Characterization of Sfp, a *Bacillus subtilis* Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases[†]

Luis E. N. Quadri,[‡] Paul H. Weinreb,[‡] Ming Lei,[‡] Michiko M. Nakano,[§] Peter Zuber,[§] and Christopher T. Walsh*[,‡]

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, and Department of Biochemistry and Molecular Biology, Louisiana State University Medical Center, Shreveport, Louisiana 71130

Received August 11, 1997; Revised Manuscript Received October 21, 1997

ABSTRACT: The *Bacillus subtilis* enzyme Sfp, required for production of the lipoheptapeptide antibiotic surfactin, posttranslationally phosphopantetheinylates a serine residue in each of the seven peptidyl carrier protein domains of the first three subunits (SrfABC) of surfactin synthetase to yield docking sites for amino acid loading and peptide bond formation. With recombinant Sfp and 16−17-kDa peptidyl carrier protein (PCP) domains excised from the SrfB1 and SrfB2 modules as apo substrates, k_{cat} values of 56−104 min^{-1} and K_m values of 1.3−1.8 μM were determined, indicating equivalent recognition of the adjacent PCP domains by Sfp. In contrast to other phosphopantetheinyl transferases (PPTases) previously examined, Sfp will modify the apo forms of heterologous recombinant proteins, including the PCP domain of *Saccharomyces cerevisiae* Lys2 (involved in lysine biosynthesis), the aryl carrier protein (ArCP) domain of *Escherichia coli* EntB (involved in enterobactin biosynthesis), and the *E. coli* acyl carrier protein (ACP) subunit, suggesting Sfp as a good candidate for heterologous coexpression with peptide and polyketide synthase genes to overproduce holo-synthase enzymes. Cosubstrate coenzyme A (CoA), the phosphopantetheinyl group donor, has a K_m of 0.7 μM. Desulfo-CoA and homocysteamine-CoA are also substrates of Sfp, and benzoyl-CoA and phenylacetyl-CoA are also utilized by Sfp, resulting in direct transfer of acyl phosphopantetheinyl moieties into the carrier protein substrate. Mutagenesis in Sfp of five residues conserved across the PPTase family was assessed for in vivo effects on surfactin production and in vitro effects on PPTase activity.

Phosphopantetheinyl transferases (PPTases)[1] play an essential role in priming polyketide synthases, nonribosomal peptide synthetases, and siderophore synthetases by covalently converting serine residues in acyl carrier protein (ACP), peptidyl carrier protein (PCP), or aryl carrier protein (ArCP) domains within those multidomainal enzymes from inactive apo-forms to active holo-forms (*1*). This occurs by tethering the phosphopantetheinyl moiety of cosubstrate coenzyme A (CoA) in phosphodiester linkage to the hydroxymethyl side chain of the conserved serine residue in the ACP, ArCP, and PCP domains (see eq 1) (*1*).

The posttranslational phosphopantetheinylation introduces a 20-Å long prosthetic group with a reactive thiol terminus on which the polyketide and polypeptide chain syntheses are initiated, and intermediates are elongated and matured as covalent acyl-, aryl-, or amino acyl-S-enzyme thioesters. Each ACP, ArCP, and PCP domain in the polyketide and polypeptide synthases is subjected to phosphopantetheinylation at a conserved serine residue embedded in the carrier protein consensus motif [Gx(D/H)S(L/I)] (*2*). Thus, the 6-deoxyerythronolide B synthase (DEBS) has seven phosphopantetheinyl (P-pant) groups posttranslationally introduced into its DEBS I, II, and III subunits (each with a molecular weight of >300 kDa) (*3*), while the cyclosporin synthetase has 11 amino acid activating modules, and so 11 P-pant prosthetic groups, posttranslationally introduced along its 1.7 mDa single polypeptide chain (*4*).

The specific PPTases responsible for posttranslationally priming these multienzyme thioltemplate synthetases have only recently begun to be identified, starting with the 14-kDa *Escherichia coli* holo-acyl carrier protein synthase (ACPS) (*5*), which converts the inactive apo-ACP subunit required for fatty acid biosynthesis to its active holo-form. ACPS will transfer the phosphopantetheinyl group from CoA to other apo-ACP substrates, including apo-NodF involved in rhizobial nodulation factor biogenesis (*6*), and the apo-ACP subunits that function, when phosphopantetheinylated, in the assembly of the aromatic polyketides tetracenomycin, frenolicin, granaticin, and oxytetracycline (*7*). In contrast, the *E. coli* ACPS will *not* recognize the apo-forms of several ArCP and PCP domains, including the apo-PCP domain of the *E. coli* protein EntF and the apo-ArCP domain of the *E. coli* protein EntB (*8*, *9*) that when phosphopantetheinylated

[†] This work was supported by National Institutes of Health Grants GM20011 (C.T.W.) and GM45898 (P.Z.). P.H.W. was supported by National Institutes of Health Postdoctoral Fellowship GM18721-01.

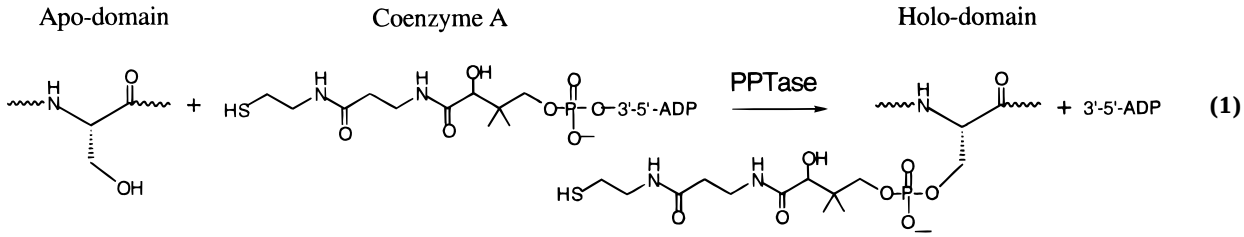
* Corresponding author e-mail, walsh@walsh.med.harvard.edu.

[‡] Harvard Medical School.

[§] Louisiana State University Medical Center.

[1] Abbreviations: ACP, acyl carrier protein; ACPS, holo-acyl carrier protein synthase; ArCP, aryl carrier protein; BSA, bovine serum albumin; CoA, coenzyme A; DEBS, 6-deoxyerythronolide B synthase; DTT, dithiothreitol; IPTG, isopropyl-1-thio-β-D-galactoside; MES, 2-(*N*-morpholino)ethanesulfonic acid; PAGE, polyacrylamide gel electrophoresis; PCP, peptidyl carrier protein; PCR, polymerase chain reaction; P-pant, 4′-phosphopantetheine; PPTase, 4′-phosphopantetheinyl transferase; SOE, splicing by overlap extension; TCA, trichloroacetic acid.

S0006-2960(97)01986-7 CCC: $15.00
Published on Web 01/22/1998



carry the seryl and dihydroxybenzoyl groups, respectively, assembled by *E. coli* into the iron chelating enterobactin.

The observed substrate specificity of ACPS led us to the expectation that a second PPTase must exist in *E. coli* with selectivity for apo-EntB and apo-EntF, and we have recently identified EntD as that enzyme (*8*, *9*). The 26-kDa EntD protein has marginal homology to ACPS as will be noted below, but was clearly related to such proteins as Sfp and Gsp from *B. subtilis* and *B. brevis*, respectively (*10*, *11*). The gene *sfp* had been identified genetically as required for production of the nonribosomally synthesized peptide antibiotic surfactin (*10*). The genes *sfp* and *gsp*, the latter being involved in synthesis of the peptide antibiotic gramicidin S and used successfully to produce a functional gramicidin synthetase module by coexpression in *E. coli* (*11*, *12*), mapped next to the surfactin synthetase and gramicidin synthetase genes, respectively. In addition, *sfp* could substitute for the *entD* gene in *E. coli entD*$^-$ mutants (*13*) and *gsp* for *sfp* in *B. subtilis sfp*$^-$ mutants (*11*), suggesting orthologous function of the three genes. We have previously reported the initial overproduction and purification of Sfp in *E. coli* and demonstrated its PPTase activity on a 143-kDa substrate containing the first module of the 401-kDa SrfB subunit of surfactin synthetase (*8*). Our findings suggest the reaction shown in Scheme 1 for Sfp; that is, conversion of apo-modules to holo-modules in surfactin synthetase, which are then competent for self-charging by the specific amino acids regioselectively incorporated into the heptapeptide antibiotic product surfactin by the thiotemplate mechanism (*14*, *15*). In the results reported here, the catalytic efficiency of Sfp is analyzed for modification of apo-PCP domains from two adjacent modules (SrfB1 and SrfB2) of the SrfB subunit of surfactin synthetase as well as for other heterologous apo-ACP, -PCP, and -ArCP domains. Utilization of CoA analogues by Sfp is also reported. We also present in vivo and in vitro analysis of Sfp mutants, generated by site-directed mutagenesis, with substitution at amino acids that constitute a signature motif for PPTases.

