Identification of a Phosphopantetheinyl Transferase for Erythromycin Biosynthesis in *Saccharopolyspora erythraea*

Kira J. Weissman,*[a] Hui Hong,[b] Markiyan Oliynyk,[a] Alexis P. Siskos,[a] and Peter F. Leadlay[a]

Phosphopantetheinyl transferases (PPTases) catalyze the essential post-translational activation of carrier proteins (CPs) from fatty acid synthases (FASs) (primary metabolism), polyketide synthases (PKSs), and non-ribosomal polypeptide synthetases (NRPSs) (secondary metabolism). Bacteria typically harbor one PPTase specific for CPs of primary metabolism ("ACPS-type" PPTases) and at least one capable of modifying carrier proteins involved in secondary metabolism ("Sfp-type" PPTases). In order to identify the PPTase(s) associated with erythromycin biosynthesis in Saccharopolyspora erythraea, *we have used the genome sequence of this organism to identify, clone, and express (in* Escherichia coli*) three candidate PPTases: an ACPS-type PPTase (*S. erythraea *ACPS) and two Sfp-type PPTases (a discrete enzyme (SePptII) and another that is integrated into a modular PKS subunit (SePptI)). In vitro analysis of these recombinant PPTases, with an acyl carrier protein-thioesterase (ACP-TE) didomain from the erythromycin PKS as substrate, revealed that only SePptII is active in phosphopantetheinyl transfer with this substrate. SePptII was also shown to provide complete modification of ACP-TE and of an entire multienzyme subunit from the erythromycin PKS in* E. coli. *The efficiency of the SePptII in phosphopantetheinyl transfer in* E. coli *makes it an attractive alternative to other Sfp-type PPTases for co-expression experiments with PKS proteins.*

Introduction

Phosphopantetheinyl transferases (PPTases) catalyze the post-translational modification of carrier proteins (CPs) in the multienzyme fatty acid synthases (FASs), polyketide synthases (PKSs), and non-ribosomal polypeptide synthetases (NRPSs). The CPs are either integrated into the synthases (type I) or are discrete subunits (type II), and depending on the nature of the attached intermediate, are called acyl carrier proteins (ACPs), aryl carrier proteins (ArCPs), or peptidyl carrier proteins (PCPs). Phosphopantetheinylation occurs by transfer of the 4′-phosphopantetheine (P-pant) prosthetic group from coenzyme A (CoASH) to a conserved serine residue in the CPs, converting the proteins from their inactive "apo" forms to their active "holo" forms.[1] The reactive thiol terminus of the P-pant provides an attachment site for chain-extension intermediates, while its length (20 Å) and flexibility allow the CPs to shuttle the growing chains between spatially separated active sites within the synthase complexes.[2] Each CP domain must be modified in order for the whole multienzyme to function; for example, the hexamodular PKS 6-deoxyerythronolide B synthase (DEBS) responsible for erythromycin biosynthesis has a phosphopantetheinyl group on each of seven ACPs distributed among DEBS 1, 2, and 3 (Figure 1).[3]

The PPTases have been classified into three structural groups, which also loosely correlate with their substrate specificity.[1, 4] The first is the "ACPS-type", exemplified by ACPS of *Escherichia coli*. These PPTases are ≈120 amino acids in size, function as homotrimers,[5, 6] and are fairly specific for CPs of primary metabolism, although the ACPSs from *E. coli*[7, 8] and *Streptomyces coelicolor*[9] can activate a wider range of substrates in vitro, including in the case of the *S. coelicolor* enzyme, ACPs from type I and type II PKSs. The second group is the "Sfp-type" PPTases, named after Sfp (surfactin phosphopantetheinyl transferase), which is required for the production of the lipoheptapeptide antibiotic surfactin in *Bacillus subtilis*.[1] These PPTases are monomeric[10] and approximately twice the size of the ACPS-type enzymes, which suggests that they evolved by gene duplication from an ACPS ancestor.[4] Although Sfp-type PPTases are thought to be optimized for CPs of secondary metabolism,[11] several exhibit a relaxed substrate specificity, and are capable of modifying both type I and type II ACPs and PCPs.[4, 7, 12] In fact, Sfp-type PPTases appear to have assumed the role of ACPSs in some organisms (e.g., *Haemophilus influenzae*, *Pseudomonas aeruginosa*, and *Synochosystis* PCC6803).[2, 11] The third type of

[a] *Dr. K. J. Weissman, Dr. M. Oliynyk, Dr. A. P. Siskos, Prof. P. F. Leadlay*
Department of Biochemistry, University of Cambridge
80 Tennis Court Road, Cambridge CB2 1GA (UK)
Fax: (+44) 1223-766-091
E-mail: kjw21@cus.cam.ac.uk

[b] *Dr. H. Hong*
Department of Chemistry, University of Cambridge
Lensfield Road, Cambridge CB2 1EW (UK)

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim DOI: 10.1002/cbic.200300775

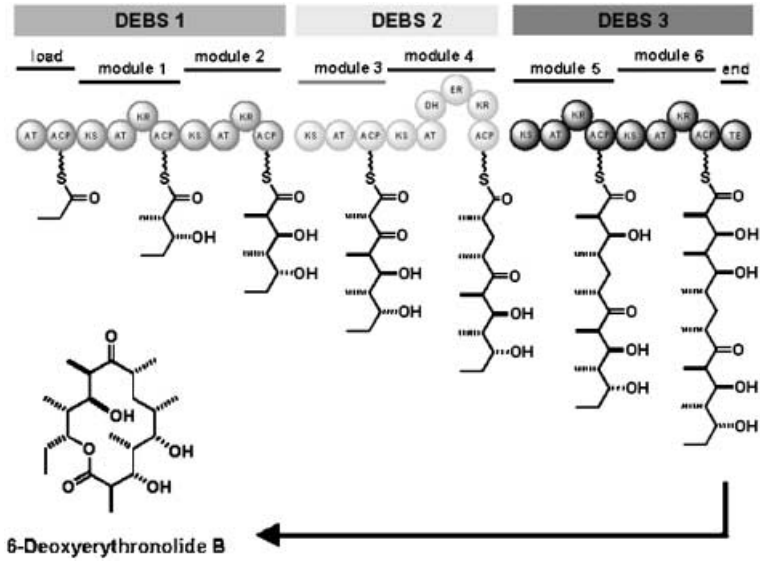

Figure 1. *The erythromycin-producing polyketide synthase 6-deoxyerythronolide B synthase (DEBS) is a hexamodular PKS incorporating seven acyl carrier protein (ACP) domains. Each ACP is post-translationally modified with a phosphopantetheinyl (P-pant) "swinging arm" (indicated by the zigzag lines). The terminal thiol of each P-pant group provides an attachment site for chain-extension intermediates onto the PKS, as shown.*

PPTase is incorporated as a domain at the carboxyl terminus of the α-subunit of fungal type I FASs, where it catalyzes the *in cis* autophosphopantetheinylation of the ACP at the amino terminus of the same polypeptide.[13]

Many organisms contain multiple P-pant-requiring pathways, and it was initially speculated that each pathway might have a dedicated PPTase.[8] Indeed, the yeast *Saccharomyces cerevisiae*, which contains three P-pant-dependent pathways, has three PPTases with mutually exclusive substrate specificities: PPT2 for the type II ACP of mitochondrial FAS, Lys5 for the type I PCP domain of the α-aminoadipate reductase Lys2, and a PPTase domain integrated into the type I FAS2.[13–15] However, the complete genome sequences of *S. coelicolor*[16] and *Streptomyces avermitilis*[17] have revealed that the number of CP-containing pathways in these bacteria significantly exceeds the number of PPTases. For example, *S. coelicolor* has a total of 22 secondary metabolic clusters (including eight PKSs, four NRPSs, and one hybrid PKS-NRPS), yet has only two Sfp-type PPTases.[18] In addition, most of the polyketide biosynthetic gene clusters identified so far do not contain a gene for their associated PPTase.[4] Taken together, these observations imply that it will be difficult to predict both the complement of PPTases for a given organism and their location within the genome.

