Chemistry & Biology 13, 625–636, June 2006 DOI 10.1016/j.chembiol.2006.04.010

Evidence for a Protein-Protein Interaction Motif on an Acyl Carrier Protein Domain from a Modular Polyketide Synthase

Kira J. Weissman,[1,*] Hui Hong,[1,2] Bojana Popovic,[1] and Filip Meersman[2,3]
[1] Department of Biochemistry
University of Cambridge
80 Tennis Court Road
Cambridge CB2 1GA
United Kingdom
[2] Department of Chemistry
University of Cambridge
Lensfield Road
Cambridge CB2 1EW
United Kingdom

Summary

During biosynthesis on modular polyketide synthases (PKSs), chain extension intermediates are tethered to acyl carrier protein (ACP) domains through phosphopantetheinyl prosthetic groups. Each ACP must therefore interact with every other domain within the module, and also with a downstream acceptor domain. The nature of these interactions is key to our understanding of the topology and operation of these multienzymes. Sequence analysis and homology modeling implicates a potential helical region (helix II) on the ACPs as a protein-protein interaction motif. Using site-directed mutagenesis, we show that residues along this putative helix lie at the interface between the ACP and the phosphopantetheinyl transferase that catalyzes its activation. Our results accord with previous studies of discrete ACP proteins from fatty acid and aromatic polyketide biosynthesis, suggesting that helix II may also serve as a universal interaction motif in modular PKSs.

Introduction

The multienzyme polyketide synthases (PKSs) biosynthesize a diverse range of clinically valuable natural products, including erythromycin A, epothilone, and FK506 (Tacrolimus). PKSs function as molecular assembly lines, as each round of chain extension and processing is (usually) performed by a specific module within the giant subunits [1–3]. Each module minimally contains a ketosynthase (KS) domain that catalyzes carbon-carbon bond formation, an acyl carrier protein (ACP) domain to which the growing chain is tethered, and an acyl transferase (AT) domain that recruits the starter and extender units. Additional domains set the level of reduction of the biosynthetic intermediates and may comprise ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) activities. Chain extension is typically terminated by a thioesterase (TE) domain [4].

*Correspondence: kjw21@cus.cam.ac.uk
[3] Present address: Katholieke Universiteit Leuven, Department of Chemistry, Celestijnenlaan 200 F, B-3001 Leuven, Belgium.

During the biosynthesis, the ACP domain must interact with all of the other domains within its module (Figure 1) (modeled after data contained in [5]). After chain extension and reductive processing, the ACP also participates in an acyl transfer reaction with a domain located in a downstream module (either a KS [KS_{n+1}] or TE domain [Figure 1]). In the "double helical" model for modular PKS architecture [5], each PKS multienzyme is a homodimer, and the individual polypeptides are twisted around each other. This arrangement places the ACP at the spatial center of the module, where it can access all of the other domains (Figure 1). This mode of operation mirrors the essential role of ACP domains in fatty acid biosynthesis and many metabolic pathways [6], in which they engage in specific interactions with multiple heterologous partners. In animals, the component enzymes of fatty acid biosynthesis, including the ACP, are joined by peptide linkages into large (type I) multifunctional enzymes resembling a single PKS module [7]. In plants and most bacteria, however, the fatty acid synthase (FAS) enzymes are present as discrete proteins (type II) [8].

There is currently no structural information available for type I PKS ACPs. However, the NMR or crystal structures of a number of type II ACPs from both FAS and PKS systems [9–12], as well as the isolated ACP domain from type I FAS of rat [13], have been solved. Although the sequence identity among ACPs from such systems is low (e.g., 14%–25% for the sequences shown in Figure 2), all solved ACP structures share a common topology consisting of a four-α helix bundle [9]. Given the resemblance among ACP domains, modular PKS ACPs are expected to have a similar global fold. Each ACP is modified by posttranslational addition of the 4′-phosphopantetheine moiety from CoA by a dedicated phosphopantetheinyl transferase (PPTase) enzyme [14], which converts the inactive *apo* protein to the active *holo* form. The active Ser, which accepts the 4′-phosphopantetheine, is part of a highly-conserved DSL motif at the base of helix II, on the surface of the protein. Covalent attachment of the acyl intermediates to the terminal cysteamine thiol of the prosthetic group allows them to be shepherded efficiently among the various active sites within the PKS.

In this context, it is interesting to consider how each PKS ACP can interact with all of the other domains within its module, with the downstream KS or TE, and also with a PPTase. In type II FAS and PKS systems, in which catalysis is performed by a collection of individual enzymes, the ACP must share interfaces with all of its partner domains in order to form the active complexes. Structural, computational, and mutagenesis approaches have identified helix II on type II FAS ACPs as a universal "recognition helix," which binds to a common 3D surface feature on its target enzymes [6, 15–17]. Modeling of the interface between the ACP and KR domains from the type II actinorhodin PKS similarly implicates helix II [18]. However, given the low overall sequence homology between ACP domains from FAS and PKS systems (Figure 2), PKS ACP domains may interact with partner

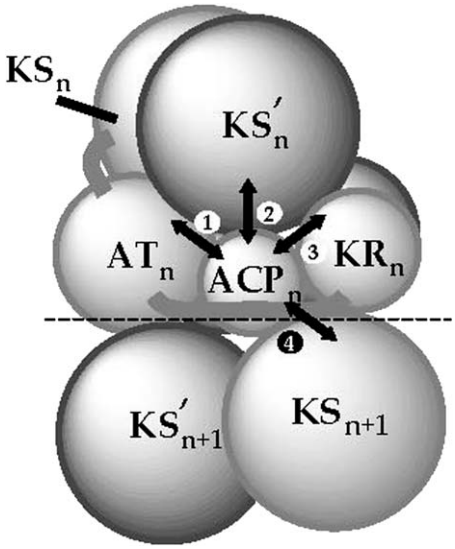

Figure 1. Protein-Protein Interactions within PKS Modules

The ACP domain is the functional center of each PKS module, as it interacts with every catalytic domain within the module during each round of chain extension and reductive processing (modeled after data contained in [5]). The specific interactions are as follows: (1) between ACP_n and AT_n during loading of the starter/extender unit; (2) between ACP_n and KS_n' during chain extension (the apostrophe indicates that the KS domain lies on a separate polypeptide from the ACP, within the PKS homodimer); (3) between ACP_n and KR_n during ketoreduction (although interaction with KR_n' has not been formally excluded); and (4) between ACP_n and KS_{n+1} during intermodular acyl transfer.

enzymes through alternative recognition regions. In addition, mutagenesis studies of the erythromycin modular PKS (DEBS) [19] and the structurally related type I animal FAS [20] suggest that at least some of the catalytic partners of each ACP domain are located on its own polypeptide, within the overall homodimeric module [5]. This arrangement may obviate the need for defined protein-protein interfaces between the ACP and some of its partner activities.

To begin to identify recognition features on type I ACP domains, we chose to investigate the interaction between a representative ACP domain from DEBS and the PPTase enzyme, SePptII, responsible for its post-translational activation [21]. In addition, we used the broad-specificity PPTase Sfp from *Bacillus subtilis* [22], which has also been shown to recognize DEBS ACP domains [21, 23]. Because the ACP and the PPTase enzymes are not covalently attached within the PKS multienzymes, they must share interfaces that can be targeted by mutagenesis. This strategy is precedented by experiments with the type II FAS ACP domains from *Escherichia coli* and *Vibrio harveyi* [24–26] as well as with various peptidyl carrier protein (PCP) domains from the tyrocidine (Tyc) nonribosomal peptide synthetase (NRPS) [27, 28]. In addition, NMR studies of individual ACP domains as both *apo* and *holo* proteins have shown that the two forms have essentially identical structures [12, 29]. This observation suggests that any recognition motif identified for the *apo* domain will be relevant to studies of the interaction between the *holo* protein and its catalytic partners within the module.

