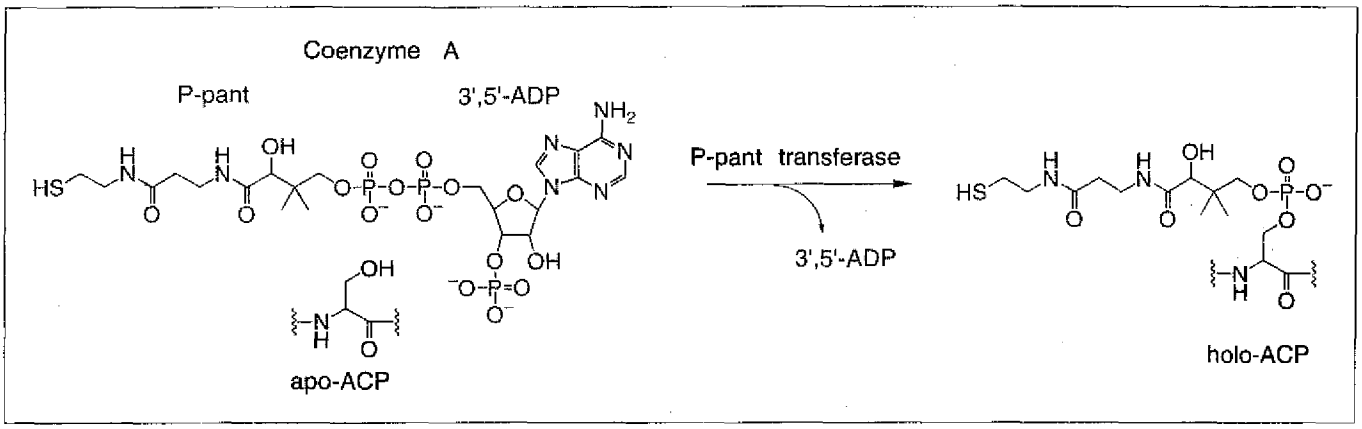
Received: **20 September 1996**
Revisions requested: **10 October 1996**
Revisions received: **22 October 1996**
Accepted: **22 October 1996**

Chemistry & Biology
November 1996, **3**:923–936

Introduction

Multienzyme complexes exist for acyl group activation and transfer reactions in the biogenesis of fatty acids, the polyketide family of natural products (e.g. erythromycin and tetracycline), and almost all non-ribosomal peptides (e.g. vancomycin, cyclosporin, bacitracin and penicillin). All of these complexes contain one or more small proteins, ~80–100 amino acids (aa) long, either as separate subunits or as integrated domains, that function as carrier proteins for the growing acyl chain. These acyl carrier protein(ACP) domains, which may be one of the domains of a multi-functional enzyme (in the type I synthases) or a separate subunit (in the type II multienzyme complex synthases), can be recognized by the conserved sequence signature motif (L,V)(G,L)(G,A,F,Y)(D,H,K,E)S(L,Q)(D,A,G) [1]. They are converted from inactive apo-forms to functional holo-forms by attack of the β-hydroxy sidechain of the conserved serine residue in the ACP signature sequence on the pyrophosphate linkage of coenzyme A (CoASH). This results in transfer of the 4′-phosphopantetheinyl (P-pant) moiety of CoASH onto the attacking serine (Fig. 1). The newly introduced -SH of the P-pant prosthetic group now acts as a nucleophile for acylation by a substrate, which may be acyl-CoA or malonyl-CoA derivatives for the fatty acid and polyketide synthases (PKS), or aminoacyl-AMPs for the peptide and depsipeptide synthetases (Fig. 2). In the PKS complexes the carboxy-activated malonyl-ACP derivative then undergoes decar-boxylation, forming a nucleophilic carbanion species that attacks a second acyl thiolester to yield a new carbon– carbon bond in one of the steps of polyketide biosynthesis. In peptide and depsipeptide

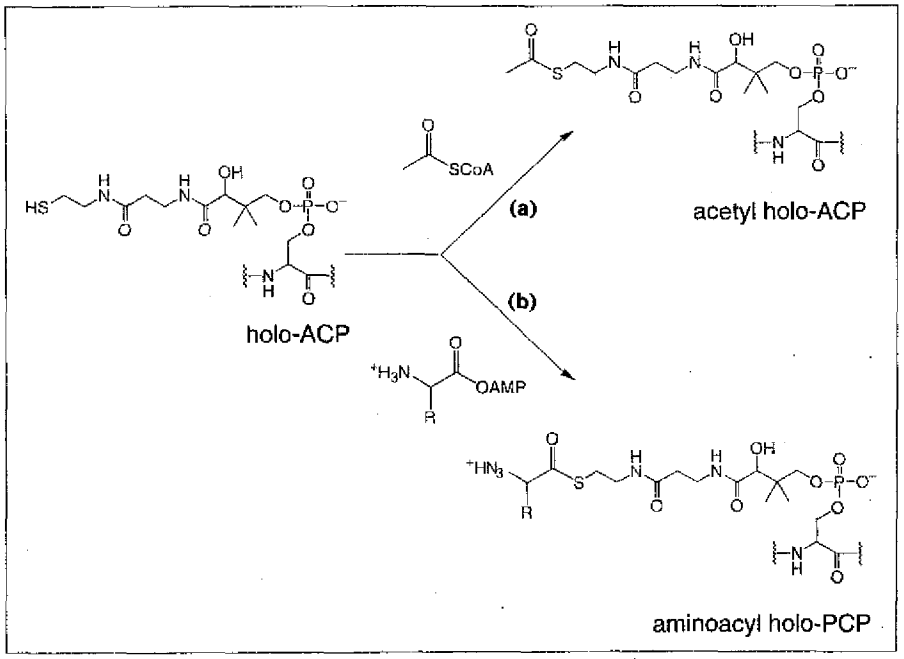
Figure 1



General reaction scheme for posttranslational phosphopantetheinylation. P-pant transferases transfer the 4′-phosphopantetheine moiety from CoA to a conserved serine residue of apo-ACP to produce holo-ACP and 3′,5′-ADP.

synthetases, the aminoacyl-ACPs or hydroxyacyl-ACPs serve as nucleophiles in amide and ester bond-forming steps respectively (Fig. 3). The posttranslational phosphopantetheinylation of apo-ACP domains is clearly essential for the activity of the multienzyme synthases responsible for the biogenesis of a vast array of natural products. We have therefore searched for and characterized enzymes with P-pant transferase activity. We recently reported the cloning and charac-terization of the first such transferase, the *Escherichia coli* holo-acyl carrier protein synthase (ACPS), which activates the fatty acid synthase ACP by converting it to its holo-form [2]. Using the conversion of *E. coli* apo-ACP to holo-ACP as an assay, we purified ACPS 70 000-fold and identified it as the product of a previously described essential *E. coli* gene of unknown function, *dpj* [3]. The *E. coli* ACPS is a 28 kDa dimer of two 125-aa subunits with a k_{cat} of 80–100 min^{-1} and a $K_M \leq 10^{-6}$ M for apo-ACP. We subsequently showed that the *E. coli* ACPS will also modify apo-forms of several type II ACP homologs including the *Lactobacillus casei* D-alanyl carrier protein (DCP) involved in D-alanylation of lipoteichoic acid [4], the *Rhizobia* protein, NodF, involved

Figure 2



The terminal cysteamine thiol of the phosphopantetheine cofactor acts as a nucleophile for acyl activation. **(a)** Fatty acid synthases and polyketide synthases transfer acyl groups from acyl-CoA to the phosphopantetheine tether attached to ACP. **(b)** Non-ribosomal peptide and depsipeptide synthetases first activate their amino-acyl or acyl substrates as their acyl-adenylates before transfer to the phosphopantetheine tether of PCP.

in the acylation of the oligosaccharide-based nodulation factors [5], and the *Streptomyces* ACPs involved in frenolicin, granaticin, oxytetracycline, and tetracenomycin polyketide antibiotic biosynthesis (AMG, RHL and CTW, unpublished results).

The *E. coli* ACPS does not detectably transfer P-pant to the apo-forms of two type I P-pant-requiring proteins involved in amino acid activation, namely apo-EntF which is involved in L-serine activation during *E. coli* enterobactin biosynthesis [6,7] and apo-PCP, a peptidyl carrier protein fragment from the *Bacillus brevis* tyrocidine synthetase (TycA) [8]. Thus other P-pant transferases, specific for the apo-forms of type I peptide synthetases, must exist. Our search in the completely sequenced *Haemophilus influenzae* [9] and *Saccharomyces cerevisiae* genomes for functional homologs of *E. coli acpS* initially failed to reveal genes with any apparent homology despite the fact that posttranslational phosphopantetheinylation of ACP domains clearly occurs in these organisms. We report here that more refined database searches yielding peptide sequences with only marginal similarity to ACPS, have in fact led us to identify a large second family of P-pant transferases including the *E. coli* EntD and *B. subtilis* Sfp proteins. The genes encoding these proteins have pre-viously been shown to be required for the production of the non-ribosomal peptides enterobactin and surfactin, respectively (Fig. 4) [10,11]. Putative P-pant transferases have also been identified in *H. influenzae* and *S. cerevisiae* (Fig. 5 and Table 1). We have overproduced and purified EntD, Sfp and a third *E. coli* protein o195 and have demonstrated the ability of each to catalyze the transfer of 4′-phosphopantetheine from CoASH to apo-protein substrates.

Results

Database search for ACP synthase homologs

BLAST searches (basic local alignment search tool) [12] with the 125-aa *E. coli* ACPS protein sequence revealed marginal similarity to the carboxy-terminal region of five fungal fatty acid synthases, suggesting that phosphopantetheinylation activity may have been subsumed as a domain in these polyenzymes (Fig. 5). We propose a scheme, based on several lines of genetic evidence [13–18], in which the carboxyl-terminus of the FAS2 subunit could be responsible for the autophosphopantetheinylation of the amino-terminal ACP domain. However, to date we have been unable to demonstrate P-pant transfer from CoASH to the *S. cerevisiae* FAS2 ACP domain (residuesAsp142–Ser230) catalyzed by the putative P-pant trans-ferase domain (residues Gly1774–Lys1894) (data not shown).