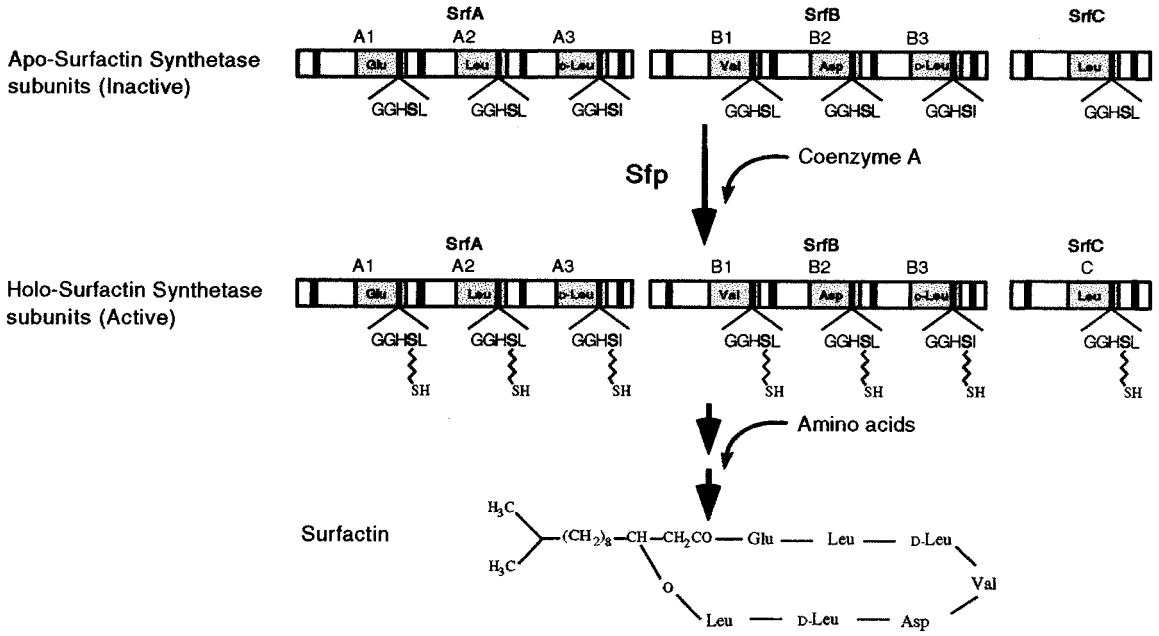
EXPERIMENTAL PROCEDURES

Materials, Bacterial Strains, and Plasmids. YT broth (2×) and DSM agar (Difco sporulation agar) were prepared and used for culturing of *B. subtilis* and *E. coli* strains as previously described (*16*, *17*). Competent cells of *E. coli* strain DH5α were purchased from GibcoBRL. Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs. The expression vector pET28b(+) and competent cells of *E. coli* strain BL21 (DE3) were purchased from Novagen. Plasmid pUC8-Sfp expressing the *sfp* gene from *B. subtilis* has been previously described (*10*). The N-terminal His_6-tagged proteins H_6SrfB1.143 (143 kDa), H_6SrfB1.73 (73 kDa), H_6SrfB1.18 (18 kDa), and H_6SrfB2.16 (16 kDa) were overproduced and purified as described in the accompanying paper (*18*). *E. coli* 12.1-kDa EntB-ArCP-H_6 (*9*) and *Saccharomyces cerevisiae* 14.4-kDa Lys2-PCP-H_6 (unpublished data) were kindly provided by Amy M. Gehring and David E. Ehmann. The *E. coli* 8.4-kDa apo-ACP subunit (*19*, *20*) was kindly provided by Roger S. Flugel. [^{3}H]CoA [~70% label in phosphopantetheinyl moiety as determined by previously described methods (*21*)] was prepared by DuPont New England Nuclear. [^{3}H]CoA was acetylated to improve storage stability, and the acetylated fraction and remaining disulfide CoA were purified as previously described (*21*). Unlabeled CoA, acetyl-CoA, benzoyl-CoA, phenylacetyl-CoA, desulfo-CoA, and 3′-dephospho-CoA were purchased from Sigma. Acetonyldethio-CoA and homocysteamine-CoA (*7*) were kindly provided by Dale G. Drueckhammer. The synthetic SrfB1 peptide G987-M1005 (GVTDNFFMIGGHSLKAMMM) was obtained from the Biopolymer Laboratory of the Department of Biological Chemistry and Molecular Pharmacology at Harvard Medical School.

Recombinant DNA Methods. Recombinant DNA techniques were performed as previously described (*16*). Plasmid DNA preparation, gel purification of DNA fragments, and purification of DNA fragments amplified by polymerase chain reaction (PCR) were performed using QIAprep, QIAEX II, and QIAquick kits, respectively (QIAGEN). PCRs were carried out using *pfu* DNA polymerase as described by the enzyme suppliers (Stratagene). Splicing by overlap extension (SOE) reactions (*22*) were performed using approximately an equimolar ratio (total amount ca. 50 ng) of each of the gel-purified PCR-amplified DNA fragments to be joined as template. The fidelity of the SOE- or PCR-amplified DNA fragments was established by nucleotide sequencing after subcloning into the corresponding expression vector. DNA sequencing was performed on double-stranded DNA by the Molecular Biology Core Facility of the Dana Farber Cancer Institute (Boston, MA). Oligonucleotide primers were obtained from Integrated DNA Technologies.

Cloning, Overproduction, and Purification of H_6SrfB1.17. The fragment (14.3 kDa, residues 960−1084) containing the PCP domain from the first module (SrfB1) of the 401-kDa three-module SrfB subunit of surfactin synthetase (Figure 2A) was PCR-amplified from the plasmid pZW6.2, which is described in the accompanying paper (*18*). The PCR primers utilized (F-44 and R-42) are shown in Table 1. The SrfB1.17 forward (F-44) and the reverse (R-42) primers introduced *Nde*I and *Hin*dIII restriction sites, respectively, in the PCR product. The amplified fragment was digested with *Nde*I and *Hin*dIII, gel-purified, and cloned into pET28b(+) digested with *Nde*I and *Hin*dIII to create the plasmid pSrfB1.17. The plasmid pSrfB1.17, containing the translational fusion (H_6SrfB1.17 protein) between the fragment of 21 amino acid residues (including a N-terminal His

Scheme 1. Schematic Representation of Phosphopantetheinylation of apo-Surfactin Synthetase by Sfp[a]



[a] The PPTase Sfp is proposed to catalyze the transfer of a phosphopantetheinyl moiety from CoA to each of the seven serine residues embedded in the PCP signature sequence [FFxxLGG(D/H)S(L/I)] located in each of the seven PCP domains, respectively, of the first three subunits (SrfABC) of apo-surfactin synthetase (upper section of the scheme). This posttranslational modification yields the active holo-surfactin synthetase (middle section of the scheme) catalytically competent for the loading of amino acid and subsequent formation of peptide bonds required for biogenesis of the lipopeptide antibiotic surfactin (lower section of the scheme) by the thiotemplate mechanism.

Table 1: Oligonucleotide Primers for Site-Directed Mutagenesis[a]

Comment	Primer	Nucleotide Sequence
G105A	F-36	5'-CAGCCGATCG**C**CATAGATATCGAAAAA-3'
G105A	R-35	5'-TTTTTCGATATCTATG**C**GATCGGCTG-3'
G105D	F-16	5'-CAGCCGATCG**A**CATAGATATCGAAAAA-3'
G105D	R-17	5'-TTTTTCGATATCTATG**T**CGATCGGCTGT-3'
D107A	F-18	5'-CAGCCGATCGGCATAG**C**TATCGAAAAA-3'
D107A	R-19	5'-TTTTTCGATA**G**CTATGCCGATCGGCTG-3'
D107E	F-38	5'-CAGCCGATCGGCATAGA**G**ATCGAAAAA-3'
D107E	R-37	5'-TTTTTCGAT**C**TCTATGCCGATCGGCTG-3'
W147A	F-20	5'-TATCATCTA**G**C**G**TCAATGAAAGAAAGC-3'
W147A	R-21	5'-GCTTTCTTTCATTGAC**G**C**T**AGATGATA-3'
W147F	F-34	5'-TATCATCTAT**T**T**T**CAATGAAAGAAAGC-3'
W147F	R-35	5'-GCTTTCTTTCATTGA**A**A**A**TAGATGATA-3'
E151A	F-24	5'-CTATGGTCAATGAAAG**C**AAGCTTTATC-3'
E151A	R-25	5'-GATAAAGCTT**G**CTTTCATTGACCATAG-3'
K155A	F-26	5'-GAAAGCTTTATCG**C**ACAGGAAGGCA-3'
K155A	R-27	5'-TGCCTTCCTGT**G**CGATAAAGCTTTC-3'
Sfp-T	F-40	5'-GCC**TCTAGA**AUGGGTGCGTTTGATTCACAGCC-3'
SrfB1.17	F-44	5'-GCGCCCAT**CATATG**CAGCCGGAATACGCAGCACCA-3'
SrfB1.17	R-42	5'-GCAG**AAGCTT**AAGCGGGGACGTTGTAGCTTGTGTTTGC-3'
Sfp	F-14	5'-AAGCAGGCAGTATCAGTTGGAC-3'
Sfp	R-15	5'-GGCGCATACAGCCATTTTGTAG-3'

[a]Mismatch mutations are underlined. The sequences in bold in the primers F-40, F-44, and R-42 indicate restriction sites for *Xba*I, *Nde*I, and *Hin*dIII, respectively.

tag) encoded in pET28b(+) and the 125 residue PCP domain from SrfB1, was introduced by transformation into *E. coli* strain BL21(DE3). For overproduction of H_6SrfB1.17 (16.6-kDa), the *E. coli* strain BL21(DE3)/pSrfB1.17 was cultivated (1 L) with shaking at 37 °C in YT broth (2×) containing 50 μg/mL kanamycin to an OD_{600} of 0.5−0.7 and induced with 1 mM isopropyl-1-thio-β-D-galactoside (IPTG). Incubation was continued for a period of 4 h before the cells were harvested by centrifugation (10 min at 2000*g*) and resuspended in 35 mL of a solution of 5 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, pH 7.9. Cells were disrupted by two passages through a French pressure cell (18 000 psi). Cellular debris was removed from the lysate by centrifugation (30 min at 17000*g*). The N-terminal His_6-tagged H_6SrfB1.17 protein was purified by nickel column chromatography using His·Bind Resin (Novagen) according to the manufacturer's instructions. Fractions of the eluant from the nickel column were analyzed by SDS−PAGE, and those containing H_6SrfB1.17 (>90% pure) were pooled and dialyzed against 10 mM Tris-HCl (pH 7.5−8.0), 1 mM EDTA, and 10% glycerol. After dialysis, the samples were concentrated to 100−200 μM with a Centriprep 3 concentrator (Amicon), aliquoted, flash-frozen in liquid nitrogen, and stored at −80 °C. The purification strategy yielded ~25 mg of pure (>90%) H_6SrfB1.17.