Results

Identification of a Putative Recognition Motif on DEBS ACPs

The sequence identities among modular PKS and type II FAS ACPs are low (14%–25%, for the sequences shown in Figure 2). However, the sequences are significantly more homologous over the region that corresponds to helix II, the recognition helix of the FAS ACPs (as high as 47% pairwise identity between the FAS and PKS sequences shown in Figure 2). To evaluate whether this

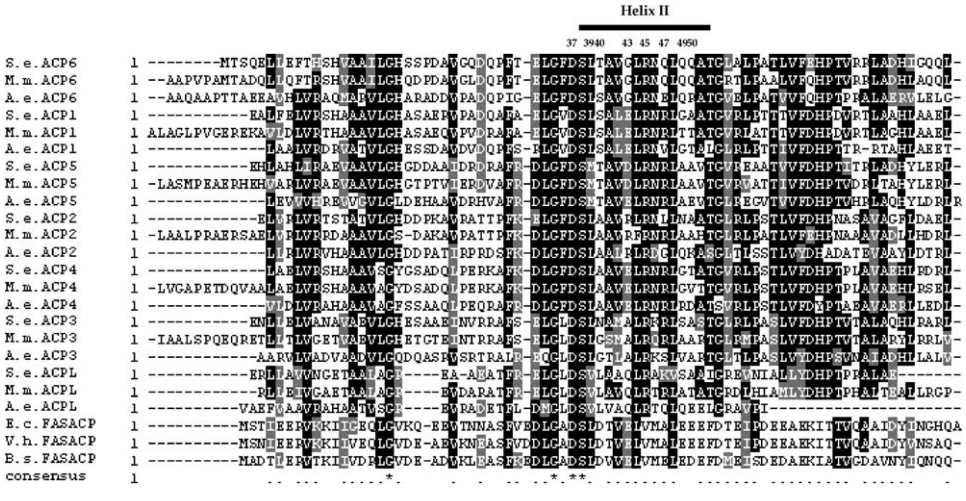

Figure 2. Sequence Alignment of PKS ACP Domains

Multiple sequence alignment of the ACP domains from the closely related PKS systems from *Saccharopolyspora erythraea* (S. e.), *Micromonospora megalomicea* (M. m.), and *Aeromicrobium erythreum* (A. e.) (all assemble the same polyketide core structure, 6-deoxyerythronolide B) with the type II fatty acid ACP domains from *Escherichia coli* (E. c.), *Vibrio harveyi* (V. h.), and *Bacillus subtilis* (B. s.) was used to select mutants within the helix II region of ACP_6. The solid bar indicates the extent of helix II in FAS ACP domains as determined by structural analysis [12]. Residues are numbered according to ACP_6. The sequences were aligned with Clustal W [60], and the figure was produced by using BoxShade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

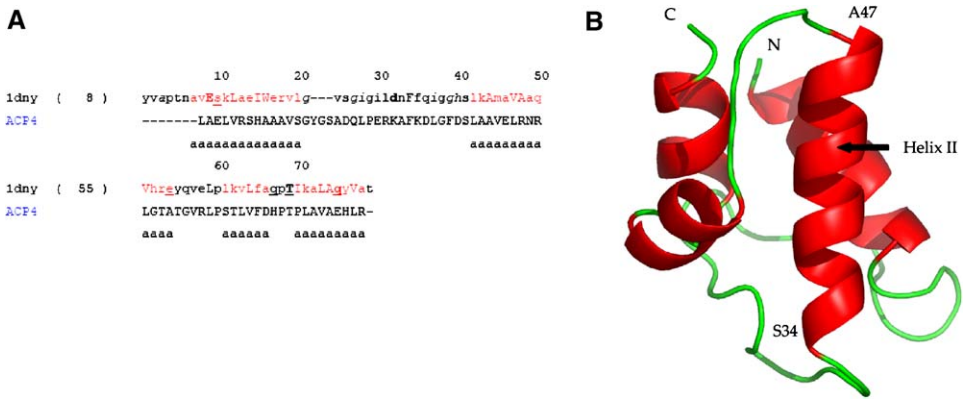

Figure 3. Sequence Alignment and Homology Model of DEBS ACP_4

(A) Sequence alignment used to build the model of ACP_4.

(B) The homology model of ACP_4, showing the predicted extent of helix II (residues S34–A47).

portion of the modular PKS ACPs is also likely to be helical, we constructed a homology model of ACP_4 from DEBS. The choice of ACP_4 was arbitrary, as the level of sequence identity among DEBS ACP domains (37%–64% pairwise) shows that their overall structures are similar [30, 31]. The sequence of ACP_4 was submitted to FUGUE [32] for both sequence and structural alignment. FUGUE identifies homologs of a query sequence by using PSI-BLAST [33], from which it calculates a sequence profile. This sequence profile is then used to screen the structural profile library within FUGUE for potential structural homologs (see Experimental Procedures). For fold assignment, the ACP_4 sequence was also submitted to the fold recognition server 3D-PSSM [34–36]. Both programs identified the PCP domain TycC3 as the preferred template for ACP_4. Although other candidate templates were obtained, none had a sequence identity with ACP_4 greater than 30%. The sequences of ACP_4 and TycC3 were therefore aligned with ClustalX and used for model building with MODELLER [37]. The 32% sequence identity between ACP_4 and the template TycC3 PCP is at the lower end of what is normally considered capable of producing a reliable homology model [30, 31]. Therefore, we confined our analysis to consider the overall fold and clear elements of secondary structure. In the model (Figure 3), ACP_4 consists of a distorted 4-α helix bundle, containing 3 major helices (I, II, and IV) (9–14 residues), a shorter helix (III) (6 residues), and a long loop region between helices I and II (21 residues). Inspection of the model shows that the predicted extent of helix II (residues 34–47) encompasses the region of highest similarity to the FAS ACP domains (Figure 2) (although the locations of the helices in ACP_4 have only been identified through an in silico experiment, for simplicity we will hereafter refer to the region between residues 34 and 47 as helix II). Therefore, both our sequence analysis and homology modeling implicated helix II of PKS ACP domains as a potential protein-protein recognition motif.

Selection of the Mutants for PPTase Interaction Studies

To evaluate the role of helix II in protein-protein recognition, we chose to carry out site-directed mutagenesis of DEBS ACP_6, covalently linked to its downstream TE partner. This didomain was first identified as a stable fragment in limited proteolysis experiments of DEBS multienzymes [38], and it can be purified in good yield from *E. coli* using nickel affinity chromatography. The putative helix II region of ACP_6 was identified by sequence alignment with DEBS ACP_4 and the FAS ACPs (Figure 2). The following mutants were selected in order to make the DEBS ACP_6 more similar to the *E. coli* FAS ACP: T40D, G43E, R45V, R45E, Q47A, Q49E, and Q50E. Residues D37 and L39, which are more highly conserved among ACPs, were both mutated to alanine. A panel of double mutants, including T40D+R45V, G43E+R45V, Q49E+R45V, and Q50E+R45V, was also created. Although V42 and N46 may also function as interface residues, the obvious mutants (such as to alanine or charged residues) at these positions occur naturally within other ACP domains of DEBS (Figure 2) [39]. In addition, the alanine variants of T40, G43, Q49, and Q50 are also present in alternative DEBS ACPs.

Although the mutagenesis strategy introduced many nonconservative mutations (for example, changes in charge on the residue), it was supported by three considerations. First, the mutant residues are present in the equivalent position within helix II of multiple FAS ACPs, so we anticipated that the mutations would not have a significant effect on the structure of the ACP_6 domain. Second, such mutations were expected to have a measurable impact on the interaction between the ACP domain and the PPTases, should they occur at a critical position within the interface. Finally, by mutating the DEBS ACP_6 domain toward that of FAS, we could evaluate whether the altered ACPs would become substrates for the PPTase which is native to the *E. coli* expression host, whose recognition region (helix I) contains multiple positively charged residues [26, 27]. A significant gain of activity would enable the expression of holo PKS proteins in *E. coli* without the need to coexpress a heterologous PPTase enzyme [21]. As proof of principle, it has been shown that reversing the charge on a single residue within the region of the TycC3 PCP corresponding to helix II (K47D, equivalent to residue T40 in the PKS ACP_6) is sufficient to make the mutant PCP a substrate for the *E. coli* PPTase [27, 28]. The double mutants of

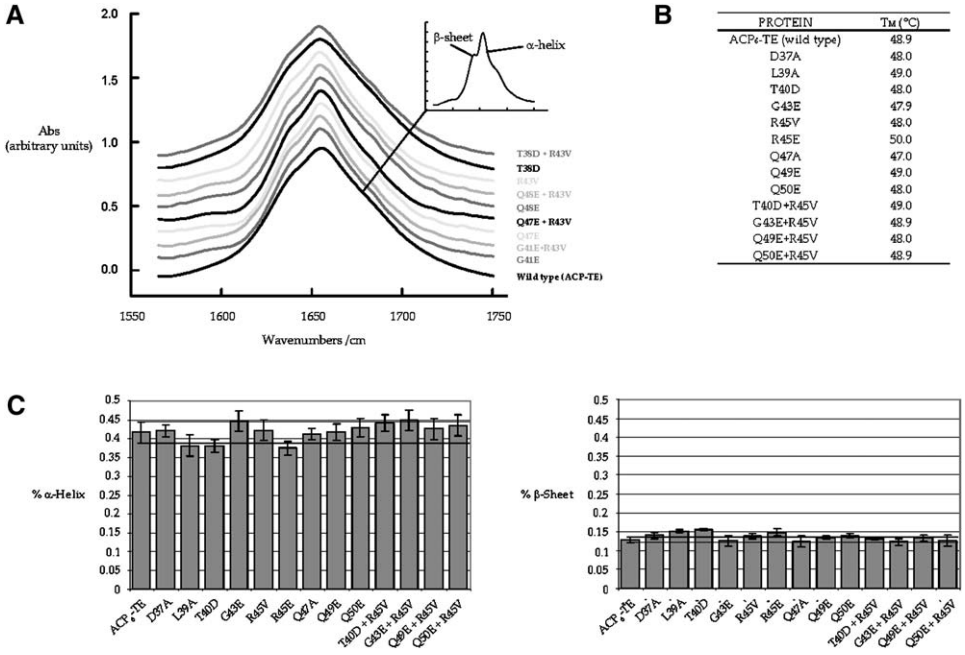

Figure 4. Secondary Structure Analysis of the Wild-Type ACP_6-TE Domain and Mutants

(A) Infrared spectra obtained of the wild-type ACP_6-TE and a selection of mutants in the amide I region (inset, the deconvoluted spectrum of the wild-type domain). The spectra have been baseline corrected.