Using the small similarity between the fungal FAS2 carboxyl-termini and ACPS as a starting point, we detected potential homology to three bacterial proteins, EntD *(E. coli)*, Sfp *(B. subtilis)*, and Gsp *(B. brevis)* which have previously been identified as genes that appear to have a common ancestor (orthologous genes) (Fig. 5) [19]. Indeed *E. coli entD* and *Bacillus brevis gsp* can complement *sfp* mutants, supporting the idea that these three proteins have similar functions [19,20]. The specific biochemical

Figure 3

Acyl-pantetheinyl thiolesters have a wide variety of fates in the biosynthesis of complex natural products. Acyl-pantetheinyl thiolesters can act as **(a)** carbanion nucleophiles for carbon skeleton assembly in fatty acid and polyketide biosynthesis or as **(b)** nitrogen or **(c)** oxygen nucleophiles to yield amide or ester bonds in peptide and depsipeptide biosynthesis.

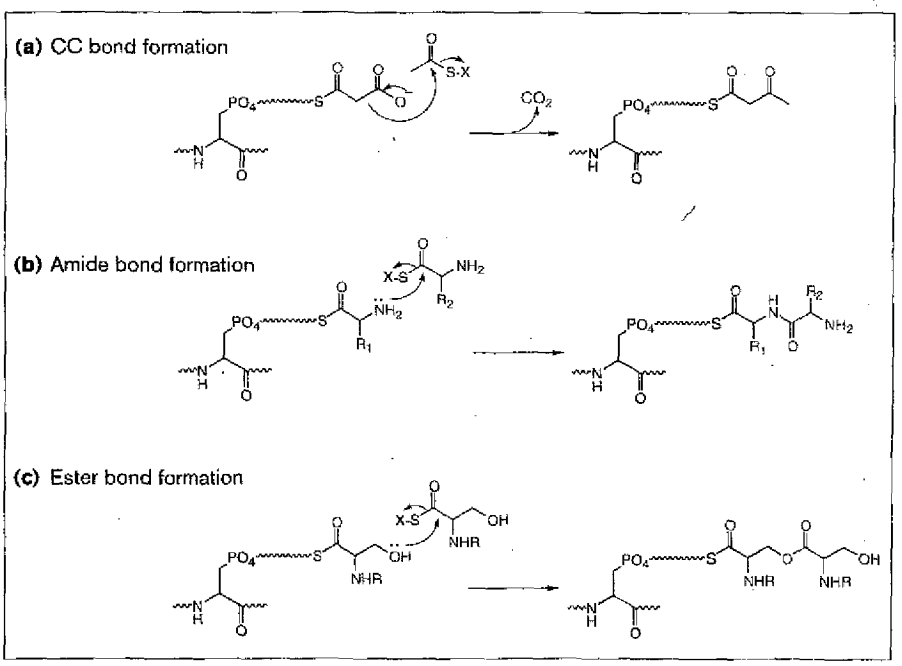
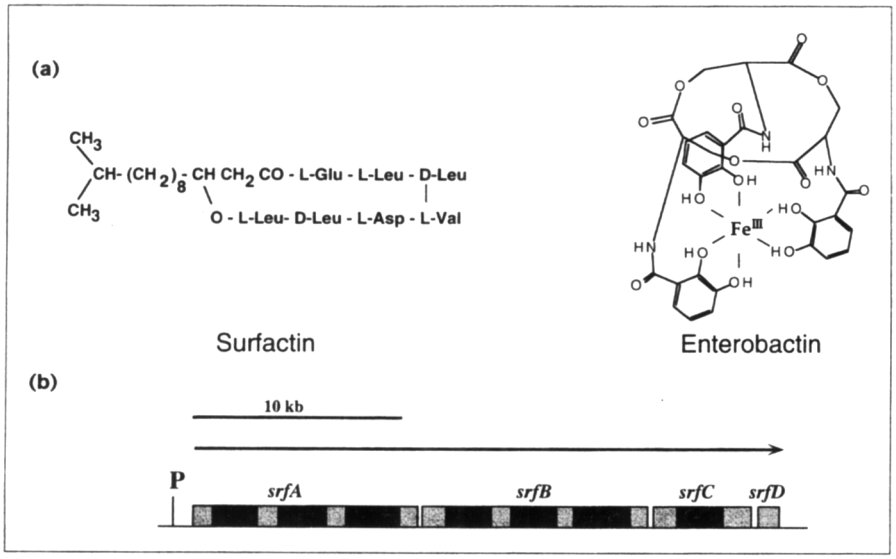


Figure 4



Non-ribosomal peptides and some of the genes involved in their synthesis. **(a)** Chemical structures of surfactin and enterobactin. **(b)** The *srf* operon consists of four open reading frames in which *srfA*, *srfB*, and *srfC* encode for the activities that activate and assemble the seven component amino acids and branched chain β-hydroxy fatty acid of surfactin.

functions of *entD*, *sfp* and *gsp* have up to now remained obscure. *Sfp* was isolated as a locus required for production of the lipopeptide antibiotic surfactin in *B. subtilis* (Fig. 4) [11] and *gsp* is similarly required for gramicidin biosynthesis [19]. Likewise, *entD* has been shown to be required for production of the Fe^{III}-chelating siderophore enterobactin in *E. coli* [10]. Further BLAST searches revealed several other proteins that share potential homology with ACPS (Table 1), including a third *E. coli* open reading frame (in addition to ACPS and EntD) of unknown function designated o195 and proteins involved in cyanobacterial heterocyst differentiation and fungal lysine biosynthesis. Local sequence alignments of the putative P-pant transferase domains reveal two sequence motifs containing several highly conserved residues (Fig. 5, highlighted in yellow).

Confirmation of sequence-predicted P-pant transferase activity

To test the sequence-predicted P-pant transferase activity of this enzyme family, we needed to overproduce and purify representative members of this family (EntD, Sfp and o195), prepare apo-forms of putative substrate proteins or subdomains (ACP, PCP, EntF, and SrfB1) and assay the catalytic competence of the putative enzymes.

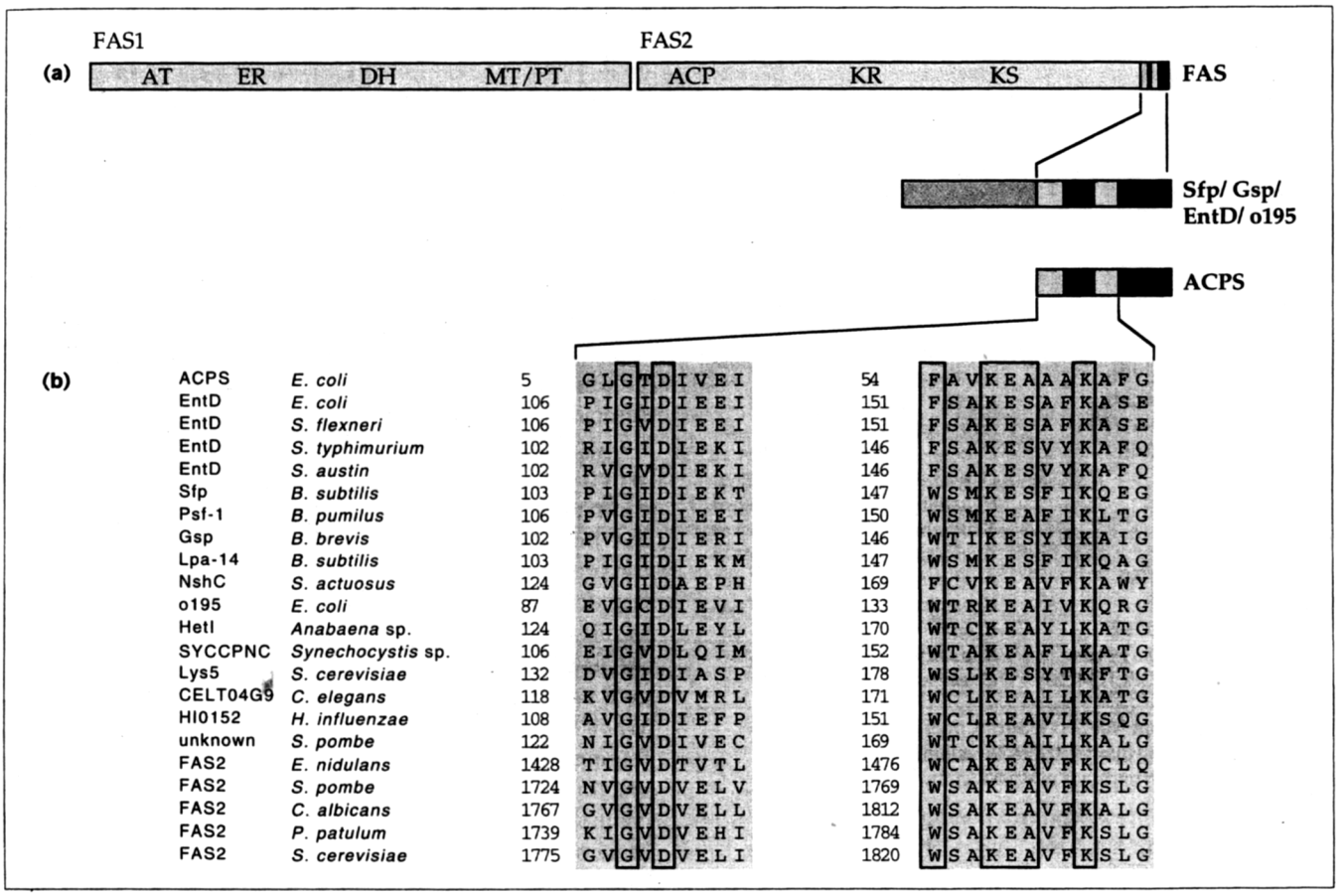
Overproduction, purification and characterization of enzymes

Sfp (26.1 kDa) was overproduced and purified using previously published procedures (Fig. 6) [11]. EntD (23.6 kDa) had previously been cloned, but its overproduction had proven difficult, presumably due to the frequency of rare codons and an unusual UUG start codon [10]. We therefore changed the UUG start to AUG and optimized the codon usage for the first six residues. The *entD* gene was PCR-amplified from wild type *E. coli* and cloned into the T7-promoter-based pET28b expression plasmid (Novagen). Induction at 25° C yielded soluble EntD, which was purified by ammonium sulfate precipitation and Sephacryl S-100 chromatography. Similarly, the o195 gene was PCR-amplified from wild type *E. coli* cells with codon optimization and cloned into pET28b. Induction at 37°C or 25° C yielded predominantly insoluble o195 protein (21.8 kDa), that could be solubilized in 8 M urea, purified by Q-Sepharose chromatography under denaturing conditions, and renatured by dialysis.