Mutagenesis of the sfp Gene. Site-directed mutants G105A, G105D, D107A, D107E, W147A, W147F, E151A, and K155A were constructed using the SOE method (*22*). In the first round of PCR, the sequence upstream and downstream of the mutation was amplified separately using the plasmid pUC8-Sfp as template and the forward/reverse (F/R) primer pairs shown in Table 1. The resulting PCR-amplified fragments were gel-purified and subjected to a second round of PCR using the Sfp forward (F-14) and reverse (R-15) primer pair (Table 1). The resulting amplified fragment was digested with *Eag*I and *Bsp*EI, gel-purified, and subcloned into pUC8−Sfp digested with *Eag*I and *Bsp*EI. The pUC8-Sfp derivatives containing the mutations in *sfp* were introduced into *E. coli* DH5α. Sfp-T (a truncated version of Sfp encompassing residues I96M−L224) was constructed from pUC8-Sfp by introducing an in-frame

deletion in the *sfp* gene that encompassed codons 2−96 using PCR. To introduce the deletion, Sfp-T forward primer F-40 (encompassing the *Xba*I site upstream of the AUG codon of *sfp*, the AUG codon, and the Gly97 codon at its 5′-end) and Sfp reverse primer R-15 were used to amplify a 344-bp fragment from the *sfp* gene. The amplified fragment was digested with *Xba*I and *Bsp*EI, gel-purified, and subcloned into pUC8-Sfp digested with *Xba*I and *Bsp*EI to create the plasmid pUC8-SfpT. This plasmid was introduced by transformation into *E. coli* DH5α.

Overproduction and Purification of Sfp Proteins. Overproduction of *B. subtilis* wild-type Sfp in the *E. coli* strain MV1190/pUC8-Sfp (*10*) and purification were performed based on previously described procedures (*8*, *10*). *E. coli* MV1190/pUC8-Sfp was cultivated overnight with shaking at 37 °C in YT broth (2×) containing 100 μg/mL ampicillin. Cells from 1 L of culture were harvested by centrifugation (10 min at 2000*g*) and resuspended in TMGD buffer (10 mM $MgCl_2$, 5 mM DTT, 5% glycerol, and 50 mM Tris-HCl, pH 8.0). After resuspension, all steps were performed at 4 °C unless otherwise noted. Cells were disrupted by two passages through a French pressure cell (18 000 psi). Cellular debris was removed from the lysate by centrifugation (30 min at 17000*g*), and crystalline ammonium sulfate was added (35% saturation) to the lysate supernatant. After 1 h at 4 °C with slight agitation, the precipitate was collected by centrifugation (30 min at 17000). The supernatant was subjected to a second round of ammonium sulfate precipitation (55% saturation). The precipitate was collected by centrifugation and dissolved in 20 mL of TMGD buffer. The solution was applied to a Sephacryl S-100 column (Pharmacia, 2.5 × 115 cm, 1 mL/min flow rate) equilibrated with TMGD buffer. The proteins were eluted with the same buffer at a flow rate of 1 mL/min. Fractions (8 mL each) were analyzed by SDS−PAGE, and those containing Sfp (>90% pure) were pooled, concentrated to 2−5 mg/mL with a Centriprep 10 concentrator (Amicon), aliquoted, and flash-frozen in liquid nitrogen for storage at −80 °C. Overproduction of Sfp mutants in the *E. coli* strain DH5α containing the pUC8-Sfp derivatives expressing the mutated *sfp* genes was performed at 30 °C. Purification and storage of soluble Sfp mutant proteins were performed in the same way as that described above for wild-type Sfp.

Protein concentration was determined by the Bradford assay (*23*) using dye concentrate reagent from Bio-Rad. Bovine serum albumin (BSA) was used as standard.

Assay for Apo-protein to Holo-Protein Conversion by [^{3}H]Phosphopantetheine Group Transfer from [^{3}H]Coenzyme A. The radioassay for determination of PPTase activity has been previously described (*8*). Typically, the apo-protein or the synthetic peptide G987-M1005 and the [^{3}H]CoA cosubstrate were incubated (30 min at 37 °C, unless otherwise indicated) with the PPTase in a 100-μL reaction mixture containing 10 mM $MgCl_2$, 5 mM DTT, and 75 mM MES/sodium acetate buffer, pH 6.0. Acetylated [^{3}H]CoA was hydrolyzed to the free thiol form as previously reported (*21*) before its addition to the reaction mixture; whereas the disulfide [^{3}H]CoA was added directly to the reaction mixture already containing DTT (5 mM) for reduction to its free thiol form. Apo-proteins were added to the reaction from stocks (100−200 μM) in 1 mM EDTA, 10% glycerol, and 10 mM Tris-HCl, pH 7.5−8.0. The synthetic peptide G987-M1005 was added from a 1 mM stock in water. Reactions were started by adding the PPTase in 10 μL of Sfp dilution buffer (1 mM EDTA, 10% glycerol, 1 mg/mL BSA, and 10 mM Tris-HCl, pH 8.0). Enzyme concentration in the assay was typically in the low nanomolar range, except for the D107E mutant used in concentrations of up to 1 μM. Reactions were quenched with 800 μL of 10% trichloroacetic acid (TCA) with BSA (375 μg) added as carrier. Precipitated proteins were pelleted by centrifugation, and the pellets were washed three times with 800 μL of 10% TCA before dissolving them in 150 μL of 1 M Tris base. The redissolved proteins were mixed with 3.5 mL of liquid scintillation cocktail, and the amount of radioactivity incorporated was quantified by liquid scintillation counting. Reactions were routinely performed in triplicate. Buffers Tris-HCl (for pH 7.0−9.5) and MES/sodium acetate (for pH 4.5−6.5) were used at 75 mM final concentration in the PPTase reaction for determination of the optimum pH for PPTase activity.

The kinetic data (PPTase rate vs substrate concentration) derived from the radioassay was fit to the Michaelis−Menten equation or to the general substrate inhibition equation (*24*) in the case of PPTase rate vs Lys2-PCP-H_6 concentration.

Utilization of Coenzyme A Analogues by Sfp. Phosphopantetheinylation reactions (100 μL) using desulfo-CoA, homocysteamine-CoA, acetyl-CoA, acetonyldethio-CoA, benzoyl-CoA, phenylacetyl-CoA or 3′-dephospho-CoA, and the apo-H_6SrfB1.17 domain or the *E. coli* apo-ACP subunit were performed as described above except for the following modification. CoA analogues were included in the reaction mixture at 250 μM, apo-proteins at 60 μM, and Sfp at 200 nM. For assessment of apo- to holo-PCP/ACP conversion, the reactions were incubated for 3 h at 37 °C and subsequently submitted for matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF) analysis (Biopolymer Laboratory, Howard Hughes Medical Institute, Harvard Medical School).

Complementation by SPβ Specialized Transduction. The pUC8-Sfp derivatives bearing the mutant *sfp* alleles were introduced into the *B. subtilis* strain JH642 (*trpC2 pheA sfp*^{0}) by phage SPβ-mediated specialized transduction (*10*, *25*, *26*) to determine if the mutant products possessed PPTase activity in vivo. Plasmids were cleaved with *Eco*RI and then ligated with the 1.6-kb *Eco*RI fragment of pMMN13 (*27*) bearing a chloramphenicol resistance gene derived from plasmid pC194. The resulting plasmids were used to transform *B. subtilis* strain ZB307 [SPβc2del2::Tn917::pSK106 (*25*)] with selection for chloramphenicol resistance (Cm^R). The plasmids will integrate into the prophage by a single, Campbell-type recombination event with the plasmid sequences present in the phage. SPβ phage lysates were generated from the transformants by heat induction (*25*) and were used to lysogenize *B. subtilis* strain ZB449 (*trpC2 pheA abrB703* SPβ^c). Two lysogens were chosen as a source of specialized transducing phage. Phage lysates were used to lysogenize cells from a mid-log phase culture of strain JH642 with selection for Cm^R. Lysogens were plated onto DSM (Difco sporulation agar), which contained 0.1 mM $Ca(NO_3)_2$. The lysogenized *B. subtilis* strain JH642 expressing the different *sfp* alleles are indicated in Table 2. Surfactin production was detected by the formation of calcium-dependent halo around the colonies due to the precipitation of the lipopeptide in the presence of calcium (*28*).