(B) Melting temperatures (T_M) obtained for the wild-type ACP_6-TE and all mutants.

(C) Percentages of α helices and β sheets determined by deconvolution of the CD spectra of wild-type ACP_6-TE and all mutants. The bars indicate the standard deviation from the mean value for each mutant, as determined by alternative deconvolution methods. The range of predicted values for the wild-type didomain is shown relative to those of the mutants (dotted lines).

ACP_6 were particularly interesting, as previous studies have shown that the greater the resemblance of the substrate to the FAS ACP within helix II, the more efficiently it is recognized by the *E. coli* PPTase [27]. The overall charge on the proteins is also likely to be important for the interaction, as the FAS ACP has a pI of 3.81 and its PPTase partner has a pI of 9.3. The double mutants of ACP_6 had pI values (4.63–4.66) that were significantly lower than the wild-type value (5.42).

Cloning and Expression of the ACP_6-TE Mutants and Expression of the PPTase Enzymes

Site-directed mutagenesis was carried out on the ACP_6-TE gene by QuickChange mutagenesis (details on the construction of this expression vector have been reported previously [21]). The mutant plasmids were sequenced between naturally occurring SacI and DraIII sites within the ACP gene, and then the correct SacI-DraIII fragments were excised and cloned back into the unmutated parent vector. The mutated and wild-type ACP_6-TE didomains were expressed in *E. coli* BL21-CodonPlus-RP cells and purified in a single step using nickel affinity chromatography. All of the mutant proteins were expressed at similar levels to wild-type (5–7 mg ACP-TE/l of culture), with the exception of D37A (1.4-fold lower yield), suggesting that the majority of the mutations had not significantly affected the expression and native folding of the proteins. The masses of the purified proteins were determined by high-pressure liquid chromatography-mass spectrometry (HPLC-MS), which revealed that all of the proteins were in their unmodified, *apo* forms. Evidently, the mutations introduced into the ACP_6 domain were insufficient to enable its recognition by the *E. coli* PPTase enzyme.

The PPTases SePptII and Sfp were expressed and purified as described previously [21], although SePptII was obtained at an improved level of purity. The presence of residual contaminating proteins in the preparation of SePptII was not expected to affect the rates of phosphopantetheinylation in vitro, as none of the native *E. coli* proteins can perform the reaction in vivo.

Structural Analysis of the Wild-Type and Mutant ACP_6-TE Didomains

In principle, the interaction between the PPTase and the substrate ACPs could be altered by structural changes to the ACP domain resulting from the mutations. Therefore, to evaluate the structural integrity of the mutants, thin films of hydrated wild-type and mutant ACP domains were analyzed by attenuated total reflection (ATR) infrared spectroscopy. Secondary structure analysis was then performed by Gaussian curve fitting of the resolution-enhanced spectra [40, 41]. This analysis revealed no significant differences in secondary structure among the mutants (Figure 4A).

To strengthen this result, far-UV circular dichroism spectra (CD) and melting temperatures (T_M) (Figure 4B) were obtained of the wild-type and each of the mutant ACP-TE didomains. Deconvolution of the spectra was carried out with Selcon3 and CDSSTR [42] by using the Dichroweb interface [43]. The percentages of α helices and β sheets for each of the proteins were then calculated by using the combined results from both Selcon3 and CDSSTR (reference sets 3 and 6) (Figure 4C).

The secondary structure was considered to be significantly altered by the mutation(s) when the average value for the percentages of α helices or β sheets for a particular mutant was greater than one standard deviation from that of the wild-type protein (wild-type: 41.6% α-helix; 13.6 % β sheet). None of the proteins showed more than a 4% change in T_M (Figure 4B), suggesting that their overall stability was unaffected by the mutations. However, the CD spectra of L39A, T40D, and R45E showed statistically significant increases in the content of β sheets, although the proportion of α helix was not substantially different (Figure 4C).

This type of analysis does not reveal any information on changes to the tertiary structure of the proteins in the region of the mutation. However, it has recently been shown that a largely α-helical 11 residue peptide (the "ybbR tag") incorporating the conserved helix II DSL motif is efficiently recognized by Sfp [44]. This result suggests that the α-helical scaffold for the active site serine is the critical structural element for the phosphopantetheinyl transfer reaction. Our results from both IR and CD indicate that this region was essentially unchanged by the mutations, though it is impossible to rule out very subtle structural alterations.

Design of the Phosphopantetheinyl Transfer Assay

We anticipated that mutations to helix II could alter the rate of the phosphopantetheinyl transfer reaction through effects on the formation or stability of the ACP-PPTase complex as well as through effects on the catalytic step in which the ACP active site serine is modified. In either case, changes to the rate upon mutation of a specific amino acid would demonstrate that the residue in question is involved in the interaction between the enzymes, and thus help to define the protein-protein interface.

Structural and mutational analysis of Sfp has identified the loop between β4 and α5 (residues T111–S124) as the region of the enzyme that binds its native PCP substrate [45]. Among these residues, K112, E117, and K120 are proposed to engage in electrostatic and hydrogen bonding interactions with the PCP. Inspection of sequence alignments of PPTase enzymes (Figure 5) suggests that the corresponding residues in SePptII are E110, G115, and D118. By this analysis, the charges of the putative recognition residues in SePptII are reversed at two positions, while a neutral residue replaces a charged residue at the third. We could therefore predict that if helix II lies at the interface with the PPTase, then introducing negatively charged residues should have a deleterious effect on the ability of the ACP to associate with SePptII and should therefore lower the rate of phosphopantetheine transfer. In contrast, the same mutations should increase the rate of reaction by Sfp. Such data would provide additional evidence that the mutations in helix II had affected the interaction between the proteins, and had not simply introduced structural defects into the ACP_6-TE didomain.

In previous experiments to evaluate phosphopantetheinylation of mutant carrier protein (CP) domains [14, 22, 28, 45–47], the extent of phosphopantetheine transfer has been measured by the rate at which radioactivity is incorporated from ^{3}H-labeled coenzyme A into the substrate ACP or PCP. This method relies on precipitating the CP with trichloroacetic acid, prior to analysis by liquid scintillation counting and/or autoradiography. However, precipitation has been shown to introduce significant variability into the data, particularly with the low concentrations of substrate (< 1 μM) required for kinetic analysis [46, 48]. As an alternative, we investigated HPLC-MS analysis of the modified ACP_6-TE in the absence of precipitation. Our experiments showed that the reaction could be stopped at various time points by the addition of 4% glacial acetic acid. Direct HPLC-MS analysis of the quenched samples could then be used to quantify the relative amounts of *apo* and *holo* proteins in the samples, yielding an initial reaction rate (Figure 6). Control reactions in which wild-type ACP_6-TE was incubated with CoA and SePptII in buffer that had been preacidified showed no detectable phosphopantetheine transfer.

To evaluate the utility of this method for carrying out kinetic analysis, we initially determined the rate of phosphopantetheinylation of the wild-type protein and the R45V mutant. Here, we selected a range of concentrations set by the detection limit of the HPLC-MS instrument (1.5, 3.1, 5.2, and 10 μM). This analysis revealed that SePptII is inhibited by its substrate, ACP_6-TE (the rate of reaction at 10.4 μM ACP_6-TE was reduced by 5-fold relative to that at 1.5 μM). The R43V mutant also exhibited inhibition kinetics, though to a lesser extent (2-fold rate reduction over the same range of concentrations). PPTase substrate inhibition has been observed previously in experiments with FAS ACPs from *E. coli* and *Streptococcus pneumoniae* [46, 48, 49], and it may result from the presence of multiple *apo* ACP conformers or improper binding of the ACP to the active

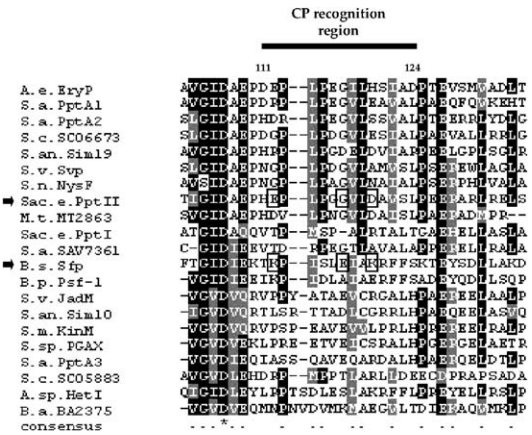

Figure 5. Multiple Alignment of PPTase Domains

The CP recognition region determined for Sfp from *Bacillus subtilis* (B. s.) is indicated by a solid bar, and three specific residues implicated by site-directed mutagenesis in electrostatic interactions with the substrate are indicated by boxes. The corresponding residues in SePptII are indicated (boxes). Abbreviations are: A. e., *Aeromicrobium erythreum*; S. a., *Streptomyces avermitilis*; S. c., *Streptomyces coelicolor*; S. an., *Streptomyces antibioticus* Tu 6040; S. v., *Streptomyces verticillus*; Sac. e., *Saccharopolyspora erythraea*; M. t., *Mycobacterium tuberculosis*; B. p., *Bacillus pumilus*; S. m., *Streptomyces murayamaensis*; S. sp., *Streptomyces* species PGA64; A. sp., *Anabena* species; B. a., *Bacillus anthracis*. The sequences were aligned with Clustal W [60], and the figure was produced by using BoxShade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