Overproduction, purification and characterization of substrates

Apo-ACP and apo-EntF were overproduced and purified as previously described [7] [21]. Apo-PCP (the peptidyl carrier protein of tyrocidine synthetase, see Fig. 7) and apo-SrfB1 (the first amino acid activation and peptidyl carrier protein domains of surfactin synthetase subunit B) were overproduced in *E. coli* and purified as hexahistidine-tagged proteins using nickel chelate chromatography. Typically, when P-pant-requiring enzymes are over-produced in *E. coli* the fraction of recombinant protein that is modified to the holo-form represents only a small percentage of the total recombinant protein [22]. We have been able to confirm that the percentage of holo-ACP present in the purified preparation is below 5 % by using analytical HPLC to resolve the apo and holo-forms of the protein (data not shown) [23]. The ratio of apo- to holo-forms of the other substrates after purification was not precisely determined. It is clear, however, as shown below, that sufficient quantities of the apo-forms of each of these proteins were obtained to act as substrates of the P-pant transferase enzymes. P-pant transferase activity toward each of these substrates was assayed by monitoring

Figure 5



The putative phosphopantetheinyl transferase family. **(a)** Schematic showing location of the proposed P-pant transferase domains (purple) and location of consensus sequences (yellow) in the fungal fatty acid synthases (FAS), the Sfp/Gsp/EntD/o195 homology family, and *E. coli* ACPS. Component FAS activities are abbreviated as AT, acyl transferase; ER, enoyl reductase; DH, dehydratase; MT/PT malonyl/palmitoyl transferase; ACP, acyl carrier protein; KR, ketoreductase; KS, ketosynthase. **(b)** Local DNA sequence alignments of the consensus sequences of the P-pant transferase enzyme superfamily. Highly conserved residues are boxed.

the transfer of [^{3}H]-4′-phosphopantetheine from [^{3}H]-(pantetheinyl)-CoASH in the presence of the putative P-pant transferase enzyme. Reactions were quenched with 10 % trichloroacetic acid (TCA), and the resulting protein pellet was washed, resolubilized, and counted by liquid scintillation to determine the extent to which the apo-substrate was modified to the holo-form by the covalent attachment of [^{3}H]-4′-phosphopantetheine.

Enzymatic activity with apo-ACP and apo-PCP as substrates

We were initially concerned that large proteins such as EntF (140 kDa) and SrfB (400 kDa) would be difficult to work with as substrates for the preliminary characterization of the putative P-pant transferases. Indeed our prior attempts to modify purified EntF with ACPS had been unsuccessful (RHL, RSF and CTW, unpublished results). Studies with the large, multifunctional chicken fatty acid synthase had shown that, following partial proteolytic digestion, functional domains representative of component synthase activities could be isolated [24–28]. Indeed, a functional ACP domain of the rat fatty acid synthase had previously been isolated in this manner (S Smith and VS Rangan, personal communication). By identifying the sequence limits of a peptidyl carrier protein (PCP) domain of tyrocidine synthetase (TycA), Marahiel and coworkers have been able to overproduce a functional 112-aa peptide synthetase carrier protein [8] (Fig. 7). This protein undergoes partial phosphopantetheinylation in *E. coli*, and can then act as an aminoacyl acceptor when incubated with its corresponding adenylation/transferase domain. The PCP substrate is easily purified from endogenous *E. coli* ACP when expressed as a hexahistidine fusion (data not shown). An analogous strategy led to construction and isolation of a hexahistidine fusion of SrfB1, a 143 kDa fragment containing the amino-acid-activating and PCP domains involved in the activation of the fourth residue (valine) in surfactin biosynthesis (Fig. 7).

Table 1

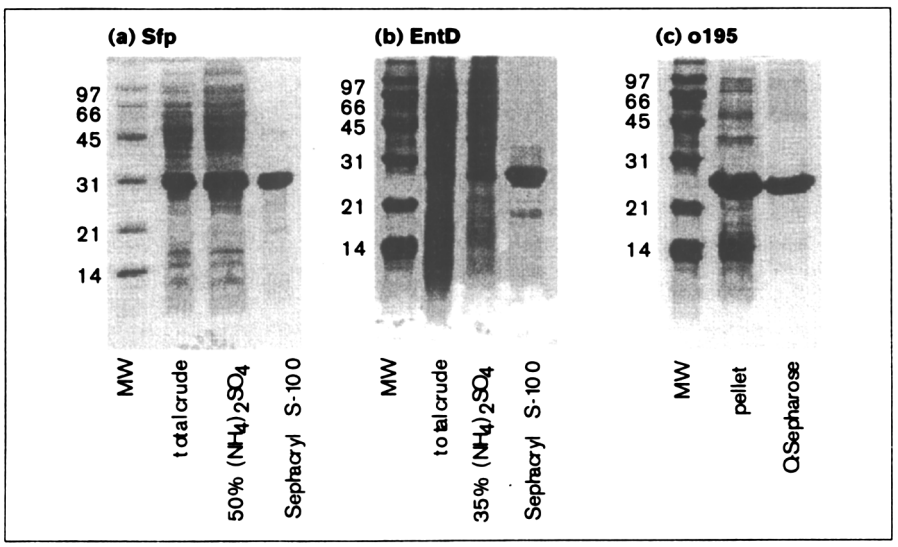
ACP synthase homologs.*

Pathway	Protein	Organism	Size
Enterobactin	EntD	*E. coli*	209 aa
		S. typhimurium	232 aa
		S. austin	232 aa
		S. flexneri	209 aa
Surfactin	Sfp	*B. subtilis*	224 aa
	Psf-1	*B. pumilus*	233 aa
Gramicidin S	Gsp	*B. brevis*	237 aa
Bacitracin	Bli	*B. licheniformis*	225 aa
Iturin A	Lpa-14	*B. subtilis*	224 aa
Nosiheptide	NshC	*S. actuosus*	253 aa
Lysine	LYS5	*S. cerevisiae*	272 aa
Fatty acids	ACPS	*E. coli*	126 aa
	HI0152	*H. influenzae*	235 aa
	FAS2	*S. cerevisiae*	1894 aa
		C. albicans	1885 aa
		P. patulum	1857 aa
		S. pombe	1842 aa
		A. nidulans	1559 aa
Differentiation	HetI	*Anabaena sp.*	237 aa
		Synechocystis sp.	246 aa
Unknown	o195	*E. coli*	195 aa
	1314154	*S. pombe*	258 aa
	CELTO4G9	*C. elegans*	297 aa

*All sequences except NshC (W Strohl, personal communication, GenBank Accession Number U75434, submitted) and Bli (M Marahiel, unpublished) are available in the GenBank, SwissProt, or EMBL databases.

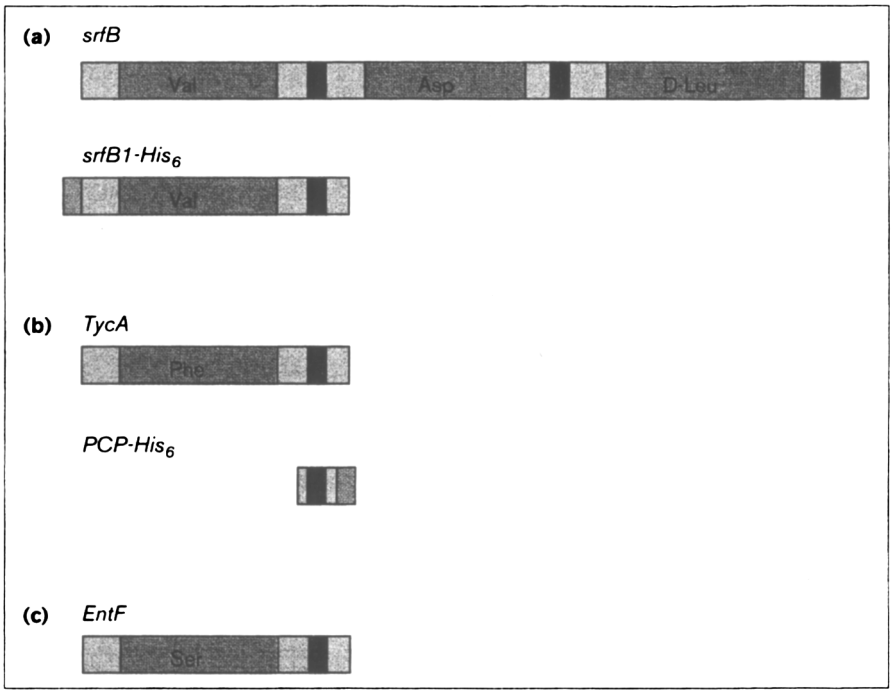
As mentioned above, recombinant PCP undergoes partial phosphopantetheinylation when expressed in *E. coli* [8]. When recombinant PCP was incubated with purified ACPS and [^{3}H]-(pantetheinyl)-CoASH *in vitro*, however, no incorporation of ^{3}H label was observed (Fig. 8). This result agreed with our earlier finding that ACPS cannot catalyze the modification of EntF, another type I peptide synthetase component. We therefore hypothesized that another *E. coli* P-pant transferase activity, probably EntD given its sequence similarity to ACPS, is specific for the phosphopantetheinylation of EntF or recombinant PCP overproduced in *E. coli*. To test this idea, we incubated each of the four pure proteins ACPS, EntD, o195, and Sfp with apo-ACP and apo-PCP in the presence of [^{3}H]CoASH. Each of the four candidate P-pant transferases generated tritiated ACP and/or PCP in TCA precipitation assays (data not shown). To verify that the ^{3}H label that coprecipitated with ACP and PCP represented covalent attachment of P-pant, the tritiated products were subjected to SDS electrophoresis and autoradiography (Fig. 8). It is clear that both ACPS and Sfp show robust phosphopantetheinylation activity (Fig. 8a). When apo-ACP is the substrate, EntD is weakly active compared to ACPS and Sfp and o195 is even less active, but both EntD and o195 give signals well above the background, showing that EntD and o195 do have P-pant transferase activity. When the 13 kDa apo-PCP was used as substrate for these four P-pant transferases in Figure 8b, Sfp and EntD are now highly active, but o195 and ACPS give no detectable modification at the single timepoint. When the much larger substrates apo-EntF and apo-SrfB1 fragment (140 kD) are used (Fig. 8c), the cognate enzymes, EntD for EntF and Sfp for SrfB1, are obviously competent for posttranslational phosphopantetheinylation. Mass spectrometry was used to confirm that the tritium incorporated into the apo-proteins represented transfer of the intact phosphopantetheinyl group. We previously validated this approach using ACPS as catalyst and holo-ACP as product [2] and used it here to examine PCP modification. Mass spectrometric analysis (MALDI-TOF) of unlabeled enzymatic holo-PCP indicated a molecular

Figure 6



Overproduction of candidate P-pant transferases. **(a)** Purification of *Bacillus subtilis* Sfp heterologously expressed in *Escherichia coli*. **(b)** Overproduction and purification of *E. coli* EntD. **(c)** Overproduction and purification of *E. coli* o195. All gels shown are SDS-PAGE (15 % acrylamide, 2.6 % bisacrylamide).