We aimed to identify the PPTase(s) responsible for activating the erythromycin PKS in *Saccharopolyspora erythraea*, in order to ensure optimal overexpression of active synthase proteins in heterologous hosts such as *E. coli*. Expression of the C-terminal didomain ACP-thioesterase (TE) of DEBS 3,[19] as well as of the multienzyme subunit DEBS 3[20] in *E. coli*, showed that, although the proteins were judged to be correctly folded by several criteria, endogenous *E. coli* PPTases were not able to use them as substrates.[19, 20] Later, catalytically active unimodular and bimodular DEBS proteins were obtained by co-expression of the broad-specificity *B. subtilis* PPTase Sfp in *E. coli*.[21–23] However, this modification is reportedly incomplete,[21] consistent with the preference of Sfp for PCP over ACP substrates, at least in vitro.[4, 20] For optimal heterologous expression, and for in vitro studies of ACP/PPTase interaction, it would clearly be preferable to clone the gene for the cognate PPTase for the DEBS proteins. This gene does not lie in the erythromycin biosynthetic cluster, but we have now used the essentially complete genome sequence of *S. erythraea* to characterize the "PPTase-ome" (PPTase complement of the genome) of this strain. We show here that (only) one of the three PPTases of *S. erythraea* readily catalyzes the complete modification of DEBS-derived proteins expressed in *E. coli*.

Results

Identification of three PPTases in *Saccharopolyspora erythraea*

A search of a curated database of BLAST[22] hits from the *Saccharopolyspora erythraea* genome yielded three contigs containing genes with homology to both ACPS-type and Sfp-type PPTases. However, because the genome sequencing was only partially completed ($\approx$95%) when we initiated the project (as estimated by the number of gaps remaining), two of the PPTase genes were only represented by partial sequences. Therefore, inverse polymerase chain reaction (PCR)[23] was used to complete the sequences with genomic DNA from the *S. erythraea* strain JC2.[24]

Analysis of the predicted protein products of the three genes revealed that one is a putative ACPS-type PPTase (designated *S. erythraea* ACPS) and the other two are Sfp-type PPTases (designated SePptI and SePptII). The *acpS* gene starts with a GTG codon, which is preceded by a potential ribosome binding site (GGAGTG). The overall (76.5%) and third codon position (76.6%) G + C content and the codon usage of *acpS* are as expected for an actinomycete gene. The deduced protein sequence includes consensus residues identified by Lambalot et al.,[1] which participate in substrate binding and catalysis:[5, 6, 10] (V/I)G(V/I)D (motif P2) and (F/W)(S/C/T)XKE(A/S)hhK (h refers to a hydrophobic residue) (motif P3) separated from each other by 40–45 residues (Figure 2B). *S. erythraea* ACPS also shows significant similarities to a large number of putative ACPSs from actinomycetes and other species (Figure 3A),[25] including the *S. coelicolor* ACPS (47% identity, 58% similarity; accession number O86785), *S. avermitilis* ACPS (46% identity, 57% similarity; accession number Q82DL2),[17] and *Mycobacterium tuberculosis* ACPS (31% identity, 45% similarity; accession number O53228). The *S. erythraea* ACPS protein is predicted to contain 123 amino acids, which is typical for these enzymes.

Because the *sepptI* gene did not have a clear ribosome binding site, the position of the start codon was not obvious. Two possible ATG start codons were identified, yielding putative products of 285 and 246 amino acids (Figure 2C). The codon

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

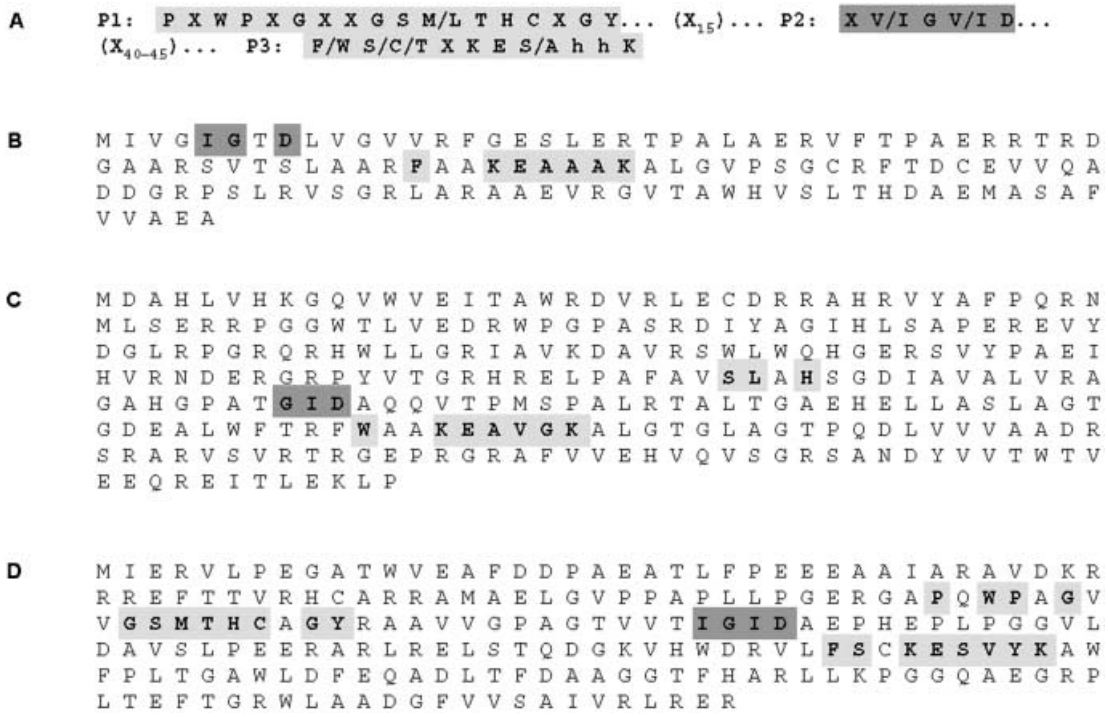

Figure 2. *Sequences of the* Saccharopolyspora erythraea *phosphopantetheinyl transferases (PPTases). A) Sequences of the PPTase consensus motifs, P1, P2, and P3. ACPS-type PPTases incorporate motifs P2 and P3, while all three motifs are present in the Sfp-type PPTases. B) Sequence of the* S. erythraea *ACPS. The conserved PPTase motifs are highlighted in gray. C) Sequence of the Sfp-type PPTase, SePptI. The first row shows residues that are only in the longer version of the protein. D) Sequence of the Sfp-type PPTase, SePptII.*

usage of either version of the gene is as expected (72.1% and 73.3% G + C content, respectively). From the typical length of Sfp-type PPTases (≈240 amino acids), however, it was thought that the 246 residue protein was the more likely candidate. While SePptI contains the most highly conserved residues of the consensus motifs P2 and P3, it is missing six amino acids from an additional motif, P1 (PXWPXGX$_2$GS(M/L)THCXGY), identified by Sánchez et al.,[4] which lies 15 amino acids upstream of P2 (Figure 2C). Nonetheless, it aligns well with a subset of the Sfp-type PPTases (Figure 3B), including *S. avermitilis* SAV7361 (45% identity, 54% similarity; accession number Q93HI0), *Streptomyces venezuelae* JadM (33% identity, 42% similarity; accession number Q9L7M4), and *S. coelicolor* SCO5883 (33% identity, 41% similarity; accession number O54146).

The gene for SePptII also begins with a GTG codon and is preceded by a putative ribosome binding site (GGAGAA). Its overall (74.5%) and third codon position (74.6%) G+C contents are as expected for an actinomycete gene. The predicted protein contains all three motifs, P1 – P3 (Figure 2D). SePptII aligns well with other Sfp-type PPTases in the database (Figure 3B),[23] including, *S. coelicolor* SCO6673 (54% identity, 63% similarity; accession number O88029), *S. avermitilis* PptA2 (54% identity, 62% similarity; accession number Q82MB2), and *Streptomyces verticillus* Svp (55% identity, 63% similarity; accession number Q9F0Q6). The SePptII protein is expected to contain 230 amino acids, within the typical size range for this type of PPTase.

Location of the PPTases in the *S. erythraea* genome

We were interested in determining the genetic context of the three PPTases in order to help identify the putative PPTase(s) for erythromycin biosynthesis. We therefore analyzed the curated database of *S. erythraea* for BLAST hits on the genes flanking the PPTases. The two genes upstream from the *acpS* within the contig are homologues of a lipase from an uncultured bacterium (accession number AF223645) and L-glutamine-D-fructose-6-phosphate amidotransferase from *Mycobacterium smegmatis* (accession number GLMS_MYCSM); the downstream genes do not match known sequences in the database. The ACPS of *S. coelicolor* (accession number SAV4964) is also located just downstream of an amidotransferase gene.