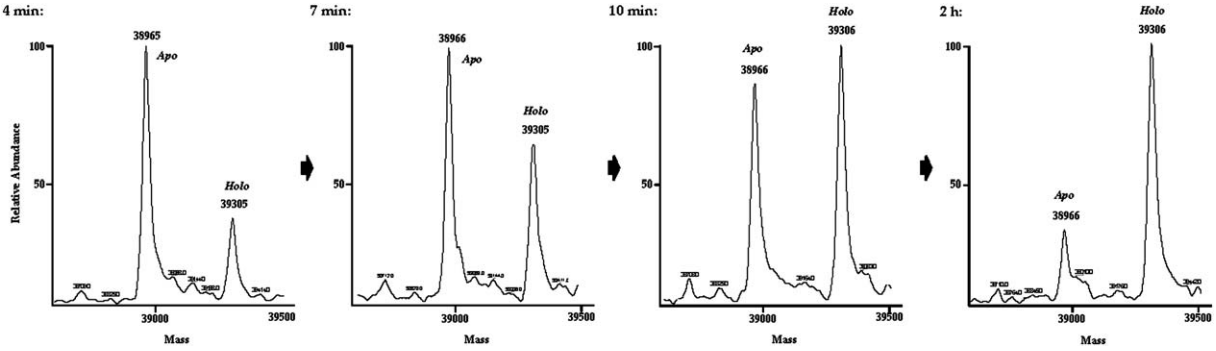

Figure 6. Representative Time Course for Phosphopantetheinylation of ACP_6-TE by SePptII

Samples (100 μl) of an assay mixture containing wild-type ACP_6-TE and SePptII were removed after 4, 7, and 10 min, quenched with 4 μl glacial acetic acid, and snap frozen in liquid nitrogen. The end point of the reaction was determined from a 2 hr incubation. All of the samples were analyzed by high-performance liquid-chromatography mass spectrometry (HPLC-MS) to determine the relative proportion of *apo* (MW 38965/6 Da) and *holo* (39305/6 Da) protein at each time point, by comparison of the corresponding peak heights.

site on the PPTase [46]. In contrast to SePptII, Sfp showed significantly less inhibition with increasing concentrations of ACP_6-TE, suggesting that its K_I value for the substrate is significantly higher than its K_M. In either case, it is unlikely that the ACP_6-TE concentrations are saturating, as the measured K_M values for natural CP substrates are in the range 0.2–5 μM, and those for heterologous CPs are typically higher [28, 45, 47].

As the substrate inhibition precluded a straightforward determination of the Michaelis constants, we instead compared the initial rates of phosphopantetheinylation of the wild-type ACP_6-TE didomain and mutants when present at the same concentration (3.1 μM). This concentration represents the minimum amount of protein required to yield a clear signal above background by HPLC-MS, and it is also not significantly inhibitory to SePptII. Based on initial studies with the wild-type ACP_6-TE, concentrations of SePptII (3.1 nM) and Sfp (15 or 31 nM, in a particular assay) were selected so that the rates of reaction with the two PPTases for all mutants were similar. Incubations were carried out under conditions reported earlier [21]: NaPi (50 mM for Sfp [pH 6] or SePptII [pH 7]), $MgCl_2$ (10 mM), DTT (5 mM), CoASH (500 μM–1 mM) in a total volume of 300 μl. The PPTase enzymes were added, and then three aliquots (100 μl) of the reaction were removed at various times within the linear portion of the rate profile (as determined by initial studies), quenched by the addition of acetic acid (4 μl), and snap frozen in liquid nitrogen. As the reaction did not always go to completion, a reaction was carried out in every case to determine the end point of phosphopantetheinylation, typically by incubating for 2 hr (a substantial decrease in the stability of ACP_6-TE was observed upon longer incubations). Even after 6 hr, the reactions with mutants L39A and R45E were incomplete; thus, the end point reached was taken as 100% modification. All samples and controls were analyzed by HPLC-MS, and the relative proportions of *apo* and *holo* ACP_6-TE at each time point were determined from the corresponding peak heights (Figure 6). The rates of reaction were then calculated by linear regression from the individual time points, incorporating a correction for the final extent of phosphopantetheinylation, if necessary.

In order that the results from each protein could be compared, the assays on the ACP_6-TE and each of the mutants were carried out simultaneously with identical samples of the PPTase enzymes. Due to the technical difficulty of analyzing the large number of samples generated in these experiments by HPLC-MS, the initial rates could only be determined twice with all of the mutants; however, in the majority of cases, rates could be determined three times. However, additional experiments with a subset of the mutants (data not shown) were always consistent with the trends observed when all of the mutants were assayed together.

Mutation of Helix II Has Opposite Effects on Phosphopantetheinylation by SePptII and Sfp

The relative rates of phosphopantetheinylation of the wild-type ACP_6-TE and the mutants by both SePptII and Sfp are shown in Table 1. In deriving these results, the average value for the initial rate for each mutant has been compared to that of the wild-type protein. The data show that changes to the rate of the reaction with both PPTase enzymes were observed for mutations along the length of helix II. Our results therefore provide strong evidence that helix II serves as a protein-protein interaction interface on the ACP domain.

The reactivity of the G43E and Q49E mutants with SePptII was essentially unchanged. Of the remaining mutants, T40D and Q47A yielded slight increases in rate; L39A, R45V/E, and Q50E had a deleterious effect; and the D37A mutant yielded no detectable reaction. D37 sits within the highly conserved "DSL" motif at the amino terminus of helix II, consistent with it playing an essential role in the reactivity of the CP domains. In the cocrystal structure of the FAS ACP from *B. subtilis* with its partner PPTase, AcpS, the corresponding Asp residues participate in an essential salt bridge with a conserved Arg residue on AcpS, which serves to correctly orient the active site Ser for modification [26]. In further support of our result, mutation of the corresponding Asp to Cys in the ACP from *E. coli* FAS completely abolished the phosphopantetheinyl transfer reaction [6].

Mutation here of the leucine within the DSL motif to alanine also markedly lowered the reaction rate (ca. 13-fold). In the ACP/AcpS cocrystal structure, the leucine

Table 1. Relative Rates of Phosphopantetheinylation of Wild-Type ACP_6-TE and Mutants by SePptII and Sfp

Protein	Reaction with SePptII (Relative to Wild-Type)	Trends in Reactivity with SePptII	Reaction with Sfp (Relative to Wild-Type)	Trends in Reactivity with Sfp
ACP_6-TE	1		1	
D37A	No reaction	N.A.	No reaction	N.A.
L39A	0.08 (0.09, 0.07)[a]	↓↓	0.020 (0.023, 0.017)[a]	↓↓
T40D	1.4 ± 0.1	↑	0.46 ± 0.05	↓
G43E	0.9 ± 0.1	–	1.30 ± 0.07	↑
R45V	0.56 (0.62, 0.51)[a]	↓	4.6 ± 0.3	↑↑
R45E	0.038 (0.032, 0.043)[a]	↓↓	0.02[b]	↓↓
Q47A	1.4 (1.3, 1.5)[a]	↑	1.4 (1.2, 1.5)[a]	↑
Q49E	0.92 ± 0.04	—	1.7 ± 0.1	↑
Q50E	0.74 ± 0.08	↓	1.2 (1.1, 1.3)[a]	↑
T40D+R45V	0.80 ± 0.08	↓	4.9 ± 0.2	↑↑
G43E+R54V	0.80 ± 0.08	↓	5.4 ± 0.4	↑↑
Q49E+R45V	0.60 ± 0.06	↓	5.2 ± 0.3	↑↑
Q50E+R45V	0.59 ± 0.04	↓	4.8 ± 0.5	↑↑

For each mutant, the initial rates obtained from multiple experiments (at least two) with each PPTase enzyme were averaged, and then compared to the average initial rate for the wild-type ACP_6-TE. Data are reported as mean ± SEM.
[a] Mean values obtained from two independent measurements (shown in parentheses).
[b] Only a single experiment with this mutant yielded useable data.

participates in an essential hydrophobic interaction with the AcpS [26], which may not be possible with a smaller Ala residue. However, our CD analysis revealed changes to the secondary structure of L39A, which might also account for the very significant decrease in rate. Of the remaining mutants, those involving residue R45 produced the most significant reductions in transfer efficiency, when present as a single mutant or within the context of double mutants. This residue is strictly conserved among ACPs from many modular PKS systems, in contrast to other residues within helix II (Figure 2; Figure S1, see the Supplemental Data available with this article online). The reductions in rate observed with this mutant cannot be attributed to an increase in substrate inhibition, as R45V gave a reduced degree of inhibition relative to wild-type over a range of concentrations. The activity of R45E was 27-fold lower than that of wild-type, although this decrease in rate could also be due to structural defects in the mutant protein. CD analysis of T40D also showed that the secondary structure was significantly altered, making the increase in reactivity seen with this mutant difficult to interpret.