Table 2: *Bacillus subtilis* strains

Strain	Relevant Genotype	*sfp* allele
LAB2735	SPβc2del2::Tn917::pSK10Δ6::pMMN352	*sfpW147A*
LAB2736	SPβc2del2::Tn917::pSK10Δ6::pMMN354	*sfpG105D*
LAB2737	SPβc2del2::Tn917::pSK10Δ6::pMMN356	*sfpW147F*
LAB2738	SPβc2del2::Tn917::pSK10Δ6::pMMN358	*sfpD107E*
LAB2740	SPβc2del2::Tn917::pSK10Δ6::pMMN360	*sfp*
LAB2741	SPβc2del2::Tn917::pSK10Δ6::pMMN362	*sfpK155A*
LAB2742	SPβc2del2::Tn917::pSK10Δ6::pMMN364	*sfpD107A*
LAB2743	SPβc2del2::Tn917::pSK10Δ6::pMMN366	*sfpE151A*
LAB2747	SPβc2del2::Tn917::pSK10Δ6::pMMN370	*sfpG105A*
LAB2749	SPβc2del2::Tn917::pSK10Δ6::pMMN378	*sfp-T*

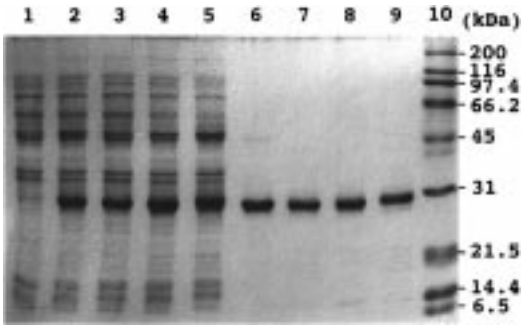


FIGURE 1: Levels of overproduction and purity of wild-type Sfp and mutants W147A, W147F, and D107E. Protein samples were resolved by SDS−PAGE (12%) and stained with Coomassie Blue. Lane 1 was loaded with cell lysate from the *E. coli* strains DH5α; lanes 2−5 were loaded with cell lysates from the *E. coli* strains overproducing wild-type Sfp, mutant W147F, mutant W147A, and mutant D107E, respectively; lanes 6−9 were loaded with wild-type Sfp and mutants W147A, W147F, and D107E, respectively, obtained after purification; lane 10 was loaded with molecular weight markers.

RESULTS

Production and Purification of B. subtilis Sfp in E. coli. Wild-type Sfp (224 amino acid residues) from *B. subtilis* as well as three specific site-directed mutants were overproduced in *E. coli* and purified. Overproduction of wild-type Sfp was conducted at 37 °C. However, overproduction of the Sfp mutants was conducted at 30 °C to increase the proportion of soluble enzyme; less soluble enzyme was recovered when the mutants were overproduced at 37 °C (data not shown). Among those mutants that complemented the inactive *sfp*^{0} allele to render transformants producing surfactin of the *B. subtilis* strain JH642 of Srf$^-$ phenotype (mutants G105A, W147A, W147F, and K155A, see below), only the mutants W147A and W147F were overproduced in *E. coli* as soluble proteins and could be purified to homogeneity. Despite the activity of the G105A mutant in vivo and in vitro (see below), the protein was insoluble and it was not further characterized. The mutant D107E was also overproduced as soluble protein and purified. Figure 1 shows the levels of overproduction of wild-type Sfp and mutants W147A, W147F, and D107E in cell lysates and the level of purity achieved by subsequent purification steps. Typical yield of wild-type Sfp and the D107E mutant was ∼20 mg/L of culture, whereas the yield of Sfp mutants W147F and W147A was 7−9 mg/L.

Phosphopantetheinylation of Apo-SrfB1 PCP Domain by Sfp. To assess the catalytic capacity of Sfp to function as a PPTase (eq 1), [^{3}H]CoA was used as cosubstrate along with the apo-form of a surfactin synthetase PCP domain containing the conserved serine residue embedded in a consensus PCP sequence motif [FFxxLGG(D/H)S(L/I)] (*29*). As shown in Scheme 1, there are seven potential substrate sites for Sfp in the first three subunits (SrfABC) of surfactin synthetase (*17*, *27*, *30*, *31*, *32*). The PCP domain in each amino acid-activating module must be converted from the apo- to holo-form (eq 1) in order for each module to activate its cognate substrate as an amino acyl-S-enzyme intermediate, the prerequisite for subsequent peptide bond formation. In our initial test to validate Sfp as a PPTase, we chose as substrates three fragments encompassing residues 92−1308 (137 kDa), residues 462−1089 (69 kDa), and residues 960−1085 (14 kDa) from the first module (SrfB1) of the 401-kDa three-module SrfB subunit. These substrates (H_6SrfB1.143, H_6SrfB1.73, and H_6SrfB1.18, respectively) were overproduced in and purified from *E. coli* as N-terminal His_6-tagged derivatives (Figure 2A). The His_6-tagged substrates were purified by affinity chromatography on a nickel column as described in the accompanying paper (*18*). We previously reported both covalent incorporation of [^{3}H]P-pant and then subsequent transfer of [^{14}C]valine to the largest (H_6SrfB1.143: residues 92−1308 of SrfB) of these substrates (*8*). The first ∼350 amino acid residues of SrfB are thought to encode a condensation domain, and the next ∼500 residues encode an adenylation domain, followed by the PCP domain of ∼100 residues and containing S999 embedded in the consensus sequence that marks a PCP domain (*2*, *33*). Figure 2B shows a time course for the [^{3}H]phosphopantetheinylation of the apo-protein H_6SrfB1.143, H_6SrfB1.73 (encompassing residues 462−1089 from SrfB), and H_6SrfB1.18 (encompassing residues 960−1085 from SrfB) as determined by measurement of TCA precipitable radioactivity. The results clearly indicate that under the conditions of the assay the protein H_6SrfB1.18, containing the 126 residue PCP domain of SrfB1, was the most efficiently modified substrate of the three tested. More detailed quantitation of the mole fraction of apo to holo conversion of the SrfB1 fragments is presented in the accompanying paper (*18*). On this basis and also because of the ease of purification and facility of mass spectrometric analysis of 18-kDa vs 143-kDa reaction products, we utilized the H_6SrfB1.18 or the pET28b-derived H_6SrfB1.17 substrate (the latter encompassing residues 960−1084 of SrfB) as preferred substrates for further kinetic analysis of Sfp PPTase activity.

Before further analysis of Sfp PPTase activity, the pH profile of the enzyme was determined. Figure 3 shows that Sfp has a clear pH optimum at pH 6.0 with H_6SrfB1.18, displaying less than 20% of V_{max} at both pH 5.0 and pH 7.0, allowing a clear choice of pH for determination of K_m and k_{cat}. Subsequent kinetic analysis indicated a K_m of 0.7 μM for CoA and a k_{cat} of 102 min^{-1} (Figure 4A). With H_6SrfB1.17 apo-PCP as the variable substrate (Figure 4B), a K_m of 1.3 μM and a k_{cat} of 104 min^{-1} were obtained. Thus, the k_{cat} values determined with either the PCP domain or CoA as variable substrate are in reasonable agreement. In contrast to the severe inhibition of *E. coli* ACPS by various apo-ACP substrates (*7*), there is not substantial inhibition of Sfp by H_6SrfB1.17 under the assay conditions.

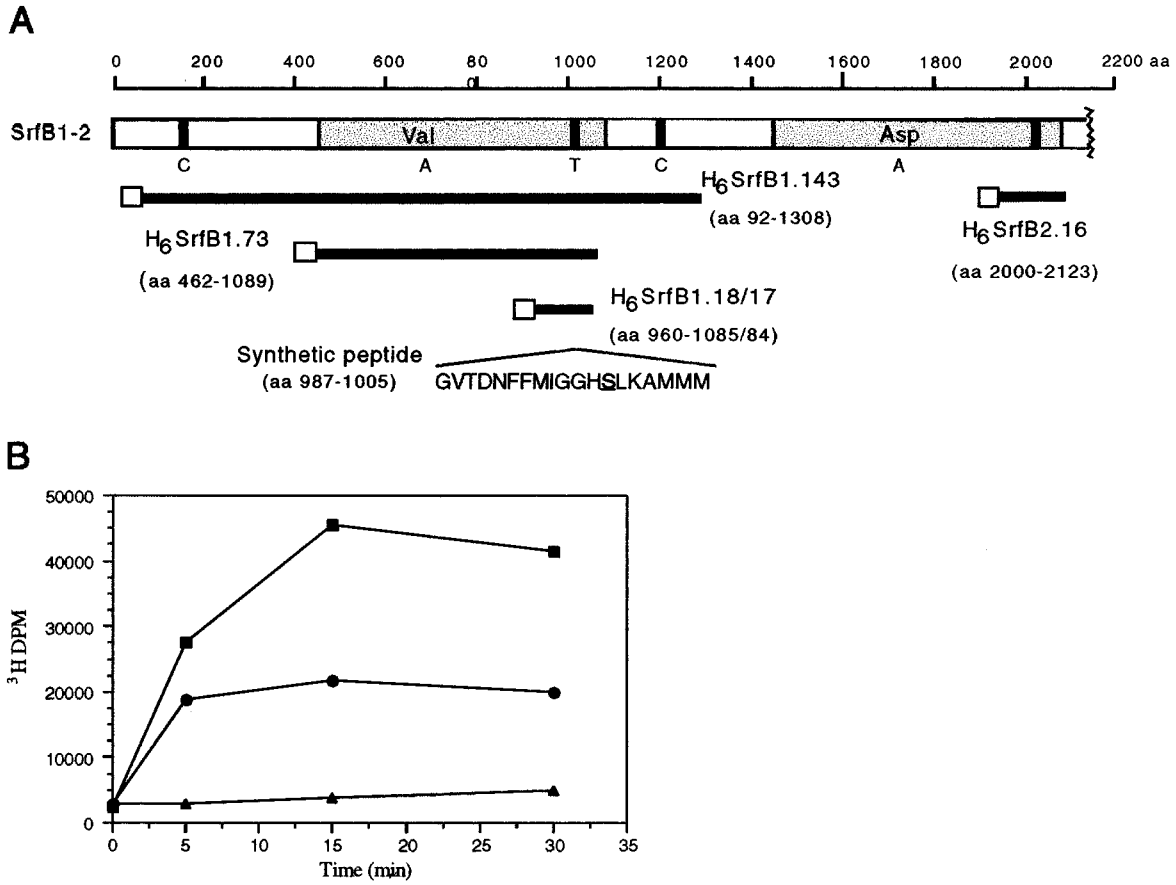


FIGURE 2: Schematic representation of recombinant fragments of SrfB used as P-pant acceptors in this study and time course for incorporation of [^{3}H]phosphopantetheine into His_6-tagged SrfB1 fragments. (A) View of modules 1 and 2 of the SrfB subunit (top of the diagram) and position of the SrfB fragments (thick lines) overproduced as N-terminal His_6-tagged proteins H_6SrfB1.143, H_6SrfB1.73, H_6SrfB1.18, H_6SrfB1.17, and H_6SrfB2.16 and the synthetic peptide G987-M1005. The conserved serine residue (S999) embedded in the peptide G987-M1005 is indicated. The His_6 tags are indicated as open squares. The positions of the condensation (C), adenylation (A), and PCP or thioesterification (T) domains in the SrfB1-2 fragment are indicated. The numbers in the scale represent amino acid residues (aa). (B) Incorporation of [^{3}H]P-pant into S999 of apo-H_6SrfB1.18 (squares), apo-H_6SrfB1.73 (triangles), and apo-H_6SrfB1.143 (circles) was measured by radioassay. Reaction mixtures included 4 μM apo-SrfB1 substrate, 40 μM [^{3}H]CoA (580 Ci/mol), 20 nM Sfp, 10 mM $MgCl_2$, 5 mM DTT, and 75 mM Tris-HCl, pH 8.8.