The *seppt1* gene is located within the same open reading frame (ORF) as genes with strong homology to PKS ketosynthase (KS) and acyltransferase (AT) domains; the KS domain is incomplete as the contig begins approximately 100 amino acids downstream of the KS start site. It therefore appears that SePptI forms the C terminus of a PKS multienzyme (Figure 4A); this PKS is as yet uncharacterized and is the third example of a modular PKS within *S. erythraea*. Consistent with this observation, the downstream genes within the contig include putative glycosyltransferases and *Streptomyces* sp. pathway-specific regulatory proteins. All three genes are highly similar to the corresponding domains in a putative modular polyketide synthase subunit (SAV7361) in *S. avermitilis* (the sequence of the integral PPTase is shown in Figure 3B). The SAV7361 subunit, which has the same

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

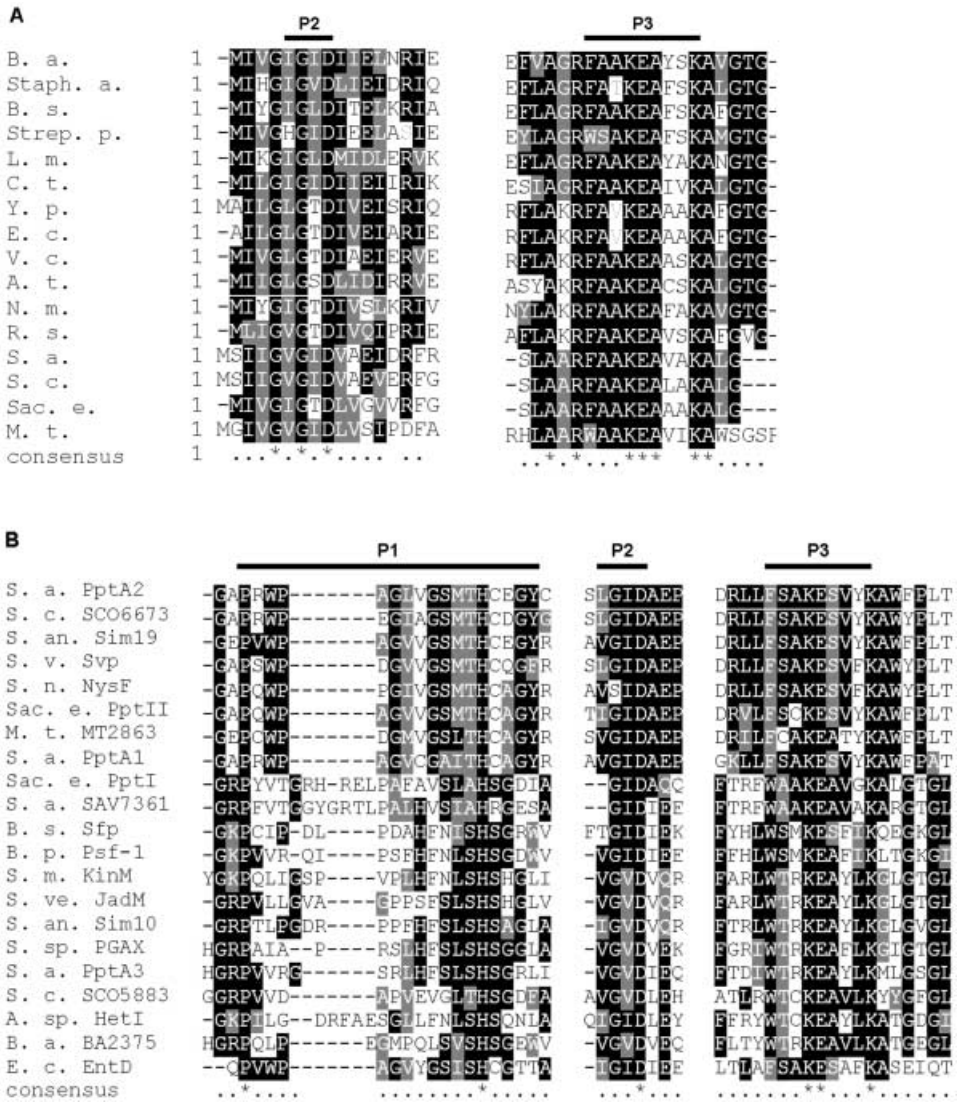

Figure 3. *Multiple alignment of PPTases in the regions of the consensus motifs. A) Alignment of ACPS-type PPTases. The consensus motifs P2 and P3 are indicated. Abbreviations are: B. a.,* Bacillus anthracis; *Staph. a.,* Staphylococcus aureus; *B. s.,* Bacillus subtilis; *Strep. p.,* Streptococcus pneumoniae; *L. m.,* Listeria monocytogenes; *C. t.,* Clostridium tetani; *Y. p.,* Yersinia pestis; *E. c.,* Escherichia coli; *V. c.,* Vibrio cholerae; *A. t.,* Agrobacterium tumefaciens; *N. m.,* Neisseria meningitidis; *R. s.,* Rickettsia sibirica; *S. a.,* Streptomyces avermitilis; *S. c.,* Streptomyces coelicolor; *Sac. e.,* Saccharopolyspora erythraea; *M. t.,* Mycobacterium tuberculosis. *B) Alignment of Sfp-type PPTases in the regions of the consensus motifs. The consensus motifs P1, P2, and P3 are indicated. Abbreviations (not given in A) are: S. an.,* Streptomyces antibioticus *Tu 6040; S. v.,* Streptomyces verticillus; *S. n.,* Streptomyces noursei; *B. p.,* Bacillus pumilus; *S. m.,* Streptomyces murayamaensis; *S. ve.,* Streptomyces venezuelae; *S. sp.,* Streptomyces sp. *PGA64; A. sp.,* Anabena *species.*

organization as the putative *S. erythraea* multienzyme, lies in the middle of an uncharacterized PKS cluster (pks1). The gene directly upstream of SAV7361 is a type II ACP, which may be the missing partner of the KS and AT domains; whether a similar arrangement occurs in *S. erythraea* remains to be determined. The location of SePptI (and the PPTase of SAV7361) within a multifunctional protein was unanticipated, because the only integral PPTases known to date are domains in the fungal type I FAS systems[13] which more closely resemble ACPS-type PPTases in length.[14] It would therefore appear that SePptI and the PPTase of SAV7361 are members of a fourth structural class of phosphopantetheinyl transferases.

The contig including *sepptII* contains only two genes, part of a transposase and *sepptII* itself; repeated efforts to sequence downstream of *sepptII* have failed. Interestingly, the transposase adjacent to *sepptII* is nearly identical (95% on the protein level) to the transposase (called IS1136) located between DEBS 1 and DEBS 2 in the erythromycin cluster.[26] The closest homologues to SePptII in *S. coelicolor* and *S. avermitilis* (SCO6673[18] and SAV1748,[27] respectively) are also located away from their target clusters.

Our analysis of the genes that are co-located with the three PPTases in *S. erythraea* pointed to SePptII as the most likely candidate for activation of the DEBS proteins.

Overproduction and purification of *S. erythraea* ACPS, SePptI, SePptII, and ACP-TE

In order to establish that SePptII was the PPTase that accomplished activation of DEBS in *S. erythraea*, we decided to express the three PPTases individually in *E. coli* for study in vitro. We chose the didomain ACP-TE from the C-terminal end of DEBS 3 as our model substrate.[19] The three PPTases were cloned into pET28b+ and pET29b+ for expression in *E. coli* BL21-CodonPlus® (RP) as both their N- and C-terminal His_6 tag fusion proteins. Analysis of total cell protein from IPTG-induced cells demonstrated that the yield of the N-terminal fusions was significantly greater than proteins which contained a C-terminal His_6 tag. In the case of SePptI (Figure 2C), both the longer (285 amino acid) and shorter (246 amino acid) versions of the protein were expressed, but as only the 246-residue SePptI was soluble, it was selected for further study. The didomain ACP-TE was expressed as a C-terminal His_6 tag fusion protein from pACP-TEHis[28] (this plasmid was constructed by cloning the gene for the ACP-TE didomain[19] into pT7–18His, a pT7–7-derived plasmid incorporating a C-terminal His_6 tag; although two of the C-terminal amino acids of the TE had been removed in construction of the

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

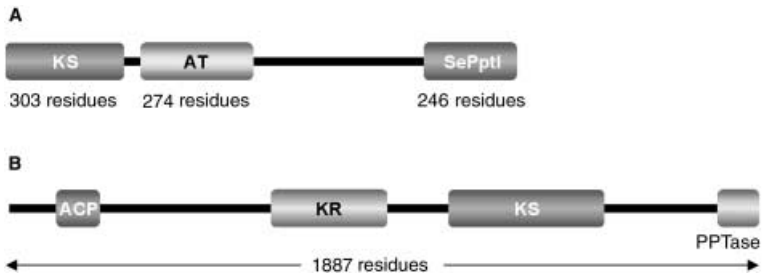

Figure 4. *Comparison of the organization of type I PKS and type I FAS multienzymes incorporating PPTases. A) Organization of the modular PKS subunit in* S. erythraea *including a C-terminal PPTase domain. The multienzyme incorporates ketosynthase (KS) and acyltransferase (AT) domains upstream of the PPTase. Domains and linkers are shown to scale. B) Organization of the yeast FAS subunit α. The FAS contains an N-terminal ACP domain followed by ketoreductase (KR) and KS activities, and terminating in the PPTase. Domain boundaries are approximate.*

plasmid, this deletion was not expected to affect the ability of the ACP domain to fold).