Figure 7



P-pant acceptor domains and the His_6-tagged constructs used for purification. Schematic diagram showing the comparative alignment of **(a)** SrfB and the SrfB1-His_6 fragment, **(b)** Tyc A and its constituent PCP domain tagged with His_6 and **(c)** EntF. Amino-acid-activating domains are shown in light purple. Phosphopantetheine attachment sites are shown in dark purple.

weight of 13 431 (calculated 13 459) in contrast to an observed molecular weight of 13 130 (calculated 13 120) for the apo-PCP substrate. These are the first data that establish that EntD, Sfp, and o195 are enzymes and that they catalyze the transfer of P-pant to the serine sidechain of an acyl carrier protein.

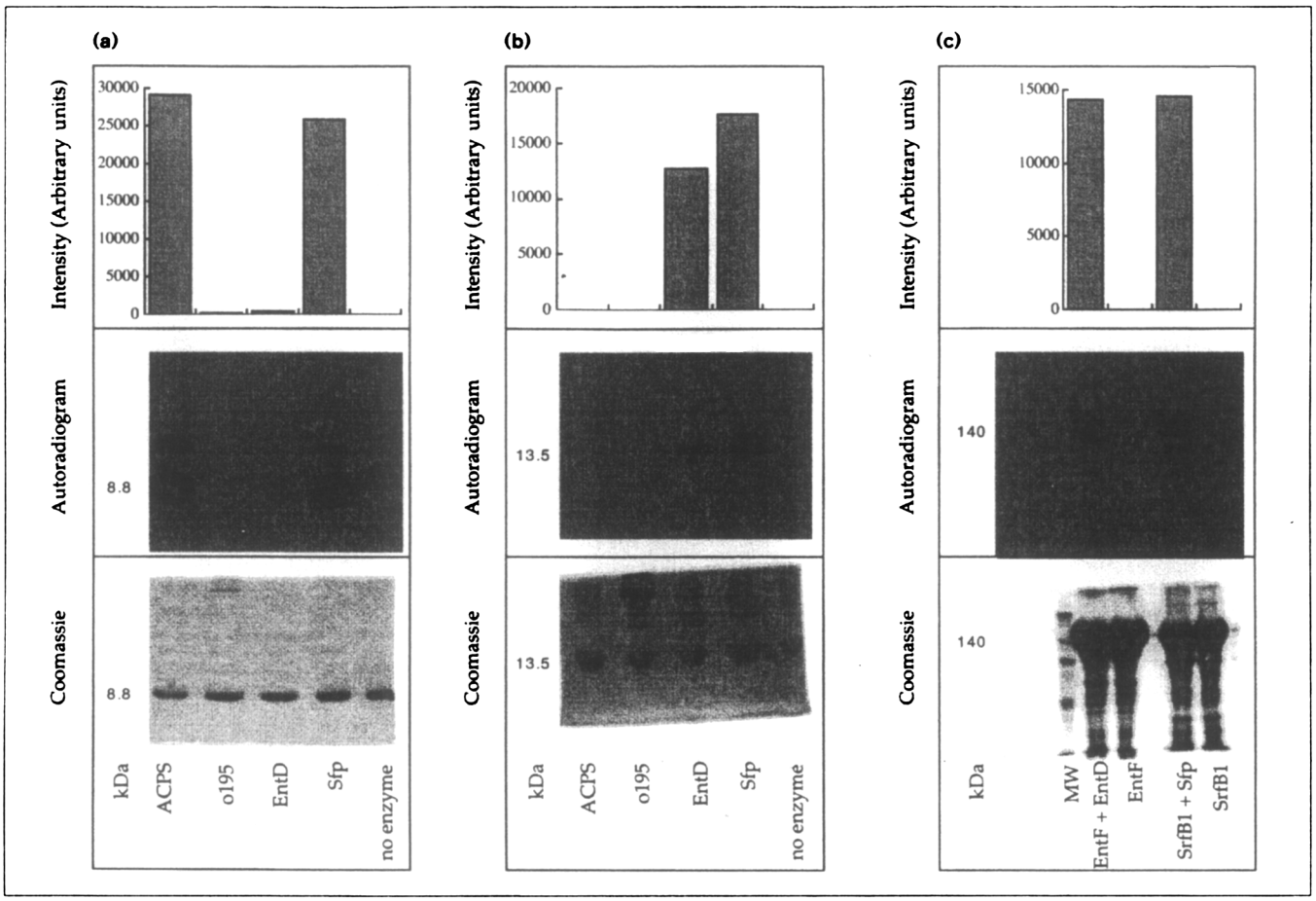
Specificity of ACPS, EntD and o195

Having demonstrated that EntD does in fact have P-pant transferase activity, we sought kinetic confirmation that it is indeed the enzyme responsible for the posttranslational modification of EntF. As described above, autoradiography of SDS gels confirmed incorporation of radiolabeled phosphopantetheine into EntF catalyzed by EntD (Fig. 8c). Furthermore, a time course of EntD-catalyzed incorporation of radiolabel into EntF provides *in vitro* evidence of at least two partner-specific P-pant transfer reactions occurring within *E. coli*. ACPS specifically catalyzes the transfer of P-pant to apo-ACP, while EntD is the transferase for its partner EntF. EntF is modified effectively by EntD (100 nM), whereas EntF undergoes almost no modification in the presence of 15-fold higher concentrations of ACPS and o195, clearly demonstrating the specificity of EntD for EntF (Fig. 9a). In contrast, apo-ACP is almost exclusively modified by ACPS (Fig. 9b), confirming that in *E.coli* ACPS is the P-pant transferase that activates the type II fatty acid synthase and EntD is the P-pant transferase that activates the type I enterobactin synthetase. The autoradiogram in Figure 8a shows, however, that both o195 and EntD can modify apo-ACP; the rate of modification is very low, yet is significantly higher than the background rate in the absence of enzyme (Fig. 8a, lane 5). This is presumably due to non-specific enzyme-catalyzed phosphopantetheinylation of the conserved serine residue. Assuming that the inclusion-bound o195 has been properly refolded and that an additional glycine introduced after the methionine start during PCR cloning has no significant effect on activity, it would appear that o195 is specific for a third, as yet unknown, substrate in *E. coli*; presumably P-pant transfer to this unknown protein would require o195 and would not be efficiently catalyzed by ACPS or EntD.

Specificity of Sfp toward apo-SrfB1, apo-PCP and apo-ACP

Sfp appears to be non-specific, efficiently catalyzing the modification of the two *Bacillus* derived type I peptide synthetase domains, apo-PCP and apo-SrfB1, the *E. coli* type II fatty acid synthase apo-ACP subunit (Fig. 8) and EntF (data not shown). Based on this evidence, Sfp would appear not to discriminate between type I peptide synthetase domains and type II fatty acid synthase subunits suggesting that there may be crosstalk between Sfp and fatty acid synthase, at least when expressed in *E. coli*. Careful kinetic analysis to determine whether Sfp selectively modifies SrfABC and not the *B. subtilis* fatty acid synthase ACP subunit must await overproduction of the *B. subtilis* ACP, however. Morbidino and co-workers [29] have been able to sequence the entire *B. subtilis* ACP protein by Edman degradation, but the intact *acpP* gene appears to be toxic to *E. coli* and has proven difficult to clone.

Figure 8



P-pant transferase reactions. Coomassie-stained gels are shown for each P-pant transferase incubation with the corresponding autoradiograms and integrated band intensities for individual P-pant transferase incubations. **(a)** Incubations of ACPS (1.8 μM), o195 (2.2 μM), EntD (1.3 μM), Sfp (1.6 μM) or no enzyme with apo-ACP (150 μM) as substrate. **(b)** Incubations of ACPS (1.8 μM), o195 (2.2 μM), EntD (1.3 μM), Sfp (1.6 μM) or no enzyme with apo-PCP (45 μM) as substrate. **(c)** Incubations of EntD (1.3 μM) and Sfp (1.6 μM) with their homologous substrates apo-EntF and apo-SrfB1.