Table 1 shows that Sfp also failed to modify the D37A mutant, and that its rate of reaction with L39A was likewise very slow (50-fold reduced relative to wild-type). Again, interpretation of these results is complicated by the observed structural defects in L39A. However, the data suggest that the DSL motif makes similar, critical interactions with both Sfp and SePptII. For the remaining mutants, the trends in reactivity with Sfp are essentially opposite to those with SePptII, as the majority produced a substantial increase in the reaction rate. Mutations such as G43E, Q49E, and Q50E, which introduced negative charge into helix II, resulted in a 1.3- to 1.7-fold increase in reactivity. Mutation of the R45 residue to valine in any context increased the rate of reaction by over 4.5-fold. The double mutants, in which R45V was combined with mutations that introduced negative charge along the length of the helix, yielded the highest overall increases in rate (up to 5.4-fold). This finding is relevant to the design of optimized small peptide motifs (such as the ybbR tag) for labeling by Sfp [44].

The differences in reactivity between SePptII and Sfp with the helix II mutants were exactly as predicted from our analysis of their putative CP binding regions. The potential recognition region of SePptII contains two negatively charged residues, E110 and D118. Therefore, mutations that introduced negatively charged residues into helix II were expected to adversely affect the formation or stability of the ACP-PPTase complex. Replacing G43 with glutamate, however, had very little effect on the rate, consistent with the fact that this position is occupied by E and D in other DEBS ACPs (Figure 2). The arginine residue at position 45 of the ACP could engage in attractive ionic or hydrogen bonding interactions with the residues on SePptII. Thus, removing the charge was expected to lower the rate of phosphopantetheine transfer. In contrast to SePptII, the recognition loop of Sfp incorporates two positively charged residues, K112 and K120, at the equivalent positions to E110 and D118. Consequently, mutations that introduced negative charge into ACP_6 helix II or removed a repulsive interaction between R45 and the two lysine residues were anticipated to promote the association of the proteins. Although the R45E mutant could, in principle, form new attractive interactions with K112 and K120, the structural changes to the mutant as observed by CD may account for its extremely low reactivity (50-fold reduced relative to wild-type).

Discussion

Reduced polyketides, such as erythromycin A, are assembled by multistep catalytic processes on modular PKSs. Throughout the biosynthesis, the chain-extension intermediates remain tethered to the phosphopantetheine swinging arm of the integral ACP domains. This mode of operation requires that the ACP interact with all of the other domains within each module, but also with a catalytic domain (either a KS or TE) located in the module downstream. Although the assembly process requires a specific sequence of catalytic events, the interactions that the ACP domain makes with the catalytic activities need not occur in any particular order. For

example, if the ACP domain were to interact with any of the reductive elements when bearing an extender unit such as malonate or methylmalonate instead of the newly extended polyketide chain, catalysis could not occur. After the chain is extended, the ACP could still engage in random, nonproductive interactions with the reductive loop activities without deleterious effects, as the DH domain can only operate after the KR has performed its reaction, and the ER domain can operate only after catalysis by the DH. In fact, the only "choice" the ACP makes is between interactions with enzymes within its own module and transfer of the group attached to its swinging arm to a domain in the next module. According to this model, there are only two critical interfaces within the PKS. The first is between the ACP domain and the KS activity within its own module, such that the geometric and stereoelectronic requirements of the condensation reaction can be satisfied. The second is with the acceptor domain in the downstream module, so that chain transfer occurs only at the appropriate time. Interactions with the reductive elements could occur at random, facilitated only by the proximity of the domains to the ACP by virtue of their covalent linkage into a multienzyme complex. This view contrasts markedly with the situation in type II FAS and PKS systems, in which the ACP must share defined, though transient, interaction interfaces with all of its catalytic partners.

In an alternative view [16], however, the interprotein interactions may be more strictly regulated. Subtle conformational changes on the ACP induced by interactions with the various forms of the chain-extension intermediate could expose new protein recognition sites on the ACP, making the CP a better substrate for one or the other of the catalytic domains at a particular point in the chain-extension process. In the absence of any high-resolution structural information on PKS modules, we cannot at present discriminate between these two models. Although encouragingly a 4.5 Å X-ray crystal structure of the related type I animal FAS was recently reported [50], it is striking that the position of the ACP is not at all defined. This observation underlines the importance of obtaining independent information on the specific interactions between the ACP and its partner domains, as a complement to structural efforts.

We aimed, therefore, to identify and characterize a protein-protein interaction motif on a typical ACP domain derived from a modular PKS. Multiple sequence alignment (Figure S1) reveals that ACPs from a range of modular PKS systems are highly similar; thus, each ACP is representative of a large group of domains. Although the overall sequence identity between PKS and type II FAS ACP domains is relatively low, the sequences show significantly higher homology over a 14-residue stretch, which corresponds to helix II in the FAS ACPs (Figure 2). In our homology model of an ACP domain from the DEBS PKS, this region also forms the second α helix within the overall four-helix bundle. Crucially, helix II from FAS ACP domains has previously been implicated as a universal "recognition" helix, suggesting that it may provide the same function for modular PKS ACPs.

To test the role of this potential helix, we investigated the interaction between DEBS ACP_6 and the discrete enzymes SePptII and Sfp, which catalyze phosphopantetheine transfer. Mutations at the interface were expected to have a detectable effect on the rate of *holo* protein formation. In addition, our results could be compared directly to data obtained with type II ACP domains and their PPTase partners [24–26]. Our analysis by HPLC-MS showed that mutations along the length of the putative ACP helix II produced changes to the rate of phosphopantetheinylation by both SePptII and Sfp, supporting a role for this region as an important protein recognition element. The highly conserved DSL motif at the N terminus of helix II appears to be critical for the reaction with both enzymes, with R45 in the middle of the helix also contributing significantly. In addition, the overall charge on the helix seems to be an important determinant of the interaction efficiency with Sfp.

It is difficult to design suitable controls for site-directed mutations in the absence of atomic resolution structures. However, the results of our experiments provide us with internal controls as to the effects of the mutations on both the structure and reactivity of ACP_6. Two of the mutations (G43E and Q49E) gave essentially no change in the reaction with PPTase SePptII, demonstrating that mutation alone does not alter the rate. Our confidence in the structural integrity of the mutants comes not only from the careful biophysical analysis by IR and CD [15, 16, 24], but also from the fact that the mutations produced opposite trends in reactivity with the two PPTase enzymes. Significantly, we could rationalize our results on the basis of charge differences in the region of the PPTase proteins expected to recognize their CP substrates. Such behavior would not be expected if the mutations had significantly altered the folding of the ACP domain.

Because amino acids within the putative helix II region are well conserved among PKS ACP domains (Figure S1), it is likely that PPTase-ACP interactions in type I systems are mediated, at least in part, by residues within this sequence. In contrast, the modeled interface between DEBS ACP_6 and the TE (to which it is covalently attached by a 65-residue linker [51]) is not predicted to involve helix II, except for the aspartate residue adjacent to the active site serine. In the ACP_6-TE model, the interface is formed instead between residues on the third helix of ACP_6 (E64, H64, and P65) that contact four arginines on the TE [4]. Experiments to evaluate the interaction between ACP_6 and the TE, as well as with its other partner domains within each module, are currently underway.

Significance

Genetic alteration of the modular or type I PKSs holds promise for generating modified polyketide products for use in drug discovery. However, a significant stumbling block to routine engineering remains the lack of detailed structural and mechanistic information on these gigantic multienzymes. For example, it is not known how the correct sequence of reduction reactions is accomplished within a module after chain extension, as each reduction event competes directly with transfer of the polyketide chain to the subsequent module. In engineered hybrid systems, aberrant chain transfer has been observed prior to full reductive processing. The ACP domain within each module must play a critical role in the "choice" of these alternative fates, as it shuttles the intermediate among all of the

catalytic domains, but also participates in acyl transfer to the downstream module. Detailed knowledge of the interaction interfaces that the ACP shares with all of its partner activities will provide significant insight into the three-dimensional relationship among the domains that facilitates their interactions, as well as illuminate aspects of interaction dynamics and regulation. Sequence analysis and homology modeling identify a potential helical region in modular PKS ACP domains that might serve as an interaction motif. Data obtained from site-directed mutagenesis of a model ACP_6-thioesterase didomain comprise, to our knowledge, the first evidence for a role for this helix at the ACP-PPTase interface. Our results, which are fully consistent with previous studies on discrete ACP domains from fatty acid and type II PKS systems, set the stage for identifying and characterizing all of the ACP interprotein interfaces.