The PPTases and the ACP-TE were purified by Ni^{2+}-nitrolotriacetic acid affinity (Ni-NTA) chromatography. This step provided the ACPS, SePptI, and ACP-TE in both good yield and good purity (Figure 5); ACPS and SePptI could be purified further to

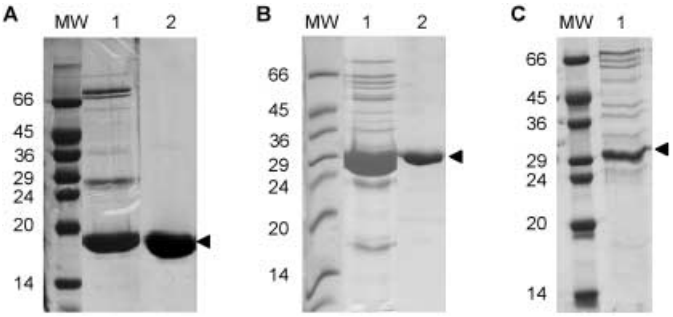

Figure 5. *Purification of* S. erythraea *ACPS, SePptI, and SePptII. A) Lane 1, purification of the ACPS after Ni-NTA affinity chromatography; Lane 2, after cation-exchange chromatography. His_6-tagged* S. erythraea *ACPS has a calculated molecular weight of 15.0 kDa. The positions of the protein markers are shown on the left. B) Lane 1, purification of SePptI after Ni-NTA affinity chromatography; Lane 2, after cation-exchange chromatography. His_6-tagged SePptI has a calculated molecular weight of 27.0 kDa. The positions of the protein markers are shown on the left. C) Lane 1, purification of SePptII after Ni-NTA affinity chromatography. His_6-tagged SePptII has a calculated molecular weight of 26.1 kDa. The positions of the protein markers are shown on the left.*

homogeneity by cation-exchange chromatography. However, the yields of SePptII from the Ni-NTA step were consistently low, despite exhaustive attempts to express the protein under a variety of conditions and by use of several different expression systems. Attempts to purify SePptII further by anion-exchange chromatography were also unsuccessful.

Phosphopantetheinylation in vitro by ACPS, SePptI, and SePptII

As SePptII was impure, only ACPS and SePptI were used initially in the in vitro assays. Incubations were carried out with 0.01 and 0.1 equiv of PPTase relative to ACP-TE, under the conditions reported by Lambalot et al. for EntD and Sfp:[1] Tris-HCl (pH 8.8, 75 mM), $MgCl_2$ (10 mM), DTT (25 mM), and CoASH (200 μM). Although the pH optimum of Sfp is 6.0, it remains active at pH 8.8.[1] Incubations were carried out for one hour at 37 °C and then quenched by the addition of 50 mM EDTA. The ACP-TE was then analyzed by liquid chromatography-mass spectrometry (LC-MS), with unmodified ACP-TE as a control; phosphopantetheinylation was expected to increase the molecular weight of ACP-TE by 339 Da. No formation of holoenzyme was observed by LC-MS in any assay with these enzymes.

On the basis of these results, all three PPTases were incubated with the ACP-TE in an alternative buffer (NaPi (50 mM), $MgCl_2$ (10 mM), DTT (5 mM), CoASH (1 mM))[4] and over a range of pH values (6.0 – 8.0). Assays were carried out with enzyme/substrate molar ratios of 0.1, 1.0, and 10 (SePptI could only be used at the two lower concentrations, as it became insoluble at 10 ×). ACP-TE was also incubated with Sfp under the same conditions (pH 6 only). The reactions were performed at 37 °C for six hours and then quenched by the addition of 50 mM EDTA. As before, the assays were analyzed by LC-MS.

This analysis revealed that the SePptII was able to modify the ACP-TE at all concentrations used and at all pHs tested (calculated: 39311 Da; found: 39304 ± 3 Da), as was Sfp. Again, ACPS and SePptI failed to modify the ACP-TE under these conditions. The assays with SePptII were repeated, but with enzyme/substrate molar ratios of 0.01 and 0.001, and at pH 7 only. As before, formation of holo-protein was complete within six hours, as judged by LC-MS. The Sfp enzyme has been reported to use CoASH analogues (desulfo, homocysteamine),[12] acyl-CoAs (benzoyl- and phenylacetyl),[12] and diketide-CoAs[29, 30] as substrates, and so it was of interest to see if SePptII could also utilize modified CoAs. The ACP-TE was therefore incubated with SePptII and a range of acyl-CoAs (acetyl, propionyl, *n*-butyryl, isobutyryl, valeryl, (2*RS*)-methylmalonyl), as well as desulfo-CoA. Only desulfo-CoA was found to be a substrate for SePptI (calculated: 39272 Da; found: 39280 Da); Svp, an Sfp-type PPTase from the bleomycin producer *Streptomyces verticillus* ATCC15003, has a similarly narrow substrate specificity, as it also failed to recognize acetyl-CoA.[4] Taken together, these assays seemed to confirm the expectation from sequence analysis that SePptII is responsible for activation of the DEBS proteins.

Although ACPS and SePptI could not modify ACP-TE under all conditions tested, they were nonetheless capable of P-pant transfer to a type II ACP. ACPS and SePptI (along with SePptII and Sfp) were incubated at fourfold excess with the ACP in buffer (Tris-HCl (pH 7.5, 50 mM), $MgCl_2$ (10 mM), DTT (2 mM), CoASH (1 mM)), at 30 °C for four hours. By LC-MS analysis, the extent of phosphopantetheinylation in all cases was approximately 50%. This experiment demonstrated that ACPS and SePptI can indeed function as PPTases, and so their inability to modify the DEBS ACP-TE probably does reflect their substrate specificity. It also showed that SePptII can recognize a type II ACP domain in addition to a type I ACP.

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

In vivo phosphopantetheinylation by SePptII

Because the partially purified SePptII was highly active against DEBS proteins in vitro, we decided to investigate whether it could also modify DEBS proteins in vivo in *E. coli*. To do this, *sepptII* was introduced into the pT7-based ACP-TE expression vector, pACP-TEHis: *sepptII* was excised from the pET28b+ based expression vector pKJW152 along with the upstream region containing the T7 promoter, lac operator, and ribosome binding site, and the downstream region containing the T7 terminator, and ligated into pACP-TEHis.[31] In this way, addition of IPTG to the cell medium was expected to induce expression of both the substrate ACP-TE and the PPTase. The ACP-TE was expressed and purified as described previously. Although SePptII was also His_6-tagged, the low levels of PPTase expression did not interfere with the purification of the ACP-TE. Analysis of the ACP-TE by LC-MS showed that the ACP was completely modified (calculated: 39311 Da; found: 39308 ± 1 Da) (Figure 6).

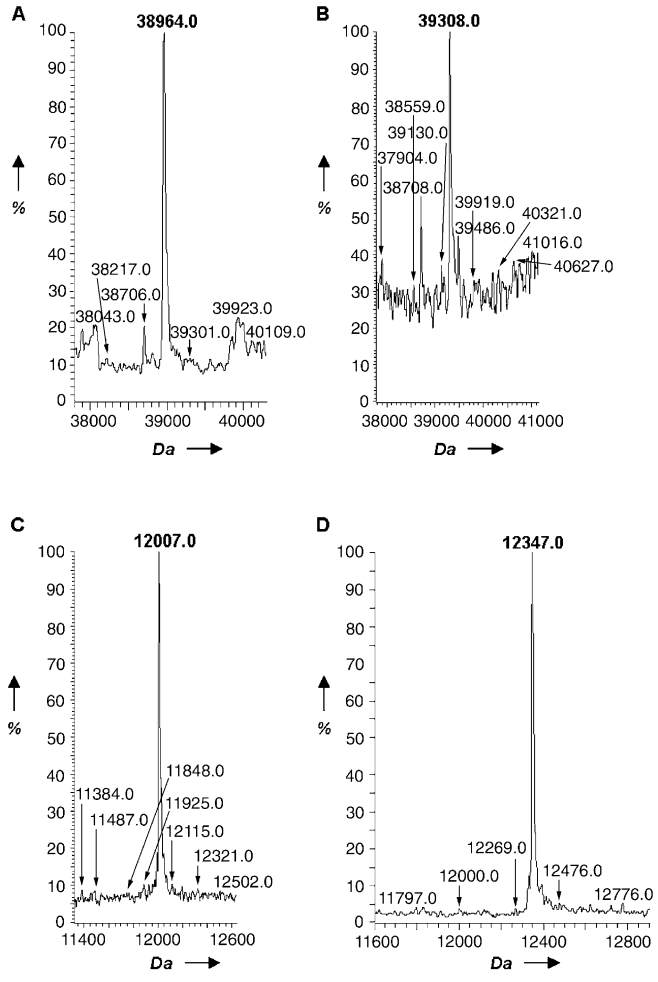

Figure 6. *LC-MS analysis of in vivo phosphopantetheinylation by SePptII. A) Mass spectrum of the didomain His_6-tagged ACP-TE expressed in the absence of SePptII (calculated size, 38972 Da). B) Mass spectrum of the didomain His_6-tagged ACP-TE co-expressed with SePptII (calculated size, 39311 Da). C) Mass spectrum of apo ACP_5 obtained by limited trypsinolysis of DEBS 3 (calculated size, 12005 Da). D) Mass spectrum of ACP_5 obtained by limited trypsinolysis of DEBS 3 after co-expression with SePptII (calculated size, 12344 Da).*