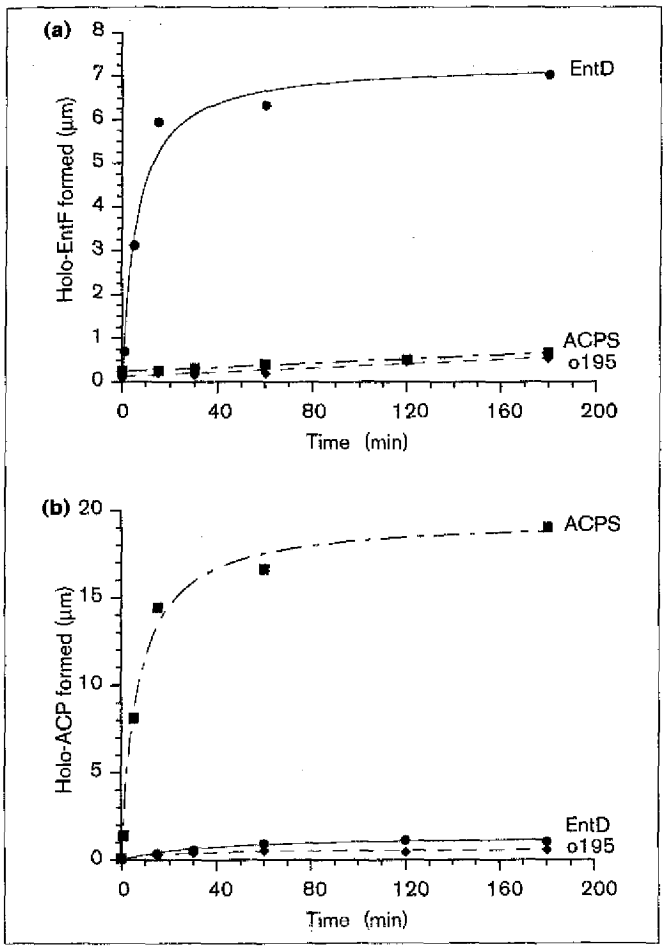
Holo-SrfB1 can activate L-valine

The action of Sfp on the 143 kDa SrfB1 fragment in conversion of the apo-form to the holo-form (Fig. 1) should generate a phosphopantetheinylated SrfB1 competent to undergo specific recognition and acylation by the amino acid L-valine, residue 4 in surfactin (Figs 4,7). Apo-SrfB1 undergoes very little acylation when incubated with [^{14}C]-L-valine, showing that the contamination of this preparation by holo-SrfB1 is small. After incubation with Sfp, however, the level of [^{14}C]-L-valine-holo-SrfB1 covalent complex formed in the complete incubation mixture increases about 14-fold, consistent with an increase in the amount of holo-SrfB1 present. The [^{14}C]-L-valine is used by the amino-acid-activating domain of holo-SrfB1 to make valyl-AMP which then undergoes intramolecular acyl-transfer to the SH group of the P-pant moiety in the adjacent PCP domain. Holo-SrfB1 cannot be covalently acylated by the non-cognate L-aspartate residue, the fifth amino acid to be activated by SrfABC, as expected given the absence of an aspartate-specific adenylation domain on SrfB1. Thus the holo-SrfB1 formed following incubation with Sfp and CoASH has both an active adenylation domain and a functional holo-peptidyl carrier protein domain, and should therefore be a useful reagent to probe peptide-bond-forming steps between adjacent sites of multienzyme, multiple thiotemplate synthases.

Discussion

The transfer of 4′-phosphopantetheine from CoASH to conserved serine residues in the signature sequences of acyl carrier protein domains (type I) or subunits (type II) is essential for the functional activation of all fatty acid synthases, polyketide synthases and non-ribosomal peptide synthetase complexes. This posttranslational phosphopantetheinylation introduces a covalently-attached

Figure 9



Time courses of P-pant transferase activity. **(a)** Time course of EntD (100 nM), ACPS (1.8 μM), or o195 (1.5 μM) incubated with apo-EntF (20 μM) as measured by radioassay. **(b)** Time course of EntD (1.6 μM), ACPS (100 nM), or o195 (1.5 μM) incubated with apo-ACP (50 μM).

nucleophilic thiol on a long tether that becomes the site of all the initiation and acyl transfer events involved in the assembly of the broad array of natural products synthesized by these enzymes. Thus, identification of the P-pant loading enzymes that create the active holo-ACP forms by posttranslational modification is important to the understanding of both the molecular mechanism of holo-ACP formation and the specificity of serine phosphopantetheinylation. These findings will aid in the design of strategies for heterologous production of functional polyketide and polypeptide synthetases (e.g. in combinatorial biosynthesis of 'unnatural' natural products), and studies aimed at the synthesis of inhibitors of specific P-pant loading reactions (e.g. in fungal lysine biosynthesis, see below).

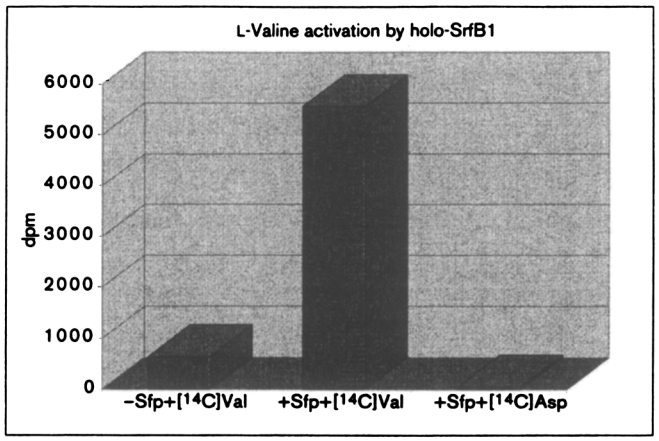
Our recent purification, characterization, and identification of the *E. coli* holo-ACPS [2] provided the first molecular information on this class of posttranslational-modifying enzymes. Somewhat to our surprise, initial database searches with the *E. coli* ACPS sequence revealed no obvious homologs in the protein databases. We eventually detected marginal similarities of 15–22 % over 120 residues in the carboxy-terminal region of three fungal fatty acid synthases (Fig. 5), indicating that the phosphopantetheinylating activity may have been integrated as a domain in these polyenzymes. For example the carboxy-terminal 121 aa of the 1894-aa yeast fatty acid synthase subunit II (yFASII) might act intramolecularly to add a P-pant unit to Ser180 on the putative ACP domain of this polyprotein. We have not yet obtained active fragments of yFASII that catalyze these reactions in *trans*, but Schweizer's group [13–18] has previously reported that two mutated fatty acid synthases, one in which the mutation is at Ser180 and the other at Gly1777, which are inactive alone, can complement each other *in vivo* and *in vitro*, consistent with this proposal.

EntD, Sfp and Gsp as specific P-pant transferases

Starting with *E.coli* ACPS, we detected three bacterial proteins EntD, Sfp, and Gsp which have previously been identified by complementation as orthologous genes [19,20]. The specific functions of *sfp*, *gsp* and *entD* have until now been obscure. The studies described here establish that Sfp has phosphopantetheinyl transferase activity and clearly assigns a catalytic loading function to Sfp. It posttranslationally modifies the conserved serine in the first subsite of SrfB, which is responsible for valine activation. We expect that Sfp will be able to modify the consensus serine residue in all seven amino-acid-activating sites in SrfABC (Fig. 4) and by extension that Gsp will catalyze P-pant transfer to the five amino-acid activating sites in GrsA and GrsB, allowing the sequential activation and polymerization of amino acids as required for the thiotemplate mechanism for non-ribosomal peptide bond assembly [30]. The *bli* and *lpa*-14 gene products most probably have an equivalent role, that is iterative P-pantetheinylation of each amino acid-activating domain in *B. licheniformis* bacitracin synthetase [31] and *B. subtilis* iturin A synthetase respectively [32]. While *in vitro* enzymatic specificity remains to be fully explored, the *in vivo* genetic studies [11,32] argue strongly for specific partner protein recognition by a distinct P-pant transferase. This may well be a general theme in non-ribosomal peptide antibiotic biosynthesis. While Sfp, Gsp and EntD are required for peptide and depsipeptide biosynthesis, these proteins are not essential for survival [10,33]. We predict, however, that there will be other as yet unidentified P-pant transferases in the *Bacillus* organisms specific for the ACP subunits of their respective fatty acid synthases which, like *E. coli* ACPS, will be essential for viability.

A third example of a partner protein-specific phosphopantetheinyl transferase is EntD, one of the proteins

Figure 10



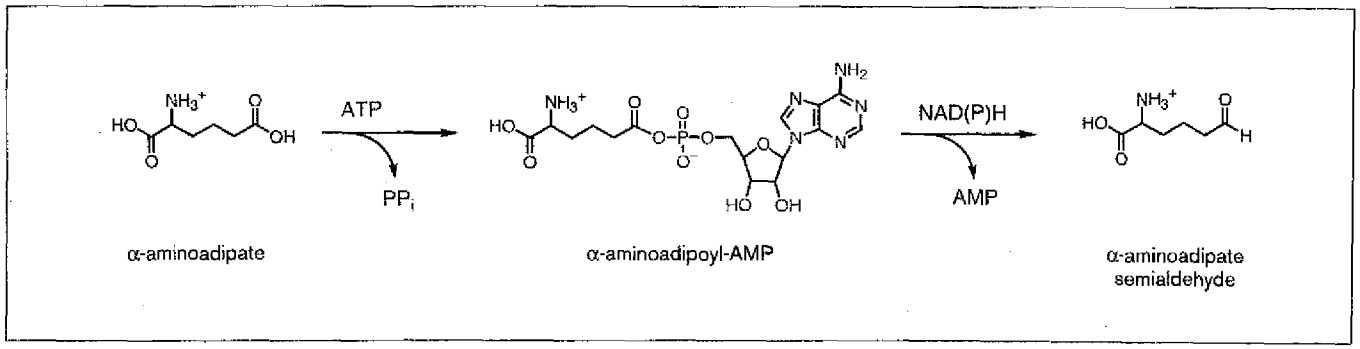
[^{14}C]Valine activation by holo-SrfB1. In the first column, SrfB1 (2 μM) was preincubated with CoA (200 μM) in the absence of Sfp before subsequent incubation with [^{14}C]-L-Valine (100 μM, 42.4 Ci mol^{-1}) and ATP (2 mM). In the second column, SrfB1 was preincubated with CoA (200 μM) in the presence of Sfp (1.3 μM) before subsequent incubation with [^{14}C]-L-Valine (100 μM, 42.4 Ci mol^{-1}) and ATP (2 mM). In the third column, SrfB1 (2 μM) was preincubated with CoA (200 μM) in the presence of Sfp (1.3 μM) before subsequent incubation with [^{14}C]-L-Aspartate (100 μM, 40.3 Ci mol^{-1}) and ATP (2 mM).

required for production and secretion of the iron-scavenging dihydroxybenzoyl-serine trilactone enterobactin in *E. coli*. We had previously cloned, sequenced, and purified EntF, a 140 kDa component of the enterobactin synthetase, and demonstrated that it activates L-serine and contains phosphopantetheine [6,7]. As EntD is required for enterobactin biosynthesis *in vivo* [10] and shows high activity for *in vitro* P-pantetheinylation of pure apo-EntF, it is now clear that EntD is defined as the specific P-pant transferase that makes active holo-EntF from apo-EntF *in vivo*. Pure ACPS from *E. coli* will not significantly posttranslationally modify EntF, consistent with the hypothesis that protein–protein recognition controls the specificity of phosphopantetheinylation *in vivo*. We predict that incubations of EntD and the enterobactin synthetase components with CoASH, L-serine and dihidroxybenzoate should reconstitute *in vitro* enterobactin production. At 140 kDa, EntF is the appropriate size for an amino-acid-activating module in a multidomain polypeptide synthetase [34]. It can be efficiently modified *in vitro* by EntD, showing that P-pant addition can occur after translation of the apo-protein, and not only co-translationally prior to folding of the apo-protein into its native structure. The NMR structure of *E. coli* apo-ACP shows that the nucleophilic Ser36 is in an accessible β-turn [35]; this may be a common architectural scaffolding for ACP domains in polyketide and polypeptide synthases and may be important in recognition by P-pant transferases.