Experimental Procedures

Abbreviations

The following abbreviations are used in this article: PKS, polyketide synthase; ACP, acyl carrier protein; KS, ketosynthase; AT, acyltransferase; KR, ketoreductase; DH, dehydratase; ER, enoyl reductase; TE, thioesterase; FAS, fatty acid synthase; DEBS, 6-doexyerythronolide B synthase; PPTase, 4′-phosphopantetheinyl transferase; SePptII, PPTase from *Saccharopolyspora erythraea* involved in erythromycin biosynthesis; Sfp, PPTase from *Bacillus subtilis* involved in surfactin biosynthesis; PCP, peptidyl carrier protein; Tyc, tyrocidine; NRPS, non-ribosomal peptide synthetase; CP, carrier protein.

Materials and General Methods

All chemicals were reagent grade or better. Ampicillin, carbenicillin, RNase, DNase, coenzyme A, $MgCl_2$, and Bradford Reagent were purchased from Sigma. IPTG (isopropyl-β-D-thiogalactopyranoside), dithiothreitol (DTT), and chloramphenicol were obtained from Melford Laboratories, Ltd. NaH_2PO_4 and NaCl were purchased from Fischer Chemicals. Complete protease inhibitor cocktail tablets were obtained from Roche Molecular Biochemicals. Nickel nitrolotriacetic acid (NTA) resin was obtained from Qiagen. PD-10 columns were purchased from Amersham-Pharmacia. *Escherichia coli* BL21-CodonPlus-RP was obtained from Stratagene. Standard procedures for DNA isolation and manipulation were performed as described previously [52]. Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs. Isolation of DNA fragments from agarose gel and purification of PCR products were carried out with the NucleoSpin 2 in 1 Extract kit (Macherey-Nagel, Düren, Germany). Mutagenic PCR was carried out by using the QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene), while standard PCR reactions were performed with Pfu polymerase (Stratagene); reactions were performed on a programmable Robo Cycler Gradient 96 (Stratagene). Synthetic oligonucleotides were purchased from TAG Newcastle, and automated DNA sequencing was carried out on double-stranded DNA templates by using an automated ABI Prism 3700 DNA Analyser (Applied Biosystems). Amino acid analysis was performed by the Protein & Nucleic Acid Chemistry Facility (PNAC) in the Department of Biochemistry, University of Cambridge.

Homology Modeling of DEBS ACP_4

The sequence of erythromycin ACP_4 was submitted to FUGUE [32] for both sequence alignment and fold assignment. FUGUE is a program developed to recognize even distant homologs of known structures [53]. Given a single query sequence, FUGUE runs PSI-BLAST [33] to perform a search against the NCBI nonredundant sequence database and collects sequence homologs. The alignment produced by PSI-BLAST is then used to calculate a sequence profile, which describes the observed amino acid distribution at each position of the query sequences. FUGUE then compares this sequence profile against each structural profile in its profile library derived from the HOMSTRAD database [54]. A structural profile consists of two matrices: a scoring matrix and a gap penalty matrix. The key feature of FUGUE is to calculate both matrices not only according to the amino acid sequence information, but also to environment-specific substitution tables and structure-dependent gap penalties. Therefore, scores for amino acid matching and insertions/deletions are evaluated not only on the basis of sequence similarity, but also on the local environment of each amino acid residue in a known structure. Based on the compatibility between the sequence profile and the structural profiles, FUGUE produces a list of potential homologs and alignments. In addition to FUGUE, a fold recognition server (3D-PSSM) [34–36] was also used to identify the closest structural homologs. Both servers identified the PCP TycC3 (pdb code: 1DNY) from the Tyc NRPS, with 32% sequence identity, as the most suitable structural homolog. The sequences of ACP_4 and TycC3 PCP were aligned with ClustalX and used for model building with MODELLER. The model was then improved by using the loop-building programs RAPPER [55] and CODA [56]. Final energy and structure minimization were performed by using the SYBYL (Tripos, Inc., St. Louis, MO, USA) force field. The model was then validated by PROCHECK [57], VERIFY3D [58], and by visual inspection with three-dimensional graphics software. The results from all of these programs revealed that the ACP_4 model did not require further modification.

Site-Directed Mutagenesis of pACP-TEHis

Plasmid pACP-TEHis [21] was mutated by QuickChange mutagenesis according to the manufacturer's instructions (Stratagene) by using the following mutagenic primers (the modified sequence is underlined; only the sense primer is shown): D37A (5′-CGAGCTC GGCTTCGCGTCGCTGACCGCGGTCG-3′); L39A (5′-CGGCTTCGAC TCGGCGACCGCGGTCGGGCTGC-3′); T40D (5′-CTTCGACTCGCTG GACGCGGTCGGGCTG-3′); G43E (5′-GACCGCGGTCGAGCTGCGC AACCA-3′); R45V (5′-CGCGGTCGGGCTGGTGAACCAGCTCCAGCA GG-3′); R45E (5′-CGCGGTCGGGCTGGAGAACCAGCTCCAGCAGG-3′); Q47A (5′-CGGGCTGCGCAACGCGCTCCAGCAGGCCACCG-3′); Q49E (5′-GCAACCAGCTCGAGCAGGCCACCG-3′); Q50E (5′-ACCA GCTCCAGGAGGCCACCGGGC-3′). The double mutants were derived from the corresponding single mutants T40D, G43E, Q49E, and Q50E by using the following mutagenic primers: T40D+R45V (5′-CGCGGTCGGGCTGGTGAACCAGCTCCAGCAGG-3′); G43E+R45V (5′-CGCGGTCGAGCTGGTGACCGCTCCAGCAGG-3′); Q49E+R45V (5′-CGCGGTCGGGCTGGTGAACCAGCTCGAGCAGG-3′); Q50E+R53V (5′-CGCGGTCGGGCTGGTGAACCAGCTCCAGGAGG-3′). The mutant plasmids were then sequenced between naturally occurring SacI and DraIII sites within the ACP-TE gene, and then the correct sequences were excised as SacI-DraIII fragments and cloned into unmutated pACP-TEHis previously digested with SacI and DraIII.

Expression and Purification of Wild-Type and Mutant ACP_6-TE Didomains

BL21-CodonPlus-RP cells were transformed with the plasmids containing the wild-type and mutant ACP_6-TE genes. Cultures were grown in LB medium (500 ml) supplemented with carbenicillin (50 μg/ml) and chloramphenicol (34 μg/ml) to an OD_{600} of >0.6, and induced with IPTG (1.0 mM); expression was carried out at 30ºC or 37ºC for 3.5–4.5 hr, and then the cells were harvested by centrifugation and frozen at −20ºC. The cell pellets were resuspended in chilled lysis buffer (5 mM imidazole, 50 mM NaH_2PO_4 [pH 8], 300 mM NaCl, 20% glycerol), containing protease inhibitor tablets, and trace DNase and RNase. Cells were ruptured by sonication (Misonix, Inc.), and the suspension was clarified by centrifugation at 21,000 × g for 30 min. The lysates were incubated with 500 μl Ni-NTA resin and incubated at 4ºC for ~1 hr. The resin was then washed stepwise with lysis buffer containing 5, 10, and 20 mM imidazole, and then the protein was eluted into lysis buffer containing 250 mM imidazole (hereafter referred to as elution buffer). The proteins were aliquoted, flash frozen in liquid nitrogen, and stored at −80ºC. By Bradford and amino acid analysis, typical yields were 5–7 mg ACP-TE/l of culture. The molecular weight of each of the purified proteins was obtained by high-performance liquid chromatography-mass spectrometry (HPLC-MS) on a Vydac Protein C4 column (30%–65% acetonitrile/0.1% trifluoroacetic acid over 35 min) on a ThermoFinigan LCQ.

Expression and Purification of SePptII and Sfp

A plasmid containing the gene for Sfp was a kind gift from Y. Li [59]. The Sfp expression vector and pKJW152 containing the gene for SePptII [21] were used to transform BL21-CodonPlus-RP cells. Cultures were grown in LB medium (500 ml) supplemented with kanamycin (30 μg/ml) and chloramphenicol (34 μg/ml) to an OD_{600} of >0.6, and were then induced with IPTG (0.1 mM [SePptII] and 0.2 mM [Sfp]); expression was carried out at 16ºC–18ºC overnight, and then the cells were harvested by centrifugation and frozen at −20ºC. Cells were ruptured by sonication (Misonix, Inc.), and the suspension was clarified by centrifugation at 21,000 × g for 30 min. The lysates were incubated with 500 μl Ni-NTA resin and incubated at 4ºC for ~1 hr. The resin was then washed stepwise with lysis buffer containing 5, 10, and 20 mM imidazole, and then the protein was eluted into lysis buffer containing 250 mM imidazole. The proteins were aliquoted, flash frozen in liquid nitrogen, and stored at −80ºC.

Infrared Analysis of the Wild-Type and Mutant ACP_6-TE Didomains

Attenuated total reflection (ATR) infrared spectra were recorded at room temperature on a Bruker Equinox 55 FTIR spectrometer equipped with a liquid nitrogen-cooled MCT (mercury cadmium telluride) detector. The sample compartment was continuously purged with dry air. For each sample, 256 interferograms were collected at a spectral resolution of 2/cm. Thin films of hydrated wild-type and mutant ACP-TE were obtained by depositing 10 μl sample (in elution buffer) on the crystal of a BioATRCell II setup (Bruker) and subsequent removal of excess water under a gentle N_2 flow. Elution buffer dried to the crystal was used as a background spectrum. Secondary structure analysis [40] was done by Gaussian curve fitting of the resolution-enhanced spectra (by using GRAMS/AI.7, Thermo Galactic). Second derivatives of the amide I band spectra were produced to determine the localization of the different spectral components. Resolution enhancement was achieved by Fourier self-deconvolution by using a band-narrowing parameter (g) of 10 and a Bessel smoothing function. Band assignments were made according to Jackson and Mantsch [41].