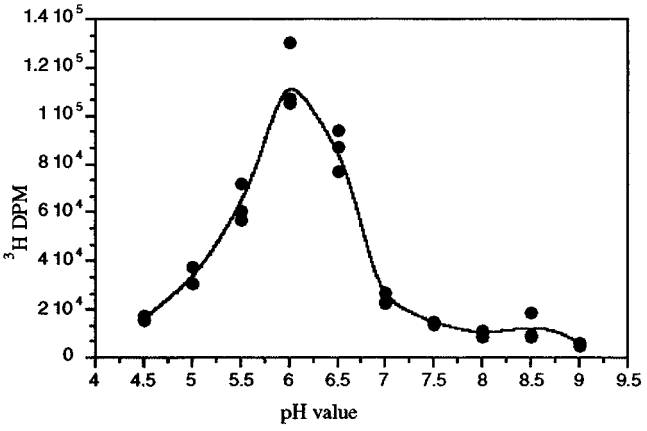


FIGURE 3: Effect of pH on the phosphopantetheinyl transferase activity of Sfp. PPTase activity was measured as incorporation of [^{3}H]P-pant into S999 of apo-H_6SrfB1.18 monitored by radioassay. Reaction mixtures included 10 μM apo-H_6SrfB1.18, 40 μM [^{3}H]-CoA (96 Ci/mol), and 2.6 nM Sfp and were incubated at 37 °C for 30 min. Buffers Tris-HCl (for pH 7.0−9.0) and MES/sodium acetate (for pH 4.5−6.5) were utilized at 75 mM final concentration in the reactions.

In addition to the initial velocity assays, we wished to utilize Sfp and [^{3}H]CoA to quantitate the amounts of apo-PCP substrates convertible to holo forms. Those results are reported in the accompanying paper (*18*), where quantitative conversion of apo-H_6SrfB1.18 PCP could be demonstrated by HPLC separation of apo- and holo-forms and validated by mass spectrometry.

Phosphopantetheinylation of Apo-SrfB2 PCP Domain by Sfp. It is probable that Sfp services all seven PCP domains in the first three subunits of surfactin synthetase for post-translational phosphopantetheinylation. As an initial comparison of Sfp recognition of a second apo-PCP domain from surfactin synthetase, we next overproduced and purified H_6SrfB2.16 (16 kDa), encompassing the fragment of 21 amino acid residues (containing the N-terminal His tag) encoded by the vector fused to the 124 amino acid residue PCP domain (residues 2006−2123) from the adjacent downstream module, SrfB2, in the SrfB subunit (Figure 2A). Apo-H_6SrfB2.16 was indeed a substrate for Sfp, with a K_m of 1.8 μM, a k_{cat} of 56 min^{-1} (Table 3), and a catalytic efficiency (k_{cat}/K_m) of 31 vs 80 μM^{-1} min^{-1} for the apo-H_6SrfB1.17 domain.

Heterologous Apo-PCP, Apo-ACP, and Apo-ArCP Domains as Substrates of Sfp. To further investigate the ability of Sfp to modify serine side chains in heterologous domains, we tested three other apo-protein fragments as substrates of Sfp. Figure 5, panels A−C, shows that the C-terminal His_6-

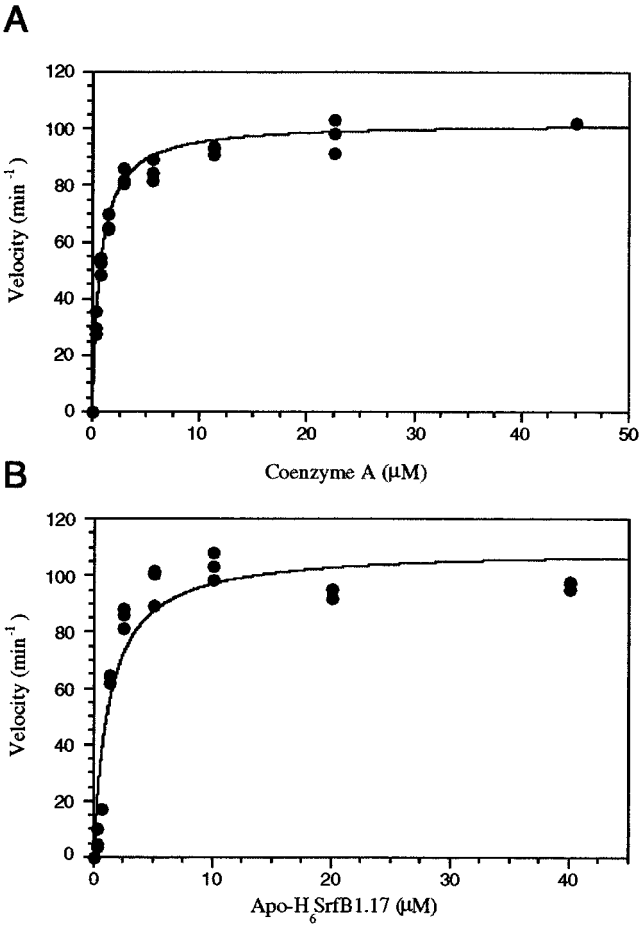


FIGURE 4: Effect of CoA and apo-H_6SrfB1.17 concentration on the rate of Sfp-catalyzed phosphopantetheinyl transfer. Plot A shows the PPTase reaction velocity as a function of CoA concentration. Reaction mixtures included [^{3}H]CoA (96 Ci/mol), 10 μM apo-H_6SrfB1.17, and 1 nM Sfp. Plot B shows the reaction velocity as a function of apo-H_6SrfB1.17 concentration. Reaction mixtures included apo-H_6SrfB1.17, 40 μM [^{3}H]-CoA (96 Ci/mol), and 1 nM Sfp. All reaction mixtures were incubated at 37 °C for 30 min. PPTase activity was measured as incorporation of [^{3}H]P-pant into apo-H_6SrfB1.17 monitored by radioassay.

Table 3: Catalytic Efficiency for Phosphopantetheinylation of Recombinant apo-PCP, apo-ACP and apo-ArCP Domains by Sfp

substrate	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (μM^{-1} min^{-1})
apo-H_6SrfB1.18 (*B. subtilis*)	1.3	104	80
apo-H_6SrfB2.18 (*B. subtilis*)	1.8	56	31
apo-EntB−ArCP−H_6 (*E. coli*)	16	65	4
apo-Lys2-PCP−H_6 (*S. cerevisiae*)	~5	>70	>14
apo-ACP (*E. coli*)	6	5.8	1

tagged apo-PCP domain (residues 805−924) of *S. cerevisiae* Lys2 (*34*), the C-terminal His_6-tagged apo-ArCP domain (residues 188−285) of *E. coli* EntB (*9*), and the *E. coli* apo-ACP subunit of 77 amino acid residues (*20*) are all processed with good K_m and high modification stoichiometries by Sfp. The kinetic parameters for phosphopantetheinylation of the heterologous substrates by Sfp are shown in Table 3. The k_{cat}/K_m data for these recombinant substrates show that, by this catalytic efficiency criterion, the *S. cerevisiae* Lys2 PCP domain is about one-sixth to one-half as good a substrate as the two SrfB1 and SrfB2 homologous PCP domains. However, and despite the mild substrate inhibition that is observed, Lys2 PCP is the heterologous substrate most