Other P-pantetheinyl transferases

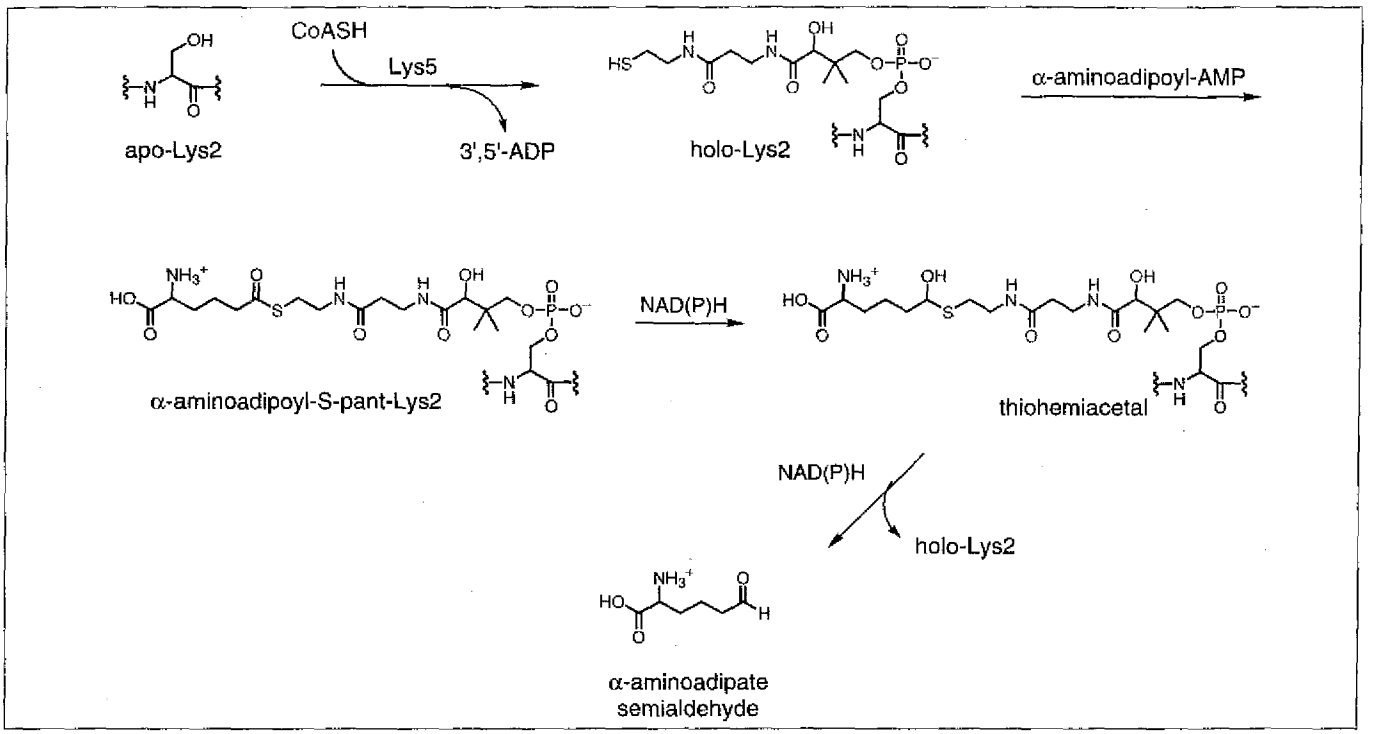
Using the EntD/Sfp/Gsp family as a base for further database searches has led to the identification of several additional candidates that are probably P-pant transferase family members (Table 1). Of these, in addition to ACPS and EntD, we have subcloned, expressed and characterized o195 as a third *E. coli* protein with P-pant transferase activity. The activity of o195 towards apo-ACP and apo-EntF is low, suggesting that o195 specifically catalyzes efficient P-pant transfer to an as yet unidentified substrate. A hypothetical protein, HI0152, in *Haemophilis influenzae* has been identified as a putative P-pant transferase. This resolves the apparent problem that no P-pant transferase in the *Haemophilis* genome had previously been found using ACPS-based searches. HI0152 is positioned directly upstream of the *H. influenzae* fatty acid synthase gene cluster, consistent with the notion that its protein product might be involved in fatty acid biogenesis. There is also some evidence that two additional proteins in cyanobacteria have similar functions (Table 1). In Anabaena, the genes *HetI*, *HetM*, and *HetN* have been implicated in the production of an unidentified secondary metabolite that inhibits heterocyst differentiation (a

Figure 11



Scheme showing the reaction previously proposed to be catalyzed by the Lys2–Lys5 complex. α-Aminoadipate is first activated to α-aminoadipoyl-AMP. This acyl-adenylate would then undergo direct reduction in a NAD(P)H dependent reaction to yield α-aminoadipate semialdehyde.

Figure 12



Scheme showing the reaction we now propose to be catalyzed by Lys2. Following phosphopantetheinylation of Lys2 catalyzed by Lys5, aminoadipate is transferred from aminoadipoyl-AMP to yield α-aminoadipoyl-S-pant-Lys2. This thioester then undergoes direct reduction in a NAD(P)H dependent reaction to yield a thiohemiacetal intermediate which then decomposes to the α-aminoadipate semialdehyde.

process occurring under low fixed nitrogen conditions in which a subset of cyanobacterial cells differentiate into the specialized heterocysts which have the ability to fix nitrogen) [36]. Sequence analysis suggests HetN is a NAD(P)H-dependent oxidoreductase like those involved in the biosynthesis of polyketides and fatty acids, while HetM has an ACP domain. HetI shows similarity to Sfp/Gsp/EntD, and is thus likely to be the HetM-specific phosphopantetheinyl transferase in the synthesis of the hypothesized secondary metabolite.

A final example is the 272-aa Lys5 protein involved in the yeast lysine biosynthetic pathway. Yeast and other fungi synthesize lysine via the unique α-aminoadipate pathway, an eight-step pathway beginning with homocitrate and proceeding via α-aminoadipate to saccharopine to lysine [37]. Complementation analysis suggests that Lys2 and Lys5 are involved in the same step in this pathway, the reduction of α-aminoadipate to aminoadipate semialdehyde [38]. Labeled pyrophosphate exchange experiments indicate that this reaction appears to proceed through an α-aminoadipoyl-AMP intermediate [39,40]. Recent sequence analysis [41] shows Lys2 to be a 155 kDa protein with homology to amino-acid-activating peptide synthetases including TycA, GrsAB, and SrfA. Like these peptide synthetases, Lys2 is believed to cleave ATP to AMP and PPi, activating α-aminoadipate to the α-aminoadipoyl-AMP which is then reduced by NADPH to the aldehyde (Fig. 11). A search for a consensus P-pant attachment site in Lys2 reveals the signature motif LGGHS around Ser880. We therefore propose, in contrast to previous suggestions, that Lys2 and Lys5 may form a two-subunit enzyme [38], that the 272-aa Lys5 is a specific phosphopantetheinyl transferase for Ser880 in Lys2. The thiol of the newly-introduced P-pant prosthetic group on Lys2 would attack the aminoadipoyl-AMP to give aminoadipoyl-S-pant-Lys2, in a similar manner to the sequential formation of aminoacyl-AMP to aminoacyl-S-pant-TycA in the homologous tyrocydine synthetase A subunit (Fig. 12). At this point, hydride addition to the acyl-S-pant-Lys2 would yield a thiohemiacetal which would readily decompose to aldehyde product and HS-pant-Lys2. This sequence has precedent in the reverse direction in the oxidation of glyceraldehyde-3-P to the acyl-S-enzyme in GAP dehydrogenase catalysis via a cysteinyl-S-enzyme hemithioacetal [42].

Significance

We have obtained evidence for a family of more than a dozen proteins with catalytic posttranslational modification activity. We anticipate that all these proteins will prove to be phosphopantetheinyl transferases with

CoASH as a common substrate but will show specificity, directed by protein–protein interactions, for the conserved serine motif in particular partner proteins. It is likely that most, if not all, of the multienzyme peptide synthetases that use the multiple thiotemplate scaffolding strategy to make peptide antibiotics nonribosomally [30] will have a partner-protein-specific posttranslational modifying enzyme that covalently adds the swinging arm thiol group required to enable acyl transfers. The new proteins in this family are 50–150 amino acid residues longer than the first one discovered, the 125-aa *E. coli* ACPS subunit; these extra amino acids may be responsible for specificity of partner-protein binding. It remains to be seen whether the many polyketide synthase complexes will use this strategy for posttranslational modification.