Circular Dichroism Analysis of the Wild-Type and Mutant ACP_6-TE Didomains

Far-UV CD spectra of the wild-type and mutant proteins, as well as the temperature-dependent melting profiles, were obtained by using an Aviv Circular Dichroism Spectrometer Model 215. Protein concentrations were ~0.2 mg/ml, and far-UV spectra were acquired between 260 and 180 nm with a bandwidth of 0.5 nm. For thermal stability experiments, the CD signal at 222 nm was monitored at 1ºC intervals from 25ºC to 80ºC. The exact concentrations of the protein samples were obtained by amino acid analysis, and then used to convert the background-subtracted CD spectra into molar ellipticity. Deconvolution of the spectra was carried out on the Dichroweb site [43] by using the analysis packages Selcon3 and CDSSTR [42] and both reference sets 3 and 6. These data were then used to calculate the average values and standard deviations for the percentages of α helices and β sheets in the wild-type and mutant proteins.

Determination of Conditions for the Phosphopantetheinyl Transfer Assay

The rates of phosphopantetheinylation of ACP_6-TE and the R45V mutant were determined at concentrations of 1.5, 3.1, 5.4, and 10 μM. To ensure that the total volume of substrate added to the assays was constant, the 5.4 and 10 μM concentrations were obtained from concentrated (4×) stocks of ACP_6-TE and R45V. Assay mixtures containing the ACP_6-TE and R45V in buffer [21, 22] (NaPi [50 mM; pH 6 for Sfp or pH 7 for SePptII], $MgCl_2$ [10 mM], DTT [5 mM], CoASH [1 mM]) were prepared in a total volume of 300 μl, and the reactions were initiated by the addition of either SePptII (3.1 nM) or Sfp (31 nM). Aliquots (100 μl) were removed at 2.5, 4.5, and 6.5 min and quenched by the addition of 4 ml glacial acetic acid. To determine the end point of the phosphopantetheinylation transfer reaction in each case, identical assays (100 μl) were carried out for 2 hr and then quenched with acetic acid. All of the quenched samples were snap frozen in liquid nitrogen and stored at −80ºC until analyzed by HPLC-MS (see above). The percentage of *holo* protein at each time point was obtained from the relative peak heights for the *apo* and *holo* species, as assigned by their relative masses, and corrected for the final extent of phosphopantetheinylation, if necessary. The data were then used to derive an initial rate of reaction by linear regression (Microsoft Excel).

Phosphopantetheinyl Transfer Assays

Assay mixtures containing the ACP_6-TE and all mutants at a concentration of 3.1 μM in buffer [21, 22] (NaPi [50 mM; pH 6 for Sfp or pH 7 for SePptII], $MgCl_2$ [10 mM], DTT [5 mM], CoASH [1 mM]) were prepared in a total volume of 300 μl, and the reactions were initiated by the addition of either SePptII (3.1 nM) or Sfp (15.5 or 31 nM in a particular set of assays). As the substrates were not eluted from the Ni-NTA resin at identical concentrations as determined by Bradford assay and amino acid analysis, differences in the volumes of protein added to the assays were equalized by the addition of pure elution buffer. Aliquots (100 μl) were removed at three time points within the linear portion of each reaction (determined during preliminary studies) and were quenched by the addition of 4 μl glacial acetic acid. In every case, an identical reaction (100 μl) was carried out in parallel to determine the final extent of modification (L39A and R45E for 6 hr; 2 hr for all other mutants). All of the quenched samples were snap frozen in liquid nitrogen and stored at −80ºC until they were analyzed by HPLC-MS (see above). The percentage of *holo* protein at each time point was obtained from the relative peak heights for the *apo* and *holo* species, as assigned by their relative masses, and corrected for the final extent of phosphopantetheinylation, if necessary. The data were then used to derive an initial rate of reaction for the wild-type ACP_6-TE and each mutant by linear regression (Microsoft Excel). The combined results from multiple initial rate determinations (at least two, but usually three) were used to derive an average rate of phosphopantetheinylation for each mutant, which was then compared to the average rate of reaction of the wild-type protein (Table 1).

Supplemental Data

Supplemental Data include an alignment of representative PKS ACP domains and are available at http://www.chembiol.com/cgi/content/full/13/6/625/DC1/.

Acknowledgments

Professor Peter Leadlay is thanked for advice and critical reading of this manuscript. We are grateful for the support of this work by a project grant from the United Kingdom Biotechnology and Biological Sciences Research Council (BBSRC). K.J.W. is a Royal Society Dorothy Hodgkin Fellow.

Received: December 13, 2005
Revised: April 24, 2006
Accepted: April 25, 2006
Published: June 23, 2006

References

1. Staunton, J., and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.
2. Walsh, C.T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science *303*, 1805–1810.
3. Weissman, K.J., and Leadlay, P.F. (2005). Combinatorial biosynthesis of reduced polyketides. Nat. Rev. Microbiol. *3*, 925–936.
4. Tsai, S.C., Miercke, L.J., Krucinski, J., Gokhale, R., Chen, J.C., Foster, P.G., Cane, D.E., Khosla, C., and Stroud, R.M. (2001). Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel. Proc. Natl. Acad. Sci. USA *98*, 14808–14813.
5. Staunton, J., Caffrey, P., Aparicio, J.F., Roberts, G.A., Bethell, S.S., and Leadlay, P.F. (1996). Evidence for a double-helical structure for modular polyketide synthases. Nat. Struct. Biol. *3*, 188–192.
6. Worsham, L.M., Earls, L., Jolly, C., Langston, K.G., Trent, M.S., and Ernst-Fonberg, M.L. (2003). Amino acid residues of

Escherichia coli acyl carrier protein involved in heterologous protein interactions. Biochemistry *42*, 167–176.