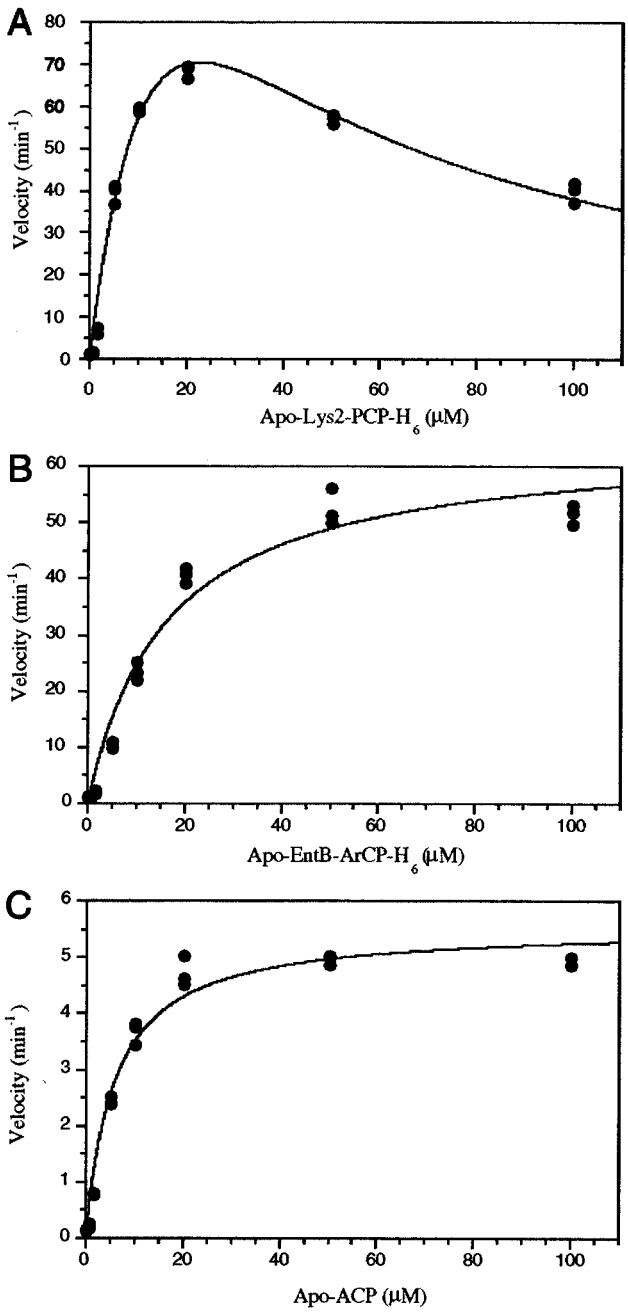


FIGURE 5: Sfp-catalyzed phosphopantetheinyl transfer to heterologous apo-PCP, apo-ACP, and apo-ArCP domains. Plots A−C show the PPTase reaction velocity as a function of apo-Lys2-PCP−H_6, apo-EntB-ArCP-H_6, and apo-ACP concentration, respectively. Reaction mixtures including the apo-protein and 100 μM [^{3}H]CoA (192 Ci/mol) were incubated at 37 °C for 30 min with 1 nM Sfp (Lys2 PCP and EntB ArCP containing reactions) or 10 nM Sfp (ACP containing reactions). PPTase activity was measured as incorporation of [^{3}H]P-pant into apo-substrate monitored by radioassay.

efficiently modified from those tested. The efficiency of EntB ArCP phosphopantetheinylation is 8−20-fold lower than that of the homologous PCP domains, and the *E. coli* ACP subunit, with k_{cat}/K_m being 8- to 30-fold lower than that of the homologous substrates, is clearly the worst substrate of those tested.

Assessment of a SrfB1-PCP Synthetic Peptide as a Potential Substrate of Sfp. To assess the ability of Sfp to recognize the consensus FFxxLGG(D/H)S(L/I) sequence in a smaller fragment, a peptide of 19 amino acids encompass-

Table 4: Mass Spectral Data Demonstrating Modification of *E. coli* Apo-ACP

prosthetic group donor	molecular mass (Da) observed[a] apo-ACP after incubation	observed holo-ACP after incubation	observed mass increase	predicted mass increase
CoASH	8488	8828	340	340
desulfo-CoA	8514	8818	304	308
homocysteamine-CoA	8501	8866	365	354
acetonyldethio-CoA[b]	8522	8886	364	364
benzoyl-CoA	8509	8947	438	444
phenylacetyl-CoA	8502	8953	451	458
acetyl-CoA	8498	8875	377	382
3′-dephospho-CoA	8483	nd[c]	nd	260

[a] The apo-ACP corresponds to the fraction of unmodified substrate remaining after the incubation. [b] DTT was not added to the reaction mixtures that contained acyl-CoA substrates. [c] nd, no holo-ACP was detected after incubation.

ing residues G987 to M1005 (GVTDNFFMIGGHSLKA-MMM) from the SrfB1-PCP domain was synthesized and tested as a substrate for phosphopantetheinylation at S999 by Sfp. Incubation for 4 h in a reaction with 65 nM Sfp (65-fold the normal assay concentration) was followed by mass spectrometric analysis. Only the apo-form of the peptide (observed molecular mass of 2089 Da; calculated, 2087 Da) was detected. The presence of the covalently bound phosphopantetheinyl group would have added 340 mass units to the peptide, a mass increase readily detectable in the apo-H_6SrfB1.18 to holo-H_6SrfB1.18 conversions as noted in the accompanying paper (*18*) or in the apo to holo conversion of the *E. coli* ACP subunit as shown in Table 4. Moreover, no [^{3}H]phosphopantetheine group transfer from [^{3}H]CoA to the synthetic peptides was detected by the TCA precipitation assay. The 19 amino acid peptide is clearly not a substrate for Sfp. Nor was it an effective inhibitor; inclusion of the peptide at 200 μM did not reduce tritium incorporation from [^{3}H]CoA into apo-H_6SrfB1.17 catalyzed by Sfp. Thus Sfp is presumed to recognize and act on a folded apo-PCP domain structure.

Recognition of Coenzyme A Analogues. While CoA is the presumed natural substrate for donation of its phosphopantetheinyl moiety to the seven serine side chains, one in each of the seven apo-PCP domains of surfactin synthetase, questions arose as to whether CoA analogues with phosphopantetheinyl arms of altered length or with an acyl group already tethered to the thiol function would be substrates of Sfp. In this context, we have previously observed that *E. coli* ACPS will transfer both modified phosphopantetheinyl moieties from CoA analogues (*7*) and acyl-phosphopantetheinyl moieties from acyl-CoA (*35*) onto apo-ACP substrates. Sfp likewise will utilize desulfo-CoA (transfer of desulfophosphopantetheine) and homocysteamine-CoA (transfer of a prosthetic group extended by one CH_2 unit) as determined by mass spectrometric assessment (Table 4). In terms of acyl-CoA recognition, the hydrolytically stable carba analogue of acetyl-CoA, acetonyl-dethio-CoA as well as the aryl-phosphopantetheinyl units from benzoyl-CoA and phenylacetyl-CoA can be loaded onto apo-H_6SrfB1.18 and onto the *E. coli* fatty acid apo-ACP subunit by Sfp. Table 4 shows the mass spectral data demonstrating the modification of *E. coli* apo-ACP, whose mass determination by MALDI-TOF

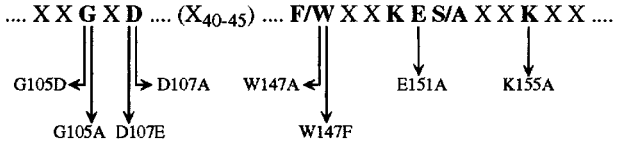


FIGURE 6: Conserved residues in the phosphopantetheinyl transferase enzyme superfamily and amino acid substitutions introduced into Sfp by site-directed mutagenesis. The PPTase signature sequence composed of two strings of conserved amino acids spaced by 40−45 undefined amino acid residues (aa) is shown in the top of the figure. The amino acid substitutions introduced in the PPTase signature sequence of Sfp are shown below the conserved sequence.

spectrometry could be obtained with more accuracy than that of apo-H_6SrfB1.18. In contrast, 3′-dephospho-CoA is not utilized by Sfp even though it has an intact phosphopantetheinyl moiety, suggesting that the 3′-phosphate group present in CoA will be an important recognition determinant.

Mutagenesis of Residues Conserved in Sfp and Other Phosphopantetheinyl Transferases. Starting with the amino acid sequence of ACPS and the C-terminal domain of *S. cerevisiae* fatty acid synthetase II (FASII), a search for homologs revealed a second family of PPTases of ca. 24−26 kDa on the basis of two short amino acid sequences shared among all the PPTases (*8*). We have validated the function of several PPTases by purification and enzymological analysis, including Sfp from *B. subtilis* (*8*), EntD from *E. coli* (*8*), and Lys5 from *S. cerevisiae* (*36*) (David E. Ehmann, Amy M. Gehring, and C.T.W., unpublished data) and Lpa14 from *B. subtilis* (P.Z. and C.T.W., unpublished data), all of which were primarily identified as PPTases on the basis of two short strings of homology with ACPS (Figure 6). In particular, G105, D107, W147, K150, E151, and K155 in Sfp are six residues conserved in all PPTases and, so far, are predictive of PPTase activity. As an initial assessment of the functional significance of these conserved residues, we have generated substitutions at five amino acid positions in Sfp as indicated in Figure 6 and conducted initial analyses both in vivo for complementation in a *sfp*$^-$ strain of *B. subtilis* for effects on surfactin production and in vitro for effects on PPTase activity. In addition, a truncated version of Sfp (Sfp-T composed of I96M−L224) corresponding to the 14-kDa *E. coli* ACPS and encoded by a *sfp* mutant gene with a deletion spanning codons 2−96 was also examined for in vivo complementation and in vitro PPTase activity.