It was also critical to demonstrate that SePptII could modify ACP domains in the context of intact PKS subunits. Our target was DEBS 3, which contains two ACP domains, ACP_5 and ACP_6 (the latter being the ACP of the didomain ACP-TE). We therefore constructed a co-expression vector for DEBS 3 and SePptII: SePptII was amplified from pKJW152, and introduced into a pET28b+-based expression vector for DEBS 3, pAPS11; again, the genes upstream and downstream of *sepptII* were also included. Here, both the DEBS 3 and the SePptII were expected to be His_6-tagged. The resulting plasmid, pKJW191R, was transformed into *E. coli* BL21-CodonPlus®, and the DEBS 3 protein was purified by Ni-NTA affinity chromatography followed by anion-exchange chromatography. The yield of purified DEBS 3 was $>2\ mg\,L^{-1}$, which compares favorably to expression levels in the absence of the PPTase gene (typically $1-1.5\ mg\,L^{-1}$).

In order to determine whether P-pant transfer had occurred, DEBS 3 was briefly subjected to limited proteolysis with trypsin. Trypsinolysis of DEBS 3 had previously been found (by LC-MS) to release ACP_5 as a distinct domain, which could be separated from the other fragments in the mixture. Analysis of the digest by LC-MS showed that ACP_5 had been quantitatively modified by SePptII (calculated: 12344 Da; found: 12347 ± 1 Da) (Figure 6). Although it was not possible to analyze ACP_6 for technical reasons, this experiment provided very good evidence of the ability of SePptII to recognize ACPs within the context of multidomain PKS subunits.

Discussion

PPTase-catalyzed phosphopantetheinyl transfer is a critical step in the post-translational activation of multienzymes involved in both primary and secondary metabolism. While some PPTases exhibit very broad specificities, the three *E. coli* enzymes (ACPS, EntD, and YhhU) are unable to modify a range of heterologous ACP and PCP domains when expressed in vivo.[19–21] Until recently, this deficiency significantly impeded in vitro research into polyketide and non-ribosomal polypeptide biosynthesis, which depends on the availability of large quantities of functional proteins. Although the *B. subtilis* PPTase Sfp can function in co-expression experiments in *E. coli*,[21, 29, 30, 32] holo-protein formation by Sfp is incomplete.[21] This inefficiency is probably due to the fact that Sfp prefers PCP substrates to ACPs and ArCPs, at least in vitro.[4, 12] We were therefore interested in identifying the PPTase from the erythromycin producer *S. erythraea* that specifically activates the DEBS PKS, in order to evaluate its utility in such co-expression experiments.

As they are small and poorly conserved domains, the PPTases are a challenging target for standard genome screening methods. In addition, even if the location of primary and secondary metabolic clusters is known, there is no guarantee that the associated PPTase genes will be co-located.[4] Sánchez et al.[4] have recently published a method for cloning PPTase genes from polyketide and non-ribosomal peptide-producing organisms. In this approach, primers are created by the use of CODEHOP (Consensus-Degenerate Hybrid Oligonucleotide Primer),[33] and the resulting PCR products are then used as probes to screen a genomic library by colony hybridization. However,

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

because the primers are designed around the conserved P1 and P3 motifs, this strategy is inherently unsuitable for identifying ACPS-type PPTases, which lack P1; attempts to amplify PPTases with degenerate primers designed against the P2 and P3 motifs were unsuccessful.[4] In addition, even when care is taken to detect weakly hybridizing signals during screening, it is unclear whether this technique can confidently identify all PPTases within a given organism.

The total sequencing of the *S. erythraea* genome, nearing completion in this laboratory, provided the candidate PPTase genes for the present study. Although the genes have not yet been annotated, the sequence data have been analyzed by using a program called "blast analyse",[34] which identifies only strong hits to proteins of known or predicted function in the SPTR database. This analysis yielded a "curated" list of proteins, which was then searched with keywords relevant to PPTases, "phosphopant" and "holo/acyl carrier protein synthase". Searching the genome sequence in this way yielded three candidate PPTases, although the sequences of two of the genes had to be completed by inverse PCR. In order to confirm that these were the only transferase sequences in *S. erythraea*, the genome was used in BLAST[22] searches with sequences both of an ACPS-type (*S. coelicolor* ACPS) and of an Sfp-type PPTase (PptA2 from *S. avermitilis*), as well as with the sequences of the three putative *S. erythraea* PPTases. These analyses and more recent searches on the essentially complete genome have revealed no additional sequences with homology to PPTases, and so these three genes appear to represent the "PPTase-ome" of *S. erythraea*.

The *S. erythraea* genome contains one ACPS-type PPTase and two Sfp-type PPTases. Although we did not expect to find a strict correlation between the location of the PPTases and their specificity, we nonetheless analyzed the genomic context of the genes for possible insight into their function. The *acpS* gene, like its counterparts in *S. coelicolor* and *S. avermitilis*, does not seem to be located near the gene for any obvious target protein. Therefore, its predicted role in primary metabolism will need to be demonstrated with studies in vitro.

Unexpectedly, we found that the Sfp-type PPTase SePptI is actually integrated into the C terminus of an unidentified PKS subunit, downstream from a KS and an AT domain. In the fungal type I FAS systems which also possess an integral PPTase, the transferase domain instead lies downstream of its target ACP, as well as a ketoreductase (KR) and a KS domain (Figure 4B).[13] Additionally, the fungal type I PPTases are more similar to the ACPS-type enzymes in length.[4] The novel organization of this erythromycin PKS is shared by a modular PKS in *S. avermitilis*,[27] which suggests that these two domains represent a distinct structural class of PPTases. The *S. erythraea* SePptI and *S. avermitilis* PPTase align most closely with Sfp-type PPTases such as JadM from *S. venezuelae* and Sfp, both of which are discrete enzymes. It is interesting that this sub-group of Sfp-type PPTases consistently lacks several of the conserved residues identified in the P1 motif (Figure 3B), although those most important in substrate binding and catalysis are present.[5, 6, 10] Sánchez et al. have speculated that the organization of the fungal FAS arose from a gene fusion event with an ACPS-type PPTase, and perhaps a similar event with one of these Sfp-type PPTases generated this new type of PKS. It will be interesting to determine whether the integral PPTase within the homodimeric structure of type I PKS[35] also forms a homodimer, since Sfp-type PPTases are normally monomeric.[5]

The short contig including *sepptII* contains only one other gene. This sequence has $>95\%$ homology (protein level) with the transposase gene located within the DEBS cluster between the ORFs *eryAI* and *eryAII*, encoding for DEBS 1 and DEBS 2, respectively.[36] This striking similarity raises the speculative possibility that *sepptII* was originally located within the cluster, but was moved elsewhere in the genome during a recombination event; the sequence adjacent to transposons is known to be a recombination "hot-spot".[37] Our efforts to purify SePptII and to demonstrate its activity in vitro, however, were hampered by poor expression levels of the PPTase in *E. coli*, despite attempts with many different conditions and expression systems, and despite alteration of the initial codons in the gene to reflect the *E. coli* codon preference. However, as $\approx 60\%$ pure protein, SePptII rapidly and efficiently catalyzed the formation of ACP-TE holo-protein over a range of pH values (6–8) and was also able to use the modified substrate, desulfo-CoA. In contrast, ACPS and SePptI were inactive under all conditions tested.