7. Asturias, F.J., Chadick, J.Z., Cheung, I.K., Stark, H., Witkowski, A., Joshi, A.K., and Smith, S. (2005). Structure and molecular organization of mammalian fatty acid synthase. Nat. Struct. Mol. Biol. *12*, 225–232.
8. Schneider, G. (2005). Enzymes in the biosynthesis of aromatic polyketide antibiotics. Curr. Opin. Struct. Biol. *15*, 629–636.
9. Findlow, S.C., Winsor, C., Simpson, T.J., Crosby, J., and Crump, M.P. (2003). Solution structure and dynamics of oxytetracycline polyketide synthase acyl carrier protein from *Streptomyces rimosus*. Biochemistry *42*, 8423–8433.
10. Holak, T.A., Kearsley, S.K., Kim, Y., and Prestegard, J.H. (1988). Three-dimensional structure of acyl carrier protein determined by NMR pseudoenergy and distance geometry calculations. Biochemistry *27*, 6135–6142.
11. Kim, Y., and Prestegard, J.H. (1989). A dynamic model for the structure of acyl carrier protein in solution. Biochemistry *28*, 8792–8797.
12. Xu, G.Y., Tam, A., Lin, L., Hixon, J., Fritz, C.C., and Powers, R. (2001). Solution structure of *B. subtilis* acyl carrier protein. Structure (Camb) *9*, 277–287.
13. Reed, M.A., Schweizer, M., Szafranska, A.E., Arthur, C., Nicholson, T.P., Cox, R.J., Crosby, J., Crump, M.P., and Simpson, T.J. (2003). The type I rat fatty acid synthase ACP shows structural homology and analogous biochemical properties to type II ACPs. Org. Biomol. Chem. *1*, 463–471.
14. 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). A new enzyme superfamily – the phosphopantetheinyl transferases. Chem. Biol. *3*, 923–926.
15. Zhang, Y.-M., Rao, M.S., Heath, R.J., Price, A.C., Olson, A.J., Rock, C.O., and White, S.W. (2001). Identification and analysis of the acyl carrier protein (ACP) docking site on β-ketoacyl-ACP synthase III. J. Biol. Chem. *276*, 8231–8238.
16. Zhang, Y.-M., Wu, B., Zheng, J., and Rock, C.O. (2003). Key residues responsible for acyl carrier protein and β-ketoacyl-ACP reductase (FabG) interaction. J. Biol. Chem. *278*, 52935–52943.
17. Zhang, Y.-M., Marrakchi, H., White, S.W., and Rock, C.O. (2003). The application of computational methods to explore the diversity and structure of bacterial fatty acid synthase. Prog. Lipid Res. *44*, 1–10.
18. Hadfield, A.T., Limpkin, C., Tertasin, W., Simpson, T.J., Crosby, J., and Crump, M.P. (2004). The crystal structure of the actIII actinorhodin polyketide reductase: proposed mechanism for ACP and polyketide binding. Structure *12*, 1865–1875.
19. Gokhale, R.S., Lau, J., Cane, D.E., and Khosla, C. (1998). Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase. Biochemistry *37*, 2524–2528.
20. Joshi, A.K., Rangan, V.S., Witkowski, A., and Smith, S. (2003). Engineering of an active animal fatty acid synthase dimer with only one competent subunit. Chem. Biol. *10*, 169–173.
21. Weissman, K.J., Hong, H., Oliynyk, M., Siskos, A.P., and Leadlay, P.F. (2004). Identification of a phosphopantetheinyl transferase for erythromycin biosynthesis in *Saccharopolyspora erythraea*. ChemBioChem *5*, 116–125.
22. Quadri, L.E., Weinreb, P.H., Lei, M., Nakano, M.M., Zuber, P., and Walsh, C.T. (1998). Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. Biochemistry *37*, 1585–1595.
23. Pfeifer, B.A., Admiraal, S.J., Gramajo, H., Cane, D.E., and Khosla, C. (2001). Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. Science *291*, 1790–1792.
24. Flaman, A.S., Chen, J.M., Van Iderstine, S.C., and Byers, D.M. (2001). Site-directed mutagenesis of acyl carrier protein (ACP) reveals amino acid residues involved in ACP structure and acyl-ACP synthetase activity. J. Biol. Chem. *276*, 35934–35939.
25. Gong, H., and Byers, D.M. (2003). Glutamate-41 of *Vibrio harveyi* acyl carrier protein is essential for fatty acid synthase but not acyl-ACP synthetase activity. Biochem. Biophys. Res. Commun. *302*, 35–40.
26. Parris, K.D., Lin, L., Tam, A., Mathew, R., Hixon, J., Stahl, M., Fritz, C.C., Seehra, J., and Somers, W.S. (2000). Crystal structures of substrate binding to *Bacillus subtilis* holo-(acyl carrier protein) synthase reveal a novel trimeric arrangement of molecules resulting in three active sites. Struct. Fold. Des. *8*, 883–895.
27. Finking, R., Mofid, M.R., and Marahiel, M.A. (2004). Mutational analysis of peptidyl carrier protein and acyl carrier protein synthase unveils residues involved in protein-protein recognition. Biochemistry *43*, 8946–8956.
28. Mofid, M.R., Finking, R., and Marahiel, M.A. (2002). Recognition of hybrid peptidyl carrier proteins/acyl carrier proteins in nonribosomal peptide synthetase modules by the 4′-phosphopantetheinyl transferases AcpS and Sfp. J. Biol. Chem. *277*, 17023–17031.
29. Wong, H.C., Liu, G., Zhang, Y.-M., Rock, C.O., and Zheng, Z. (2002). The solution structure of acyl carrier protein from *Mycobacterium tuberculosis*. J. Biol. Chem. *277*, 15874–15880.
30. Goldman-Fischman, S., and Honig, B. (2003). Structural genomics: computational methods for structure analysis. Protein Sci. *12*, 1813–1821.
31. Aloy, P., Pichaud, M., and Russel, R.B. (2005). Protein complexes: structure prediction challenges for the 21st century. Curr. Opin. Chem. Biol. *15*, 15–22.
32. Shi, J., Blundell, T.L., and Mizuguchi, K. (2001). FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. *310*, 243–257.
33. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.
34. Fischer, D., Barret, C., Bryson, K., Elofsson, A., Godzik, A., Jones, D., Karplus, K.J., Kelley, L.A., Maccallum, R.M., Pawowski, K., et al. (1999). CAFASP-1: critical assessment of fully automated structure prediction methods. Proteins (Suppl. 3), 209–217.
35. Kelley, L.A., Maccallum, R., and Sternberg, M.J. (1999). Recognition of remote protein homologies using three-dimensional information to generate a position specific scoring matrix in the program 3D-PSSM. In RECOMB 99, Proceedings of the Third Annual Conference on Computational Molecular Biology, S. Istrail, P. Pevzner, and M. Waterman, eds. (New York: The Association for Computing Machinery), pp. 218–225.
36. Kelley, L.A., MacCallum, R.M., and Sternberg, M.J. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. J. Mol. Biol. *299*, 499–520.
37. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., and Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. Proteins *23*, 318–326.
38. Caffrey, P., Green, B., Packman, L.C., Rawlings, B.J., Staunton, J., and Leadlay, P.F. (1991). An acyl-carrier-protein-thioesterase domain from the 6-deoxyerythronolide B synthase of *Saccharopolyspora erythraea*. High-level production, purification and characterisation in *Escherichia coli*. Eur. J. Biochem. *195*, 823–830.
39. Brikun, I.A., Reeves, A.R., Cernota, W.H., Luu, M.B., and Weber, J.M. (2004). The erythromycin biosynthetic gene cluster of *Aeromicrobium erythreum*. J. Ind. Microbiol. Biotechnol. *31*, 335–344.
40. Byler, D.M., and Susi, H. (1996). Examination of the secondary structure of proteins by deconvoluted FTIR spectra. Biopolymers *25*, 469–487.
41. Jackson, M., and Mantsch, H.H. (1995). The use and misuse of FTIR spectroscopy in the determination of protein structure. Crit. Rev. Biochem. Mol. Biol. *30*, 95–120.
42. Sreerama, N., Venyaminov, S.Y., and Woody, R.W. (2000). Estimation of protein secondary structure from CD spectra: inclusion of denatured proteins with native protein in the analysis. Anal. Biochem. *287*, 243–251.
43. Whitmore, L., and Wallace, B.A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res. *32*, W668–W673.
44. Yin, J., Straight, P.D., McLoughlin, S.M., Zhou, Z., Lin, A.J., Golan, D.E., Kelleher, N.L., Kolter, R., and Walsh, C.T. (2005). Genetically encoded short peptide tag for versatile protein labeling

by Sfp phosphopantetheinyl transferase. Proc. Natl. Acad. Sci. USA *102*, 15815–15820.

45. Mofid, M.R., Finking, R., Essen, L.O., and Marahiel, M.A. (2004). Structure-based mutational analysis of the 4′-phosphopantetheinyl transferase Sfp from *Bacillus subtilis*: carrier protein recognition and reaction mechanism. Biochemistry *43*, 4128–4136.
46. Flugel, R.S., Hwangbo, Y., Lambalot, R.H., Cronan, J.E., Jr., and Walsh, C.T. (2000). Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in *Escherichia coli*. J. Biol. Chem. *275*, 959–968.
47. Sánchez, C., Du, L., Edwards, D.J., Toney, M.D., and Shen, B. (2001). Cloning and characterization of a phosphopantetheinyl transferase from S*treptomyces verticillus* ATCC15003, the producer of the hybrid peptide-polyketide antitumor drug bleomycin. Chem. Biol. *8*, 725–738.
48. McAllister, K.A., Peery, R.B., Meier, T.I., Fischl, A.S., and Zhao, G. (2000). Biochemical and molecular analyses of the *Streptococcus pneumoniae* acyl carrier protein synthase, an enzyme essential for fatty acid biosynthesis. J. Biol. Chem. *275*, 30864–30872.
49. Lambalot, R.H., and Walsh, C.T. (1995). Cloning, overproduction, and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. J. Biol. Chem. *270*, 24658–24661.
50. Maier, T., Jenni, S., and Ban, N. (2006). Architecture of mammalian fatty acid synthase at 4.5 Å resolution. Science *311*, 1258–1262.
51. Ansari, M.Z., Yadav, G., Gokhale, R.S., and Mohanty, D. (2004). NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. Nucleic Acids Res. *32*, W405–W413.
52. J. Sambrook, E.F. Fritsch, and T. Maniatis, eds. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
53. Miguel, R.N., Shi, J., and Mizuguchi, K. (2002). Protein fold recognition and comparative modeling using HOMSTRAD, JOY and FUGUE. In Protein Structure Prediction: Bioinformatic Approach, I. Tsigelny, ed. (La Jolla, CA: International University Line [IUL] Publishers), pp. 143–169.
54. Mizuguchi, K., Deane, C.M., Blundell, T.L., and Overington, J.P. (1998). HOMSTRAD: a database of protein structure alignments for homologous families. Protein Sci. *7*, 2469–2471.
55. de Bakker, P.I., DePristo, M.A., Burke, D.F., and Blundell, T.L. (2003). *Ab initio* construction of polypeptide fragments: accuracy of loop decoy discrimination by an all-atom statistical potential and the AMBER force field with the Generalized Born solvation model. Proteins *51*, 21–40.
56. Deane, C.M., and Blundell, T.L. (2001). CODA: a combined algorithm for predicting the structurally variable regions of protein models. Protein Sci. *10*, 599–612.
57. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
58. Eisenberg, D., Luthy, R., and Bowie, J.U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. *277*, 396–404.
59. Li, Y., Llewellyn, N.M., Giri, R., Huang, F., and Spencer, J.B. (2005). Biosynthesis of the unique amino acid side chain of butirosin: possible protective-group chemistry in an acyl carrier protein-mediated pathway. Chem. Biol. *12*, 665–675.
60. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *11*, 4673–4680.