Materials and methods

Overproduction, purification and characterization of EntD, Sfp, and o195

B. subtilis Sfp was overproduced and purified from *E. coli* strain MV1190/pUC8-sfp as previously described by Nakano *et al.* [11] (Fig. 6). EntD was PCR-amplified from wild-type *E. coli* K-12 by colony PCR using the forward primer 5′-ATTATATCCATGGgtTCcTCcGTtTCcAAcATGGTCGATATGAAAACTACGCA-3′ and the reverse primer 5′-TGATGTCAAGCTTATTAATCGTGTTGGCACAGCGTTAT-3′ (IDT). The forward primer introduced an *Nco*I restriction site (underlined) which allowed mutation of the TTG start to an ATG start and inserted a Gly codon (GGT) after the Met initiator. In addition the forward primer optimized codon usage for the first six codons of the *entD* gene (modified bases shown in lower case). The reverse primer incorporated a *Hin*dIII restriction site (underlined). The *Nco*I/*Hin*dIII digested PCR product was cloned into pET28b (Novagen) and transformed into competent *E. coli* DH5α. The recombinant *entD* sequence was confirmed by DNA sequencing (Dana-Farber Molecular Biology Core Facility, Boston, MA). Competent cells of the overproducer strain *E. coli* BL21(DE3) were then transformed with the supercoiled pET28b-*entD*. Induction of a 2-l culture of BL21(DE3)pET28b-*entD* with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) followed by growth at 25°C for 5 h yielded predominantly inclusion-bound EntD, although a modest amount of the overproduced protein was soluble. The overproduction of soluble EntD may be complicated by the fact that the wild type Ent proteins are synthesized in detectable quantities only under iron-starved conditions. Furthermore, although the recombinant EntD is functional as a soluble protein, the wild type EntD has been reported to be membrane bound [43]. The induced cell paste was resuspended in 50 mM Tris, 1 mM EDTA, 5 % glycerol, pH 8.0 (40 ml) and lysed by two passages through the French press at 15 000 psi. Cellular debris and inclusion bound protein was removed by centrifugation at 8000 × g for 30 min. Pulverized ammonium sulfate was added to 35 %, 65 % and 80 % saturation. The 35 % fraction containing the largest fraction of EntD was applied to a 2.5 × 115 cm Sephacryl S-100 column. The column was eluted at a flow rate of 1 ml min^{-1} using the same buffer as above, collecting 8 ml fractions to obtain homogeneous protein.

Similarly, o195 was PCR-amplified from wild-type *E. coli* K-12 by colony PCR using the forward primer 5′-ATTATATCCATGGgtTAcCGGATAGTTCTGGGGAAAGTT-3′ and the reverse primer 5′-TGATGTCAAGCTTATCAGTTAACTGAATCGATCCATTG-3′(IDT). The forward primer with its *Nco*I restriction site (underlined) gave insertion of a Gly codon (GGT) after the Met initiator codon of the o195 sequence; codon usage for the succeeding codon was also optimized (base change shown in lower case). The reverse primer incorporated a *Hin*dIII restriction site (underlined). The *Nco*I/*Hin*dIII-digested PCR product was cloned into pET28b (Novagen) and transformed into competent *E. coli* DH5α. The recombinant o195 sequence was confirmed by DNA sequencing (Dana-Farber Molecular Biology Core Facility, Boston, MA). Competent cells of the overproducer strain *E. coli* BL21(DE3) were then transformed with the supercoiled pET28b-o195. Induction of a 2-l culture (2 × YT media) of BL21(DE3)pET28-o195 with 1 mM IPTG followed by growth at 37° C for 3.5 h yielded predominantly inclusion-bound o195 protein. The cell paste was resuspended in 50 mM Tris·HCl, 1 mM EDTA, 5 % glycerol, pH 8.0 (40 ml) and lysed by two passages through a French pressure cell at 15 000 psi. Cellular debris and inclusion-bound protein was pelleted by centrifugation at 27 000 × g for 30 min. The inclusion-bound protein pellet was resuspended in 30 ml of 50 mM Tris·HCl, pH 8.0, 1 mM EDTA, and 5 % glycerol and incubated for 30 min at room temperature with 10 mg lysozyme and 30 mg deoxycholate. The pellet was reobtained by centrifugation for 15 min at 27 000 × g and solubilized in 30 ml of 8 M urea, 50 mM Tris·HCl, pH 8.0, 10 mM dithiothreitol (DTT). Residual solid material was removed by centrifugation for 15 min at 27 000 × g. The urea-solubilized solution (30 ml) was then applied to a 2.5 × 10 cm Q-Sepharose column equilibrated with 8 M urea, 50 mM Tris·HCl, pH 8.0. The column was washed with 50 ml of the equilibration buffer and then a gradient of 250 ml 0–0.25 M NaCl in 8 M urea, 50 mM Tris·HCl pH 8.0 followed by 200 ml of 0.25–1 M NaCl in the same buffer was applied. The o195 protein eluted at ~200 mM NaCl as determined by 15 % SDS·PAGE. The purified o195 was renatured by diluting a portion of it 10-foid in 8 M urea, 50 mM Tris·HCl, pH 8.0, 10 mM DTT and dialyzing overnight at 4° C against 10 mM Tris·HCl, pH 8.0, 1 mM DTT. Two liters of culture grown in 2 × YT media yielded 3.1 g of cells from which ~80 mg of o195 protein was obtained.

Production of apo-protein substrates, apo-ACP, apo-PCP, apo-EntF, and apo-SrfB

The *E. coli* fatty acid synthase ACP was overproduced and purified in its apo-form from *E. coli* strain DK554 [21] following the procedure of Rock and Cronan [44] with the exception that following cell disruption and centrifugation (30 min at 28 000 × g), the crude extract containing 10 mM $MgCl_2$ and 10 mM $MnCl_2$ was incubated for 60 min at room temperature. In this manner, minor amounts of holo-ACP were hydrolyzed to the apo-form using the endogenous *E. coli* ACP phosphodiesterase [45]. The PCP domain of TycA was overproduced with a hexahistidine tag using *E. coli* strain SG13009(pREP4)/ pQE60-PCP [8]. Following lysis of the induced culture the His_6-tagged protein was purified by nickel-chelate chromatography. *E. coli* apo-EntF was purified as previously described [7].

Apo-SrfB1 was cloned from plasmid p120-21E [46]. Briefly, p120-21E was digested with EcoRV to release a 3648-base-pair fragment encoding the SrfB1, valine-activating domain of surfactin synthetase. This fragment was inserted into *Stu*I-cleaved pPROEX-1 (Gibco/BRL Life Sciences Technologies) to give plasmid pML118 which codes for a amino-terminal His_6-tagged SrfB1 domain (142.7 kDa). His_6-SrfB1 was overproduced using *E. coli* strain AG1574 (courtesy A. Grossman) [47]. Cells were grown at 25° C in 2 × YT media (2 l) to an O.D. of 0.4 at which point they were induced with 1 mM IPTG and allowed to grow for an additional 4 h. Cells were harvested by centrifugation (3 g), resuspended in 35 ml of 5 mM imidazole, 500 mM NaCl, 20 mM Tris·HCl, pH 7.9 and lysed by two passages through a French pressure cell. This crude extract was clarified by centrifugation for 30 min at 27 000 × g. More than 50 % of the overproduced SrfB1 was obtained in the soluble fraction as determined by 6 % SDS-PAGE. His_6-tagged SrfB1 was purified on His-Bind resin (Novagen) following the manufacturer's recommendations.

Assay for apo-protein to holo-protein conversion by ^{3}H-P-pant group transfer from ^{3}H-coenzyme A

P-pant transferase activity (Fig. 1) was measured by radioassay. Enzyme preparations (final enzyme concentrations of 0.1–2.2 μM) were incubated with 75 mM Tris·HCl, pH 8.8, 10 mM $MgCl_2$, 25 mM DTT, 200 μM [^{3}H]-(pantetheinyl)-CoASH (5.3 × 10^6 dpm total activity)

and substrate (apo-ACP, apo-PCP, apo-EntF or apo-SrfB1, at final concentrations of 10–150 μM) for various times at 37° C in a final volume of 100 μl. The incubations were quenched with 10 % TCA and 500 μg bovine serum albumin (BSA) was added as a carrier. The protein was precipitated by centrifugation, washed 3 times with 10 % TCA, and the protein pellet solubilized with 150 μl 1 M Tris base. The resuspended protein was added to 3 ml liquid scintillation cocktail and the amount of [^{3}H]-phosphopantetheine incorporated into the substrate protein was quantified by liquid scintillation counting. Assays for autoradiography were performed as described above except 20 μM [^{3}H]-(pantetheinyl)-CoASH (2.6×10^6 dpm total activity) was used in the assay, no BSA was added to the TCA precipitate, and pellets were solubilized in SDS or native PAGE sample buffer titrated with 1 M Tris base. Assays using apo-PCP as substrate were resolved by 15 % SDS-PAGE, assays using *E. coli* ACP were resolved by 20 % native PAGE, and assays using SrfB1 or EntF were resolved on 8 % SDS-PAGE. Gels were Coomassie-stained, soaked for 30 min in Amplify (Amersham), dried at 80° C under vacuum and exposed to X-ray film for 24–150 h at –70°C (Fig. 8). The autoradiograms were scanned using a digital scanner and relative intensities of the radiolabeled bands were quantified using NIH Image 1.59 software (National Institutes of Health, USA).

Assay for activation of L-valine by holo-SrfB1

Apo-SrfB1 (2 μM) was incubated with 200 μM CoASH, 75 mM Tris·HCl pH 8.0, 10 mM $MgCl_2$, 25 mM DTT and 1.3 μM Sfp for 15 min at 37° C to generate holo-SrfB1. To the SrfB1-Sfp reaction mixture, ^{14}C-labeled amino acid (valine, 42.4 Ci mol^{-1}; aspartic acid, 40.3 Ci mol^{-1}) was added to 100 μM final concentration. ATP was added to a final concentration of 2 mM and the reaction (115 μL) was incubated for 15 min at 37° C, then stopped by the addition of 800 μL 10 % TCA with 15 μl of a 25 mg ml^{-1} BSA solution as carrier. The precipitate was collected by centrifugation, washed with 10 % TCA, dissolved in 150 μl Tris base, and then counted by liquid scintillation.

Acknowledgements

The authors thank Professor William Strohl (Ohio State University) for sharing the *S. actuosus* NshC protein sequence prior to publication. This work was supported by National Institutes of Health grants GM20011 (CTW), GM45898 (PZ) CA66736 (CK) and 5T32-GM08313-07 (RSF). CK was also supported by grant MCB-9417419 from the National Science Foundation. MAM was supported by the Deutsche Forschungsgemeinschaft and the European Commission. RHL was supported by National Institutes of Health Post-Doctoral Fellowship GM16583-03. AMG is a Howard Hughes Medical Institute Predoctoral Fellow.