Nonetheless, it was critical to show that SePptII could also function as a PPTase in vivo in *E. coli* with DEBS proteins as substrates. Co-expression plasmids were therefore constructed for SePptII and the didomain ACP-TE, as well as the entire PKS subunit, DEBS 3, from which ACP-TE is derived. Efforts to express these two proteins in the absence of SePptII had yielded only apo-ACPs.[19, 20] Analysis of ACP-TE and the ACP_5 from trypsinolyzed DEBS 3 demonstrated complete holoation of the ACP domains, in contrast to the incomplete phosphopantetheinyl transfer achieved with Sfp.[21] Co-expression of SePptII with DEBS 3 had no apparent effect on the yield of purified DEBS protein. The yield of DEBS 3 is also high relative to other bimodular DEBS proteins that have been expressed with Sfp in *E. coli* (e.g., module 1 + module 5-TE (0.2 mg mL^{-1}) and M1 M6-TE (1 mg mL^{-1})).[29] The low expression levels of SePptII in *E. coli* are apparently compensated for by its high catalytic efficiency.

The *S. erythraea* genome contains four uncharacterized clusters for non-ribosomal polypeptide biosynthesis, and three modular PKSs, all of which require an associated PPTase for activity. The *S. erythraea* ACPS is likely to be specific for ACPs of primary metabolism; SePptI is expected to activate the PKS subunits in the cluster to which it belongs. Therefore, SePptII is the best candidate for a "promiscuous" PPTase in this bacterium, capable of modifying not only type I PKS ACP domains, but also NRPS PCP activities. The ability of SePptII to achieve complete holoprotein formation of intact PKS multienzymes in *E. coli* without substantial effects on yield due to overexpression of the PPTase itself makes SePptII a very attractive alternative to Sfp for co-expression experiments in *E. coli*.

Experimental Section

General methods and materials: *E. coli* DH10B was grown on 2 × TY medium and *E. coli* BL21-CodonPlus® (RP) (Stratagene) was grown on

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

2 × TY or LB media. Antibiotics were used at the following concentrations: ampicillin (100 μg mL^{-1}) and kanamycin (50 μg mL^{-1}) for *E. coli* DH10B, and ampicillin (50 μg mL^{-1}), kanamycin (30 μg mL^{-1}), and chloramphenicol (34 μg mL^{-1}) for *E. coli* BL21-CodonPlus. For transformation and plasmid preparation from *E. coli*, standard protocols were used.[38] Restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase, and DNA polymerase I large (Klenow) fragment were obtained from New England Biolabs. Reagents used for protein purification were purchased from Roche Molecular Biochemicals or Sigma.

Recombinant DNA methods: Purification of DNA fragments amplified by polymerase chain reaction (PCR) was performed with QIAEX II (QIAGEN) or Nucleospin® Extract kits (Machery – Nagel). PCR reactions were carried out with *Pwo* (Roche) or *Pfu* (Stratagene) DNA polymerases as described by the enzyme suppliers. DNA sequencing was performed by the DNA Sequencing Facility of the Department of Biochemistry, University of Cambridge. Oligonucleotides were purchased from QIAGEN Operon, and the expression vectors pET28b + and pET29b + were purchased from Novagen. The plasmid pACP-TEHis is a derivative of pERX1,[19] and was a kind gift from T. Schwecke.[28]

Identification of PPTase genes in *Saccharopolyspora erythraea*: The genome sequence of *S. erythraea* (> 95% complete) was used in BLAST (basic local alignment search tool)[22] searches against the SPTR (Swissprot + TrEMBL) protein databases. The results were analyzed by using "blast analyse",[34] a program which selects only strong hits to proteins of known or predicted function. This processed (curated) list was then searched with keywords related to PPTases ("phosphopant" and "holo/acyl carrier protein synthase") effectively selecting all hits in the *S. erythraea* genome with similarity to PPTase sequences in the current SPTR. This analysis yielded three contigs containing genes with homology to PPTases; however, at that stage only one of the PPTase sequences was complete. In addition, the genome sequence was BLAST searched with the *S. coelicolor* ACPS (accession number O86785) and the *S. avermitilis* PptA2 (Sfp-type PPTase; accession number Q82MB2) sequences, in order to verify the matches found by using the above technique, and also with the three *S. erythraea* PPTase sequences in order to confirm that there were no other transferase genes in the bacterium. The nucleotide sequences reported in this paper have been submitted to the GenBankTM/EBI Data Bank with accession numbers AY430049 for ACPS, AY430050 for SePptI, and AY430051 for SePptII.

Sequencing of the genes for *S. erythraea* ACPS by inverse PCR:[23] Genomic DNA from *S. erythraea* JC2[24] was digested with *Nar*I, *Sac*I, *Sac*II, *Sal*I, or *Xmn*I. The digested DNA was analyzed by agarose gel electrophoresis, and fragments in the size range ≈0.3 – 1.6 kb were excised and purified. The fragments were then ligated at high dilution in order to circularize the DNA. The gene for the ACPS was then successfully amplified from the *Sac*I-digested DNA by inverse PCR by use of primers located within the known sequence (5′-cccccgcaccggcgggatcaccgccgacac-3′ and 5′-ccggtggctcagccagactgtaggccgcacaacc-3′) to yield a product of ≈750 bp. The PCR product was then sequenced. The gene was then re-amplified (5′-gccgagatcgcgcgtgcgcgcggctacga-3′ (forward) and 5′-ggcggttgtgcggcctacagtctggctg-3′ (reverse)), phosphorylated with T4 polynucleotide kinase, and cloned into *Sma*I-cut pUC18, and the sequence was confirmed.

Sequencing of the genes for SePptII by inverse PCR:[23] Genomic DNA from *S. erythraea* JC2 was digested with *Nar*I, *Sac*I, *Sal*I, or *Xmn*I. The digested DNA was analyzed by agarose gel electrophoresis, and fragments in the size range ≈1 – 4 kb were excised and purified. The fragments were then ligated at high dilution in order to circularize the DNA. The gene for SePptII was successfully amplified from the *Sal*I-digested DNA by inverse PCR by use of primers located within the known sequence (5′-cggtgcgtggtcaccgaggagcactcgggatcggct-3′ and 5′-gtgcttactggtgatcggcaatcgtgtaaaatga-3′) to yield a product of ≈1.5 kb. The PCR product was then sequenced. The gene was then re-amplified (5′-cggtgcgtggtcaccgaggagcactcgggatcggct-3′ (f) and 5′-ggacgggcggccgccacgcgggcgcccgttcggcgta-3′ (r)), phosphorylated with T4 polynucleotide kinase, and cloned into *Sma*I-cut pUC18, and the sequence was confirmed.

Construction of expression plasmids for *S. erythraea* ACPS: All plasmids used are summarized in Table 1. *S. erythraea* ACPS was amplified from JC2 genomic DNA, with introduction of the restriction sites *Nde*I and *Hind*III (5′-tggagtgagcgg*catatg*atcgtcggcatcggca-3′ (f) and 5′-cgac*aagctt*tcaggcttcggccaccacgaacgcc-3′ (r) or 5′-cac-g*aagctt*ggcttcggccaccacgaacgccga-3′ (r)). The PCR products were phosphorylated with T4 polynucleotide kinase, and were then cloned into *Sma*I-cut pUC18. The sequences of the PCR products were confirmed by sequencing. The PCR products were then cloned into *Nde*I/*Hin*dIII-cut pET28b + and pET29b + to yield pKJW141 and pKJW146, respectively.

Construction of expression plasmids for SePptI: Expression constructs for two versions of SePptI (Figure 2) were created. The 285 amino acid version was amplified from JC2 genomic DNA, with introduction of the restriction sites *Nde*I and *Hin*dIII (5′-tccacgctg*catatg*gacgcccatctggtccacaa-3′ (f) and 5′-ttgg*aagctt*tggcagcttctccaaggtgatctcccgt-3′ (r) or 5′-gggt*aagctt*tcatggcagcttctccaaggtgatctccc-3′ (r)). The PCR products were phosphorylated with T4 polynucleotide kinase, and then cloned into *Sma*I-cut pUC18. The sequences of the PCR products were confirmed by sequencing. The PCR products were then cloned into *Nde*I/*Hin*dIII-cut pET29b + and