Complementation by SPβ specialized transduction was performed to investigate whether the *sfp* allelic variants encoding the Sfp mutants G105A, G105D, D107A, D107E, E151A, W147A, W147F, K155A, and Sfp-T were able to restore the Srf$^+$ phenotype to the *B. subtilis* strain JH642 (Srf$^-$ phenotype) by complementing the *sfp*^{0} allele. This allele expresses a C-terminal truncated inactive Sfp variant. In the complementation analysis the Sfp mutants K155A, W147A, W147F, and G105A complemented the *sfp*^{0} allele rendering transformants producing surfactin (Figure 7). In contrast, the mutants G105D, D107A, D107E, E151A, and Sfp-T (truncated to the 14-kDa size of the presumed core PPTase represented by *E. coli* ACPS) failed to complement. Therefore, these mutants are likely to be defective either in stability (e.g., folding) or catalysis. The results of the in vivo complementation are most informative in the case of the Sfp mutants such as K155A, G105D, and D107A, which were overproduced as insoluble proteins in the heterologous host

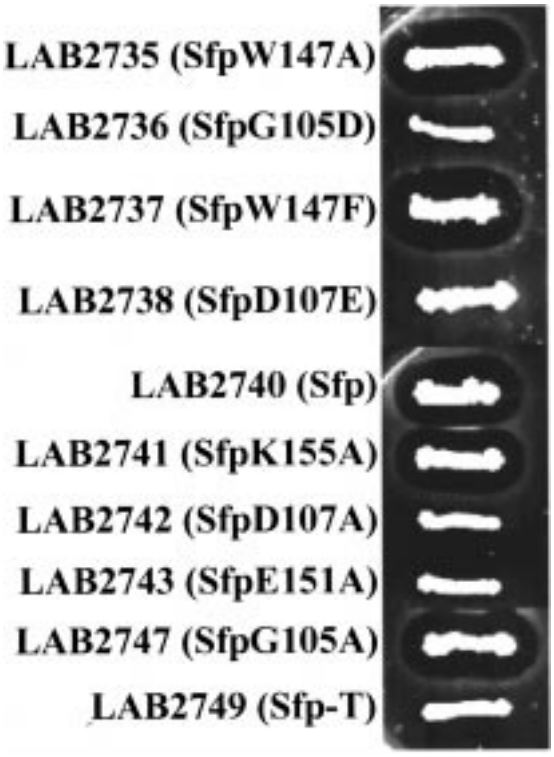


FIGURE 7: Complementation of *sfp*^{0} mutation by the SPβ-borne pUC8-Sfp derivatives bearing mutants alleles of *sfp*. The strains are listed on the left along with the name of the Sfp product (in parentheses) produced by each strain. On the right is a photograph of the colonies of each of the strains on DSM agar, showing the presence (indicating the Srf$^+$ phenotype of the strain) or absence (indicating the Srf$^-$ phenotype of the strain) of the calcium-dependent halo formation, due to the Ca^{2+} chelating activity of surfactin.

Table 5: Catalytic Properties of Sfp Mutants

protein	PPTase activity in lysate	k_{cat} (min^{-1})
Sfp	++++	104
G105A	+	nd[a]
G105D	−	nd
D107A	−	nd
D107E	−	0.02−0.04
W147F	+++	20
W147A	++	3.2−5.6
E151A	−	nd
K155A	−	nd
Sfp-T	−	nd

[a] nd, not determined.

E. coli and had no detectable PPTase activity in vitro (Table 5, see below). We could gauge the activity of these proteins by determining the ability of the corresponding mutant *sfp* alleles to complement the inactive *sfp*^{0} allele in *B. subtilis* strain JH642.

Parallel in vitro studies involved overproduction of the Sfp mutant proteins and assessment of whether the Sfp variants were overproduced as soluble proteins (at both 37 and at 30 °C, data not shown), determination of the PPTase activity in crude extracts, and purification and determination of catalytic properties of active Sfp variants that were overproduced in soluble form (Figure 1). In crude lysate, the Sfp mutants W147A and W147F were clearly active as was the G150A mutant (Figure 8). Other mutants, including K155A that complemented in vivo, had no detectable PPTase activity. The two W147 mutants (W147A and W147F) were

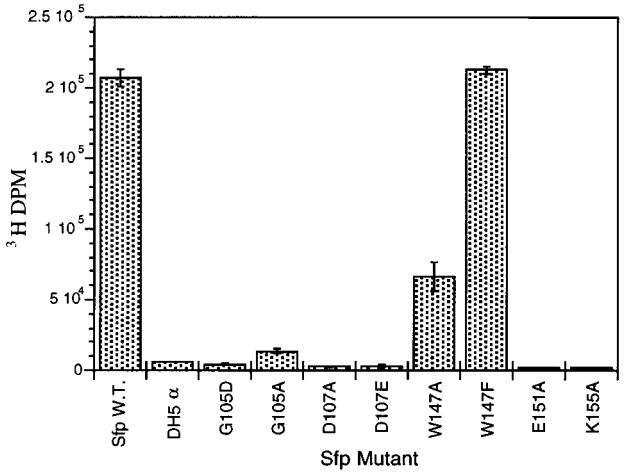


FIGURE 8: PPTase activity in cell lysates. PPTase activity in lysates of the strains overproducing the Sfp proteins was measured as incorporation of [^{3}H]P-pant into apo-H_6SrfB1.18 as monitored by radioassay. Reactions included 15 μM apo-H_6SrfB1.18 and 40 μM [^{3}H]CoA (96 Ci/mol) and were initiated by addition of ~10 μL of diluted lysate containing 3−4 μg of protein for the Sfp mutants or 3−4 ng for wild-type Sfp lysates, respectively. A reaction containing the lysate (3−4 μg of protein) of *E. coli* DH5α was included as control for host-endogenous PPTase activity. Reaction mixtures were incubated at 37 °C for 30 min. Each data point represents the average of two separate determinations.

purified to 90% homogeneity as was the D107E mutant (Figure 1). For these three mutants the K_m values for CoA (0.84, 0.82, and 1.3 μM for the mutants W147F, W147A, and D107E, respectively) and for apo-H_6-SrfB1.17 (1.3, 1.6, and 1.2 μM for the mutants W147F, W147A, and D107E, respectively) were indistinguishable from that of wild-type Sfp (0.7 μM for CoA and 1.3 μM for apo-H_6-SrfB1.17). However, there were discernible effects on k_{cat} values (Table 5). The k_{cat} of the W147F mutant was down 5-fold at 20 min^{-1}, while that of the W147A mutant was only 4% of the k_{cat} of wild-type Sfp. The purified D107E mutant had detectable activity but a k_{cat} value depressed some 3000-fold compared to that of wild-type Sfp, while the K_m values for CoA (1.3 μM) and apo-H_6SrfB1.17 (1.2 μM) were essentially unperturbed. These in vitro assays for PPTase activity only provided information on the catalytic ability of the Sfp mutants to modify the purified apo-H_6SrfB1.17; however, the in vivo complementation provided information as to whether a specific Sfp mutant is able to support surfactin production. This, in turn, indicated whether the mutant can efficiently modify all of the PCP domains in the surfactin synthetase, as is the case for the mutants K155A, W147A, W147F, and G105A.

DISCUSSION

The lipoheptapeptide antibiotic surfactin produced by *B. subtilis* is synthesized nonribosomally by an enzyme encoded, in part, by the *srfABCD* genes (*17*, *27*, *30*−*32*). The proteins SrfABC encoded in the operon activate amino acids 1−3, 4−6, and 7 of surfactin, respectively, while the protein SrfD appears to be a thioesterase homologue presumably involved in macrolactonization to the initiating β-hydroxy-acyl moiety. Genetic analysis revealed another gene, mapped just downstream of *srfD*, required for *sur*factin *p*roduction and hence termed *sfp* (*10*). This work extends our initial finding indicating that Sfp is an enzyme with catalytic ability

to posttranslationally phosphopantetheinylate the SrfB1 module (*8*) at the serine residue at position 999 (*18*). This study builds on that initial observation and delineates some of the elements of recognition and catalytic scope of the apo-PCP PPTase Sfp.

Chemically, the serine side chain modification catalyzed by Sfp and cognate PPTases generates a phosphodiester linkage and has some formal analogy to the posttranslational modification of serine residues by protein kinases. Thus, in these reactions the hydroxymethyl side chain of the serine residue embedded in a consensus sequence acts as nucleophile to attack a phosphoryl group in pyrophosphate linkage to the cofactor, namely, ATP for protein serine kinases and CoA for serine PPTases. Functionally, the apo- to holo-carrier domain conversion serves a priming function for all nonribosomal peptide synthetases, polyketide synthases, and depsipeptide synthetases in that acylation of the newly introduced cysteamine terminus of the P-pant prosthetic group is the central stratagem for monomer activation and condensation to give C−C, C−N, and C−O bond formation in the synthetase-catalyzed elongation steps of the respective biosynthetic pathways.

Partner-specific interaction has previously been demonstrated between such synthetases and their cognate phosphopantetheinyl transferases by both genetic and, more recently, biochemical evidence. The genetic evidence for partner specific recognition of apo-surfactin synthetase by Sfp was noted above, and there are corresponding data for the PPTases EntD and Gsp, involved in the biosynthesis of enterobactin in *E. coli* (*8*) and gramicidin S in *B. brevis*(*11*), respectively. We have shown that as a PPTase EntD will modify the apo-domains of both EntF (a PCP domain) and of EntB (an ArCP domain) (*8*, *9*), but will *not* act on the fatty acid synthase subunit apo-acyl carrier protein (*8*). The corresponding protein−protein based selectivity also exists in the other direction: *E. coli* ACPS phosphopantetheinylates the apo-ACP subunit but not the apo-EntB ArCP (*9*) or the apo-EntF PCP (*8*).

Against this backdrop, we have now begun to characterize *B. subtilis* Sfp as a presumed prototypic PPTase that recognizes apo-PCP domains. It is presumed that Sfp has to posttranslationally modify the conserved serine residues in each of the seven PCP domains with approximately equal catalytic efficiency for surfactin to be assembled in any significant yield. In this work, we have overproduced PCP domains as equivalent positional and sized fragments (∼124 residues) from two adjacent modules (B1 and B2) of the three-module SrfB subunit of surfactin synthetase and shown chemical and kinetic competence of Sfp toward each PCP domain with about a 2.5-fold variance in catalytic efficiency. Presumably, the other five PCP domains (SrfA1, A2, A3, B3, and C) would also be approximately equivalent substrates for Sfp. Although one characterizes phosphopantetheinylation as a *post*translational modification and we have now shown that several PPTases act on folded apo-PCP domains, the possibility of in vivo *co*translational modification of large, multidomainal peptide synthetases has not been established or ruled out. The differential phosphopantetheinylation efficiency by Sfp for the 18-kDa PCP domain (H_6SrfB1.18) > 143-kDa condensation/adenylation/PCP domain (H_6SrfB1.143) > adenylation/PCP domain (H_6SrfB1.73) may be an in vitro problem resulting from overproducing fragments that fold with unequal stability and/or unequal yields of native architecture.

When Sfp was tested with three heterologous apo-protein domains as substrates, *S. cerevisiae* Lys2, *E. coli* EntB, and *E. coli* apo-ACP, representing heterologous apo-forms of PCP, ArCP, and ACP domain variants, respectively, Sfp was able to phosphopantetheinylate all three, albeit with about an 80-fold range in catalytic efficiency. This is in marked contrast to both *E. coli* EntD and ACPS, which would not cross over between PCP and ACP apo-domains in *E. coli*. It remains to be seen if other PCP-recognizing PPTases will be as promiscuous as Sfp, but at the very least these results advance Sfp as a leading candidate for coexpression with heterologous peptide synthetases and even polyketide synthases to generate active, holo-forms *in vivo*, for example, in situations where combinatorial biosynthetic schemes are undertaken (*37*, *38*). In that context, the ability of Sfp to load acyl-pantetheinyl groups directly onto both apo-ACP or apo-PCP domains as shown here may augur well for introduction of different starter units, e.g., in polyketides, as recently shown for aromatic polyketides with *E. coli* ACPS (*35*), or in mixed polyketide/depsipeptide synthases (e.g., rapamycin synthase).

We have previously noted a small set of completely conserved residues in two regions of all PPTases [...GxD....(W/F)xxKE(S/A)xxK...] (*8*). The results reported here of the initial studies on mutagenesis of the PPTase signature, where the activity of Sfp mutants was investigated both in vivo and in vitro, correlate with each other for the most part and indicate only conservative substitutions at G105 and W147 are well tolerated. The D107E substitution reduced k_{cat} by a factor greater than 3 orders of magnitude without affecting K_m, an effect possibly consistent with a role of residue D107 in catalysis. Secondary structure predictions reveal nothing about the GXD tripeptide sequence but consistently suggest the (W/F)xxKE(S/A)xxK string is in an α-helical conformation in each PPTase analyzed (data not shown). Thus, the mutations in this string reported here may be disrupting an α-helical element crucial for enzyme−PCP interaction. Likewise, the drop in activity of D107E mutant and the 5- and 25-fold drops in activity of the mutants W147F and W147A, respectively, are only a k_{cat} effect with no change in K_m for either CoA or H_6SrfB1.17 substrates. Further insight will await structural studies of PPTases.

Given that neither the SrfB1 synthetic PCP peptide of 19 amino acid residues containing the consensus serine is phosphopantetheinylated by Sfp nor a corresponding synthetic ACP peptide of 15 residues is phosphopantetheinylated by *E. coli* ACPS (R. H. Lambalot and C.T.W., unpublished data), and that the peptides do not show significant inhibition of the PPTase activity of the enzymes, it is likely that Sfp and the more selective ACP-modifying PPTase ACPS are scanning for an architectural element in which the consensus serine is embedded. NMR structures of the *E. coli* ACP subunit (*39*), and more recently *Streptomyces coelicolor* actinorhodin ACP (*40*), show the modifiable serine residue at a bend at the end of an α-helix held in a fixed orientation by interactions with distal residues, probably creating a recognition platform. There is some analogy to this specific recognition of the 80−100 amino acid residue folded apo-ACP and apo-PCP domains by partner PPTases seen in two other posttranslational modifications, namely, lipoylation and

biotinylation, where analogous recognition of 80−100 amino acid sized folded domains/subunits within larger protein complexes by lipoyl and biotinyl transferases has been deciphered (*41*, *42*). These may reflect a common evolution for recognition and covalent posttranslational modification of multidomainal apo-proteins converted by such covalent marking into active catalysts enabled to conduct acyl transfers (P-pant-proteins), α-ketol transfers (lipoyl-proteins), and carboxyl transfers (biotinyl-proteins).

ACKNOWLEDGMENT

We thank Professor Dale G. Drueckhammer for kindly providing CoA analogues, A. M. Gehring and D. E. Ehmann for providing EntB-ArCP-H_6 and Lys2-PCP-H_6, and Roger. S. Flugel for providing *E. coli* apo-ACP.

REFERENCES

1. Walsh, C. T., Gehring, A. M., Weinreb, P. H., Quadri, L. E. N., and Flugel, R. S. (1997) *Curr. Opin. Chem. Biol. 1*, 309−315.
2. de Crécy-Lagard, V., Marlière, P., and Saurin, W. (1995) *C. R. Acad. Sci. Paris 318*, 927−936.
3. Leadlay, P. F., Staunton, J., Aparicio, J. F., Bevitt, D. J., Caffrey, P., Cortes, J., Marsden, A., and Roberts, G. A. (1993) *Biochem. Soc. Trans. 21*, 218−222.
4. Weber, G., Schorgendorfer, K., Schneider-Scherzer, E., and Leitner, E. (1994) *Curr. Genet. 26*, 120−124.
5. Lambalot, R. H., and Walsh, C. T. (1995) *J. Biol. Chem. 270*, 24658−24661.
6. Ritsema, T. (1997) Ph.D. Thesis, Leiden University, Leiden, The Netherlands.
7. Gehring, A. M., Lambalot, R. H., Vogel, K. W., Drueckhammer, D. G., and Walsh, C. T. (1997) *Chem. Biol. 4*, 17−24.
8. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., Lacelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) *Chem. Biol. 3*, 923−936.
9. Gehring, A. M., Bradley, K. A., and Walsh, C. T. (1997) *Biochemistry 36*, 8495−8503.
10. Nakano, M. M., Corbell, N., Besson, J., and Zuber, P. (1992) *Mol. Gen. Genet. 232*, 313−321.
11. Borchert, S., Stachelhaus, T., and Marahiel, M. A. (1994) *J. Bacteriol. 176*, 2458−2462.
12. Ku, J., Mirmira, R. G., Liu, L., and Santi, D. V. (1997) *Chem. Biol. 4*, 203−207.
13. Grossman, T. H., Tuckman, M., Ellestad, S., and Osburne, M. S. (1993) *J. Bacteriol. 175*, 6203−6211.
14. Kleinkauf, H., and von Döhren, H. (1996) *Eur. J. Biochem. 236*, 335−351.
15. Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R., and Morris, H. R. (1996) *J. Biol. Chem. 271*, 15428−15435.
16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor, New York.
17. Nakano, M. M., Marahiel, M. A., and Zuber, P. (1988) *J. Bacteriol. 170*, 5662−5668.
18. Weinreb, P. H., Quadri, L. E. N., Walsh, C. T., and Zuber, P. (1998) *Biochemistry, 37*, 1575−1584.
19. Magnuson, K., Jackowski, S., Rock, C. O., and Cronan, J. E. J. (1993) *Microbiol. Rev. 57*, 522−542.
20. Prescott, D. J., and Vagelos, P. R. (1972) *Adv. Enzymol. 36*, 269−311.
21. Elovson, J., and Vagelos, P. R. (1968) *J. Biol. Chem. 243*, 3603−3611.
22. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene 77*, 51−59.
23. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
24. Cleland, W. W. (1970) in *The Enzymes* (Boyer, P., Ed.) pp 1−65, Academic Press, New York.
25. Zuber, P., and Losick, R. (1987) *J. Bacteriol. 169*, 2223−2230.
26. Nakano, M. M., and Zuber, P. (1990) in *Genetics and Biotechnology of Bacilli* (Hoch, J., and Ganesan, A. T., Eds.) Academic Press Inc., New York.
27. Nakano, M. M., Magnuson, R., Myers, A., Curry, J., Grossman, A. D., and Zuber, P. (1991) *J. Bacteriol. 173*, 1770−1778.
28. Arima, K., Kakinuma, A., and Tamura, G. (1968) *Biochem. Biophys. Res. Commun. 31*, 488−494.
29. Stachelhaus, T., and Marahiel, M. A. (1995) *FEMS Lett. 125*, 3−14.
30. Cosmina, P., Rodriguez, F., de Ferra, F., Grandi, G., Perego, M., Venema, G., and van Sinderen, D. (1993) *Mol. Microbiol. 8*, 821−831.
31. Fuma, S., Fujishima, Y., Corbell, N., D'Souza, C., Nakano, M. M., Zuber, P., and Yamane, K. (1993) *Nucleic Acids Res. 21*, 93−97.
32. Nakano, M. M., and Zuber, P. (1989) *J. Bacteriol. 171*, 5347−5353.
33. Marahiel, M. A., Stachelhaus, T., and Mootz, H. (1997) *Chem. Rev. 97*, 2651−2673.
34. Morris, M. E., and Jinks-Robertson, S. (1991) *Gene 98*, 141−145.
35. Carreras, C. W., Gehring, A. M., Walsh, C. T., and Khosla, C. (1997) *Biochemistry 36*, 11757−11761.
36. Miller, K. G., and Bhattacharjee, J. K. (1996) *Gene 172*, 167−169.
37. McDaniel, R., Ebert-Khosla, S., Hopwood, D. A., and Khosla, C. (1995) *Nature 375*, 549−554.
38. Stachelhaus, T., Schneider, A., and Marahiel, M. A. (1995) *Science 269*, 69−72.
39. Holak, T. A., Nilges, M., Prestegard, J. H., Gronenborn, A. M., and Clore, G. M. (1988) *Eur. J. Biochem. 175*, 9−15.
40. Crump, M. P., Crosby, J., Dempsey, C. E., Parkinson, J. A., Murray, M., Hopwood, D. A., and Simpson, T. J. (1997) *Biochemistry 36*, 6000−6008.
41. Wallis, N. G., Allen, M. D., Broadhurst, R. W., Lessard, I. A. D., and Perham, R. N. (1996) *J. Mol. Biol. 263*, 463−474.
42. Wallis, N. G., and Perham, R. N. (1994) *J. Mol. Biol. 236*, 209−216.

BI9719861