References

1. Schlumbohm, W., *et al.*, & Wittmann-Liebold, B. (1991). An active serine is involved in covalent substrate amino acid binding at each reaction center of gramicidin S synthetase. *J. Biol. Chem.* **266**, 23135–23141.
2. Lambalot, R.H. & Walsh, C.T. (1995). Cloning, overproduction, and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. *J. Biol. Chem.* **270**, 24658–24661.
3. Takiff, H.E., Baker, T., Copeland, T., Chen, S.M. & Court, D.L. (1992). Locating essential *Escherichia coli* genes by using mini-Tn10 transposons: the *pdxJ* operon. *J. Bacteriol.* **174**, 1544–1553.
4. Debabov, D.V., Heaton, M.P., Zhang, Q., Stewart, K.D., Lambalot, R.H. & Neuhaus, F.C. (1996). The D-alanyl carrier protein in *Lactobacillus casei*: cloning, sequencing, and expression of *dltC*. *J. Bacteriol.* 178, 3869–3876.
5. Ritsema, T., Geiger, O., van Dillewijn, P., Lugtenberg, B.J.J. & Spaink, H.P. (1994). Serine residue 45 of nodulation protein NodF from *Rhizobium leguminosarum bv. viciae* is essential for its biological function. *J. Bacteriol.* **176**, 7740–7743.
6. Rusnak, F., Sakaitani, M., Drueckhammer, D., Reichart, J. & Walsh, C.T. (1991). Biosynthesis of the *Escherichia coli* siderophore enterobactin: sequence of the *entF* gene, expression and purification of EntF, and analysis of covalent phosphopantetheine. *Biochemistry* **30**, 2916–2927.
7. Reichert, J., Sakaitani, M. & Walsh, C.T. (1992). Characterization of EntF as a serine-activating enzyme. *Prot. Sci.* 1, 549–556.
8. Stachelhaus, T., Huser, A. & Marahiel, M. (1996). Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chemistry & Biology* **4**, 913–921.
9. Fleischmann, R.D., *et al.* & Venter, J.C. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae Rd*. *Science* **269**, 496–512.
10. Coderre, P.E. & Earhart, C.F. (1989). The *entD* gene of the *Escherichia coli* K12 enterobactin gene cluster. *J. Gen. Microbiol.* **135**, 3043–3055.
11. Nakano, M.M., Corbell, N., Besson, J. & Zuber, P. (1992). Isolation and characterization of *sfp*: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. *Mol. Gen. Genet.* **232**, 313–321.
12. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
13. Kuhn, L., Castorph, H. & Schweizer, E. (1972). Gene linkage and gene–enzyme relations in the fatty-acid-synthetase system of *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **24**, 492–497.
14. Schweizer, E., Kniep, B., Castorph, H. & Holzner, U. (1973). Pantetheine-free mutants of the yeast fatty-acid-synthetase complex. *Eur. J. Biochem.* **39**, 353–362.
15. Schweizer, E. (1977). Biosynthese und Struktur des Fettsaure-synthetase-Komplexes der Hefe. *Naturwissenschaften* **64**, 366–370.
16. Schweizer, E., *et al.*, & Zauner, J. (1987). Genetic control of fatty acid synthetase biosynthesis and structure in lower fungi. *Fat Sci. Tech.* **89**, 570–577.
17. Werkmeister, K., Wieland, F. & Schweizer, E. (1980). Coenzyme A: fatty acid synthetase apoenzyme 4'-phosphopantetheine transferase in yeast. *Biochem. Biophys. Res. Commun.* **96**, 483–490.
18. Schorr, R., Mittag, M., Muller, G. & Schweizer, E. (1994). Differential activities and intramolecular location of fatty acid synthase and 6-methylsalicylic acid synthase component enzymes. *Journal of Plant Physiology* 143, 407–415.
19. Borchert, S., Stachelhaus, T. & Marahiel, M.A. (1994). Induction of surfactin production in *Bacillus subtilis* by *gsp*, a gene located upstream of the gramicidin S operon in *Bacillus brevis*. *J. Bacteriol.* **176**, 2458–2462.
20. Grossman, T.H., Tuckman, M., Ellestad, S. & Osburne, M.S. (1993). Isolation and characterization of *Bacillus subtilis* genes involved in siderophore biosynthesis: relationship between *B. subtilis sfpo* and *Escherichia coli entD* genes. *J. Bacteriol.* **175**, 6203–6211.
21. Keating, D.H., Carey, M.R. & J. E. Cronan, J. (1995). The unmodified (apo) form of *Escherichia coli* acyl carrier protein is a potent inhibitor of cell growth. *J. Biol. Chem.* **270**, 22229–22235.
22. Crosby, J., Sherman, D.H., Bibb, M.J., Revill, W.P., Hopwood, D.J. & Simpson, T.J. (1995). Polyketide synthase acyl carrier proteins from *Streptomyces*: expression in *Escherichia coli*, purification and partial characterization. *Biochim. Biophys. Acta* **1251**, 32–42.
23. Hill, R.B., MacKenzie, K.R., Flanagan, J.M., J. E. Cronan, J. & Prestegard, J.H. (1995). Overexpression, purification, and characterization of *E. coli* acyl carrier protein and two mutant proteins. *Protein Expr. Purif.* **6**, 394.
24. Mattick, J.S., Tsukamoto, Y., Nickless, J. & Wakil, S.J. (1983). The architecture of the animal fatty acid synthetase I. Proteolytic dissection and peptide mapping. *J. Biol. Chem.* **258**, 15291–15299.
25. Mattick, J.S., Nickless, J., Mizugaki, M., Yang, C.Y., Uchiyama, S. & Wakil, S.J. (1983). The architecture of the animal fatty acid synthetase II. Separation of the core and thioesterase functions and determination of the N–C orientation of the subunit. *J. Biol. Chem.* **258**, 15300–15304.
26. Wong, H., Mattick, J.S. & Wakil, S.J. (1983). The architecture of the animal fatty acid synthetase III. Isolation and characterization of the β-ketoacyl reductase. *J. Biol. Chem.* **258**, 15305–15311.
27. Tsukamoto, Y., Wong, H., Mattick, J.S. & Wakil, S.J. (1983). The architecture of the animal fatty acid synthetase complex IV. Mapping of active centers and model for the mechanism of action. *J. Biol. Chem.* **258**, 15312–15322.
28. Tsukamoto, Y. & Wakil, S.J. (1988). Isolation and mapping of the β-hydroxyacyl dehydratase activity of chicken liver fatty acid synthetase. *J. Biol. Chem.* **263**, 16225–16229.
29. Morbidoni, H.R., De Mendoza, D. & Cronan, J.E., Jr. (1996). *Bacillus subtilis* acyl carrier protein is encoded in a cluster of lipid biosynthesis genes. *J. Bacteriol.* **178**, 4794–4800.
30. Lipmann, F. (1971). Attempts to map a process evolution of peptide biosynthesis. *Science* **173**, 875–884.
31. Gaidenko, T.A., Belitsky, B.R. & Haykinson, M.J. (1992). Characterization of a new pleiotropic regulatory gene from *Bacillus licheniformis*. *Biotechnologia*, 13–19.

32. Huang, C.-C., Ano, T. & Shoda, M. (1993). Nucleotide sequence and charateristics of a gene, *lpa-14*, responsible for the biosynthesis of the lipopeptide antibiotics iturin A and surfactin from *Bacillus subtilis* RB14. *J. Ferment. Bioeng.* **76**, 445–450.
33. Nakano, M.M., Marahiel, M.M. & Zuber, P. (1988). Identification of a genetic locus required for biosynthesis of the lipopeptide antibiotic surfactin in *Bacillus subtilis*. *J. Bacteriol.* **170**, 5662–5668.
34. Stachelhaus, T. & Marahiel, M.A. (1995). Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiol. Lett.* **125**, 3–14.
35. Holak, T.A., Nilges, M., Prestegard, J.H., Gronenborn, A.M. & Clore, G.M. (1988). Three-dimensional structure of acyl carrier protein in solution determined by nuclear magnetic resonance and the combined use of dynamical simulated annealing and distance geometry. *Eur. J. Biochem.* **175**, 9–15.
36. Black, T.A. & Wolk, C.P. (1994). Analysis of a Het$^-$ mutation in Anabaena sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing. *J. Bacteriol.* **176**, 2282–2292.
37. Bhattacharjee, J.K. (1985). α-Aminadipate pathway for the biosynthesis of lysine in lower eukaryotes. *CRC Crit. Rev. Microbiol.* **12**, 131–151.
38. Storts, D.R. & Bhattacharjee, J.K. (1989). Properties of revertants of *Lys2* and *Lys5* mutants as well as α-aminoadipate-semialdehyde dehydrogenase from *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* 161, 182–186.
39. Sagisaka, S. & Shimura, K. (1960). Mechanism of activation and reduction of α-aminoadipic acid by yeast enzyme. *Nature* **188**, 1189–1190.
40. Sinha, A.K. & Bhattacharjee, J.K. (1971). Lysine biosynthesis in *Saccharomyces*. Conversion of α-aminoadipate into α-aminoadipic δ-semialdehyde. *Biochem. J.* **125**, 743–749.
41. Morris, M.E. & Jinks-Robertson, S. (1991). Nucleotide sequence of the *LYS2* gene of *Saccharomyces cerevisiae*: homology to *Bacillus brevis* tyrocidine synthetase 1. *Gene* **98**, 141–145.
42. Walsh, C.T. (1979) Enzymatic Reaction Mechanisms. W.H. Freeman and Company, NY, USA.
43. Armstrong, S.K., Pettis, G.S., Forrester, L.J., & McIntosh, M. (1989). The *Escherichia coli* enterobactin biosynthesis gene *entD*: nucleotide sequence and membrane localization of its protein product. *Mol. Microbiol.* **3**, 757–766.
44. Rock, C.O. & Cronan, J.E., Jr. (1981). Acyl carrier protein from *Escherichia coli*. *Methods Enzymol.* **71**, 341–351.
45. Fischl, A.S. & Kennedy, E.P. (1990). Isolation and properties of acyl carrier protein phosphodiesterase of *Escherichia coli*. *J. Bacteriol.* **172**, 5445–5449.
46. Nakano, M.M., Magnuson, R., Myers, A., Curry, J., Grossman, A.D. & Zuber, P. (1991). *srfA* is an operon required for surfactin production, competence development, and efficient sporulation in *Bacillus subtilis*. *J. Bacteriol.* **173**, 1770–1778.
47. Frisby, D. & Zuber, P. (1991). Analysis of the upstream activating sequence and site of carbon and nitrogen source repression in the promoter of an early-induced sporulation gene of *Bacillus subtilis*. *J. Bacteriol.* **173**, 7557–7564.