Table 1. *Plasmids used in this study.*

Plasmid name	Relevant characteristics	Origin or reference
pKJW127	pET29b + -based expression plasmid for SePptI-His_6 (long version)	This work
pKJW130	pET29b + -based expression plasmid for SePptI-His_6 (short version)	This work
pKJW133	pET28b + -based expression plasmid for His_6-SePptI (long version)	This work
pKJW136	pET28b + -based expression plasmid for His_6-SePptI (short version)	This work
pKJW141	pET28b + -based expression plasmid for His_6-ACPS	This work
pKJW146	pET29b + -based expression plasmid for ACPS-His_6	This work
pKJW149	pET29b + -based expression plasmid for SePptII-His_6	This work
pKJW152	pET28b + -based expression plasmid for His_6-SePptII	This work
pKJW177	Co-expression plasmid for ACP-TE and SePptII	This work
pKJW191R	Co-expression plasmid for DEBS 3 and SePptII	This work
pACP-TEHis	pT7-based expression plasmid for ACP-TE-His_6	Unpublished, gift from T. Schwecke
pAPS11	pET28b + -based expression plasmid for His_6-DEBS 3	This work

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

pET28b+ to yield pKJW127 and pKJW133, respectively. The 246 amino acid version was amplified from JC2 genomic DNA, with introduction of the restriction sites *Nde*I and *Hin*dIII (5′-ccgcagcgg-*catatg*ctctctgagcgcaggcccggtggc-3′ (f) and 5′-ttgg*aagctt*tgg-cagcttctccaaggtgatctcccgt-3′ (r) or 5′-gggt*aagctt*tcatggcagcttctc-caaggtgatctccc-3′ (r)). The PCR products were phosphorylated with T4 polynucleotide kinase, and were then cloned into *Sma*I-cut pUC18. The sequences of the PCR products were confirmed by sequencing. The PCR products were then cloned into *Nde*I/*Hin*dIII-cut pET29b+ and pET28b+ to yield pKJW130 and pKJW136, respectively.

Construction of expression plasmids for SePptII: SePptII was amplified from JC2 genomic DNA, with introduction of the restriction sites *Nde*I and *Hin*dIII (5′-ataggagaa*catatg*atcgagcgagtgctgcccga-3′ (f) and 5′-gtat*aagctt*ccgctcccgcagccgcacgatggcgct-3′ (r) or 5′-ggcg*aagctt*tcaccgctcccgcagccgcacgat-3′ (r)). The PCR products were phosphorylated with T4 polynucleotide kinase, and were then cloned into *Sma*I-cut pUC18. The sequences of the PCR products were confirmed by sequencing. The PCR products were then cloned into *Nde*I/*Hin*dIII-cut pET29b+ and pET28b+ to yield pKJW149 and pKJW152, respectively.

Construction of a co-expression plasmid for SePptII and ACP-TE: Plasmid pKJW152 was digested with *BspE*I, and was then end-filled by use of DNA polymerase I (Klenow). The resulting linear DNA was cut with *Bgl*II and ligated into pACP-TEHis digested with *Bgl*II and *BsaA*I to yield pKJW177.

Construction of a co-expression plasmid for SePptII and DEBS 3: SePptII (along with the T7 promoter, *lac* operator, and T7 terminator sequences) were amplified from pKJW152, with introduction of *Sph*I restriction sites (5′-gaggatcg*gcatgc*cgatcccgcgaaatta-3′ (f) and 5′-ttcgccaa*gcatgc*tatagttcctcctttca-3′ (r)) at both ends. The PCR fragment was ligated into *Sma*I-cut pUC18, and the sequence was confirmed by sequencing. The PPTase gene was then ligated into *Sph*I-cut pAPS11 (DEBS 3 in pET28b+) to yield pKJW191R (in which the genes for the SePptII and DEBS 3 are in opposite orientations).

Screening for overproduction of PPTases: Plasmids pKJW127, 130, 133, 136, 141, 146, 149, and 152 were used to transform *E. coli* BL21-Codon Plus cells. Cultures of all transformants (35–50 mL, LB) were grown at 30 °C to an OD_{600} of >0.6, and induced with 0.1 mM 1-thio-β-D-galactoside (IPTG). The cells were incubated overnight at 18 °C. Sodium dodecyl sulfate (SDS, 5%, 50 μL) was added to the cell pellet from 1 mL of culture, the pellet was boiled for 15 min, and the solution was then centrifuged for 1 h. The total cell proteins were then analyzed by SDS-PAGE.

Expression and purification of PPTases: Plasmids pKJW136, 141, and 152 were used to transform *E. coli* BL21-Codon Plus cells. Cultures of *E. coli* pKJW136 and pKJW141 (500 mL, LB) were grown at 30 °C to an OD_{600} of >0.6, induced with 0.1 mM IPTG, and grown overnight at 14–18 °C. Cultures of *E. coli* pKJW152 (6 × 50 mL, LB) were grown at 37 °C to an OD_{600} of >0.6; half of the cultures were induced with 0.1 mM (IPTG) and grown at 37 °C for 6 h, while the remainder were induced with 1 mM IPTG and grown overnight at 14–18 °C. All cells were lysed by sonication in buffer A (NaPi (pH 8.0, 50 mM), NaCl (300 mM), glycerol (20%)), containing imidazole (5 mM), RNAase, DNAase, and protease inhibitor tablet (Roche)), and the lysates were clarified by centrifugation. The lysates were then incubated with 50% Ni^{2+}-nitrolotriacetic acid (Ni-NTA) resin for 1 h at 4 °C. The resin was washed multiple times with 2 × resin volumes of buffer A containing 10 mM and then 20 mM imidazole, and the proteins were eluted in buffer A containing 50 mM and 250 mM imidazole. This purification step afforded the proteins at the following purities (as estimated from Coomassie-stained SDS-PAGE gels): 136 (90–95%), 141 (95%), and 152 (60%).

The ACPS and SePptI could be further purified on a HiTrap SP cation column (1 mL, Amersham Pharmacia). The ACPS was exchanged into buffer B (KOH (10 mM), DTT (1 mM), EDTA (1 mM), glycerol (20%), pH adjusted to 6.7 with solid 2-(*N*-morpholino)ethanesulfonic acid (MES)) on a PD-10 column (Amersham), loaded onto the HiTrap SP column in buffer B, and eluted with a gradient to 100% buffer C (buffer B + NaCl (1 M)). The SePptI was exchanged into buffer D (KOH (10 mM), DTT (1 mM), EDTA (1 mM), glycerol (20%), pH adjusted to 7.7 with solid *N*-(tris(hydroxymethyl)methyl)-2-aminoethanesulfonic acid (TES)) on a PD-10 column, loaded onto the HiTrap SP column in buffer D, and eluted with a gradient to 40% buffer E (buffer D + NaCl (1 M)).

Expression and purification of ACP-TE: Plasmids pACP-TEHis and pKJW177 were used to transform *E. coli* BL21-Codon Plus cells. Cultures of *E. coli* pACP-TEHis (500 mL, LB) and pKJW177 (2 × 50 mL, LB) were grown at 37 °C to an OD_{600} of >0.6, induced with 1.0 mM IPTG, and grown for 5–5.5 h at 26–30 °C. ACP-TE was purified by Ni-NTA chromatography as for the PPTases.

Expression and purification of DEBS 3: Plasmid pKJW191R was used to transform *E. coli* BL21-Codon Plus cells. Cultures of *E. coli* pKJW191R (8 × 500 mL, LB) were grown at 30 °C to an OD_{600} of 1.2, induced with 1.0 mM IPTG, and grown for 4 h at 30 °C. The cells were lysed by sonication in buffer A containing 1 mM imidazole, RNAase, DNAase, protease inhibitor tablet (Roche), and lysozyme (1 mg mL^{-1}), and the lysate was clarified by centrifugation. Ammonium sulfate was then added to 55% saturation, and the precipitated proteins collected by centrifugation. The pellet was resuspended in buffer A, and applied to a HiTrap chelating column (5 mL, Amersham), pre-equilibrated with Ni_2SO_4 as per manufacturer's instructions. The column was washed with 2.5 column volumes of buffer A, and was then washed with 22 column volumes of buffer A containing 5 mM imidazole. The protein was then eluted with buffer A containing 250 mM imidazole. The fractions containing DEBS 3 were pooled and concentrated to 2.5 mL (Amicon® Ultra, 100 kDa MW cut-off), and then exchanged into buffer F (NaPi (pH 7.0, 50 mM), DTT (1 mM), EDTA (1 mM), glycerol (10%)) on a PD-10 column. The desalted protein was then loaded onto a HiTrap Q column (5 mL, Amersham) in buffer F. The column was washed with 13 volumes of 23% buffer G (buffer F + NaCl (1 M)), and the DEBS 3 was then eluted with a gradient to 100% buffer G. The purification yielded 8.9 mg of DEBS 3 (>2 mg L^{-1}) as estimated by Bradford assay.

In vitro analysis of PPTases with ACP-TE: In initial experiments, ACP-TE (65 μg) was incubated with 0.01 and 0.1 molar equivalents of ACPS or of SePptI in assay buffer (Tris-HCl (pH 8.8, 75 mM), $MgCl_2$ (10 mM), DTT (5 mM)) containing CoASH (200 μM) for 1 h at 37 °C (total reaction volume 1 mL) and then quenched with 50 mM EDTA. ACP-TE (6.5 μg) was alternatively incubated with 0.1, 1.0, and 10 equivalents of ACPS, SePptI, and SePptII (amount estimated from Coomassie-stained gel) in assay buffer (50 mM NaPi (pH 6, 6.5, 7, 7.5, and 8), $MgCl_2$ (10 mM), DTT (5 mM)) containing CoASH (1 mM) for 6 h at 37 °C (total reaction volume 100 μL) and quenched with 50 mM EDTA. Sfp (the kind gift of Y. Li[39]) was incubated under the same conditions at pH 6 only. The ACP-TE was also incubated with SePptII and acyl-CoAs (acetyl, propionyl, *n*/*i*-butyryl, valeryl, (2*RS*)-methylmalonyl) and desulfo-CoA (all at 250 μM)),[12] under identical conditions. In all cases, the ACP-TE was analyzed by liquid chromatography-mass spectrometry (LC-MS) on a Vydac Protein C4 column (30–65% acetonitrile/0.1% trifluoroacetic acid over 35 min) on a ThermoFinnigan LCQ.

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

In vitro analysis of PPTases with a type II ACP:[39] The ACP (20 – 40 μg) was incubated with 4 equivalents of ACPS, SePptI, SePptII, and Sfp in buffer (Tris-HCl (pH 7.5, 50 mM), $MgCl_2$ (10 mM), DTT (2 mM)) containing CoASH (1 mM) for 4 h at 30 °C, and then analyzed by LC-MS.

Limited trypsinolysis of DEBS 3 and analysis by LC-MS: DEBS 3 (350 μg) was incubated with 0.013 (w/w) equivalents of trypsin in buffer (KPi (pH 7.0, 50 mM), EDTA (1 mM), DTT (1 mM), glycerol (10%)) in a total volume of 100 μl. The digest was carried out at 30 °C for 5 min, and the reaction mixture was then immediately injected onto a Vydac Protein C4 column, pre-equilibrated with 35% acetonitrile/0.1% trifluoroacetic acid. The ACP_5 was then separated from the other components by use of a gradient of 35 – 55% acetonitrile/0.1% trifluoroacetic acid over 40 min (flow rate 0.7 mL min^{-1}).

Acknowledgements

We thank William P. Cawthorn for help in cloning pKJW191R, Torsten Schwecke for the kind gift of pACP-TEHis, Yanyan Li for purified Sfp and for carrying out assays with the type II ACP, and John Lester and Tatiana Mironenko for sharing unpublished data from the S. erythraea sequencing project. This work was supported by a grant from the BBSRC. K.J.W. is a Royal Society Dorothy Hodgkin Fellow.

Keywords: biosynthesis · erythromycin · polyketides · *Saccharopolyspora erythraea* · transferases

[1] R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chem. Biol.* **1996**, *3*, 923 – 936.
[2] H. D. Mootz, R. Finking, M. A. Marahiel, *J. Biol. Chem.* **2001**, *276*, 37289 – 37298.
[3] P. F. Leadlay, J. Staunton, J. F. Aparicio, D. J. Bevitt, P. Caffrey, J. Cortés, A. Marsden, G. A. Roberts, *Biochem. Soc. Trans.* **1993**, *21*, 218 – 222.
[4] C. Sánchez, L. Du, D. J. Edwards, M. D. Toney, B. Shen, *Chem. Biol.* **2001**, *8*, 725 – 738.
[5] N. Y. Chirgadze, S. L. Briggs, K. A. McAllister, A. S. Fischl, G. Zhao, *EMBO J.* **2000**, *19*, 5281 – 5287.
[6] K. D. Parris, L. Lin, A. Tam, R. Mathew, J. Hixon, M. Stahl, C. C. Fritz, J. Seehra, W. S. Somers, *Struct. Fold Des.* **2000**, *8*, 883 – 895.
[7] A. M. Gehring, R. H. Lambalot, K. W. Vogel, D. G. Drueckhammer, C. T. Walsh, *Chem. Biol.* **1997**, *4*, 17 – 24.
[8] C. T. Walsh, A. M. Gehring, P. H. Weinreb, L. E. N. Quadri, R. S. Flugel, *Curr. Opin. Chem. Biol.* **1997**, *1*, 309 – 315.
[9] R. J. Cox, J. Crosby, O. Daltrop, F. Glod, M. E. Jarzabek, T. P. Nicholson, M. Reed, T. J. Simpson, L. H. Smith, F. Soulas, A. E. Szafranska, J. Westcott, *J. Chem. Soc. Perkin Trans.* **2002**, *1*, 1644 – 1649.
[10] L. Reuter, M. R. Mofid, M. A. Marahiel, R. Ficner, *EMBO J.* **1999**, *18*, 6823 – 6831.
[11] R. Finking, J. Solsbacher, D. Konz, M. Schobert, A. Schäfer, D. Jahn, M. A. Marahiel, *J. Biol. Chem.* **2002**, *277*, 50293 – 50302.
[12] L. E. N. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry* **1998**, *37*, 1585 – 1595.
[13] F. FichtIscherer, C. Wellein, M. Mittag, E. Schweizer, *Eur. J. Biochem.* **2000**, *267*, 2666 – 2671.
[14] D. E. Ehmann, A. M. Gehring, C. T. Walsh, *Biochemistry* **1999**, *38*, 6171 – 6177.
[15] H.-P. Stuible, S. Meier, C. Wagner, E. Hannappel, E. Schweizer, *J. Biol. Chem.* **1998**, *273*, 22334 – 22339.
[16] S. D. Bentley, K. F. Chater, A.-M. Cerdeño-Tárraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C.-H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Niel, E. Rabbinowitsch, M.-A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, D. A. Hopwood, *Nature* **2002**, *417*, 141 – 147.
[17] S. Omura, H. Ikeda, J. Ishikawa, A. Hanamoto, C. Takahashi, M. Shinose, Y. Takahashi, H. Horikawa, H. Nakazawa, T. Osonoe, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12215 – 12220.
[18] http://www.sanger.ac.uk/Projects/S_coelicolor
[19] P. Caffrey, B. Green, L. C. Packman, B. J. Rawlings, J. Staunton, P. F. Leadlay, *Eur. J. Biochem.* **1991**, *195*, 823 – 830.
[20] G. A. Roberts, J. Staunton, P. F. Leadlay, *FEBS Lett.* **1993**, *214*, 305 – 311.
[21] R. S. Gokhale, S. Y. Tsuji, D. E. Cane, C. Khosla, *Science* **1999**, *284*, 482 – 485.
[22] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403 – 410.
[23] H. Ochman, A. S. Gerber, D. L. Hartl, *Genetics* **1988**, *120*, 621 – 623.
[24] C. J. Rowe, J. Cortés, S. Gaisser, J. Staunton, P. F. Leadlay, *Gene* **1998**, *216*, 215 – 223.
[25] J. D. Thompson, D. G. Higgins, T. J. Gibson, *Nucleic Acid Res.* **1994**, *22*, 4673 – 4680.
[26] S. Donadio, M. J. Staver, *Gene* **1993**, *126*, 147 – 151.
[27] http://avermitilis.ls.kitasato-u.ac.jp
[28] T. Schwecke, unpublished results.
[29] N. Wu, S. Y. Tsuji, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **2001**, *123*, 6465 – 6474.
[30] N. Wu, D. E. Cane, C. Khosla, *Biochemistry* **2002**, *41*, 5056 – 5066.
[31] R. J. Cox, T. S. Hitchman, K. J. Byrom, S. C. Findlow, J. A. Tanner, J. Crosby, T. J. Simpson, *FEBS Lett.* **1997**, *405*, 267 – 272.
[32] B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **2001**, *291*, 1790 – 1792.
[33] T. M. Rose, E. R. Schultz, J. G. Henikoff, S. Pietrokovski, C. M. McCallum, S. Henikoff, *Nucleic Acids Res.* **1998**, *26*, 1628 – 1685.
[34] M. Oliynyk, unpublished results.
[35] J. Staunton, P. Caffrey, J. F. Aparicio, G. A. Roberts, S. S. Bethell, P. F. Leadlay, *Nat. Struct. Biol.* **1996**, *3*, 188 – 192.
[36] S. Donadio, L. Katz, *Gene* **1992**, *111*, 51 – 60.
[37] A. T. Hagemann, N. L. Craig, *Genetics* **1993**, *133*, 9 – 16.
[38] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, **1989**.
[39] Y. Li, unpublished results.

Received: September 23, 2003 [F775]

 © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim