BIOCHEMISTRY

 Volume 45, Number 50 December 19, 2006

Current Topics

Carrier Protein Structure and Recognition in Polyketide and Nonribosomal Peptide Biosynthesis†

Jonathan R. Lai, Alexander Koglin, and Christopher T. Walsh*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115

Received September 22, 2006; Revised Manuscript Received October 31, 2006

ABSTRACT: Carrier proteins, 80−100 residues in length, serve as information-rich platforms to present growing acyl and peptidyl chains as covalently tethered phosphopantetheinyl−thioester intermediates during the biosynthesis of fatty acid, polyketide, and nonribosomal natural products. Carrier proteins are recognized both in cis and in trans by partner catalytic domains that effect chain-elongating condensations, redox adjustments, other tailoring steps, and finally kinetically controlled disconnection and release of the mature natural product. Dissection of regions of carrier proteins that are specifically recognized by upstream and downstream catalytic partner proteins is deciphering the logic for multiprotein assembly line construction of these large classes of natural products.

The fatty acyl chains in primary metabolites, and the scaffolds of thousands of polyketide (PK)[1] and nonribosomal peptide (NRP) secondary metabolites are assembled by elongation of a small number of carboxylic acid building blocks (*1*−*5*). These monomer units are all activated as acyl/aminoacyl thioesters to drive condensation steps that result in production of a peptide or ketide chain. While a small subset of polyketide synthases [type III PKSs (*6*)] use acyl-CoAs as activated monomers, all fatty acid synthases (FASs), type I and II PKSs, and nonribosomal peptide synthetases (NRPSs) tether the acyl and aminoacyl monomer units on the terminal thiol of a phosphopantetheine prosthetic group that is covalently attached to protein domains (*1*−*5*, *7*, *8*). (Type I systems have catalytic and carrier protein domains organized in functional modules with up to 12 modules per subunit, while type II enzymes have catalytic and carrier protein domains on separate subunits.) Thus, post-translational priming of all the carrier proteins in FAS, PKS, and NRPS assembly lines is a prerequisite for biosynthesis of these natural product classes. The phosphopantetheine arm is installed by a set of priming enzymes known as phosphopantetheinyl transferases (PPTases). The phosphodiester bond of coenzyme A is attacked by the side chain of a specific serine residue to convert inactive apo forms of the assembly lines to active holo forms containing the nucleophilic terminal thiolate group (Figure 1A) (*7*−*9*). A comparable logic of post-translational modification by a prosthetic group acting

† C.T.W. is supported by the National Institutes of Health (AI042738, GM020011, and GM047467). J.R.L. acknowledges a postdoctoral fellowship from the Helen Hay Whitney Foundation.

* To whom correspondence should be addressed. Phone: (617) 432-1715. Fax: (617) 432-0438. E-mail: christopher_walsh@hms.harvard.edu.

[1] Abbreviations: PK, polyketide; NRP, nonribosomal peptide; PKS, polyketide synthase; FAS, fatty acyl synthase; NRPS, nonribosomal peptide synthetase; CoASH, coenzyme A (free thiol); ACP, acyl carrier protein; DEBS, deoxyerythronolide B synthase; PCP, peptidyl carrier protein; ArCP, aryl carrier protein; PPTase, phosphopantetheinyl transferase; KS domain, ketosynthase domain; C domain, condensation domain; AT domain, acyl transferase domain; A domain, adenylation domain; NMR, nuclear magnetic resonance; TE domain, thioesterase domain; DHB, 3,5-dihydroxybenzoate; WT, wild-type; CLF, chain length factor.

10.1021/bi061979p CCC: $33.50
Published on Web 11/23/2006

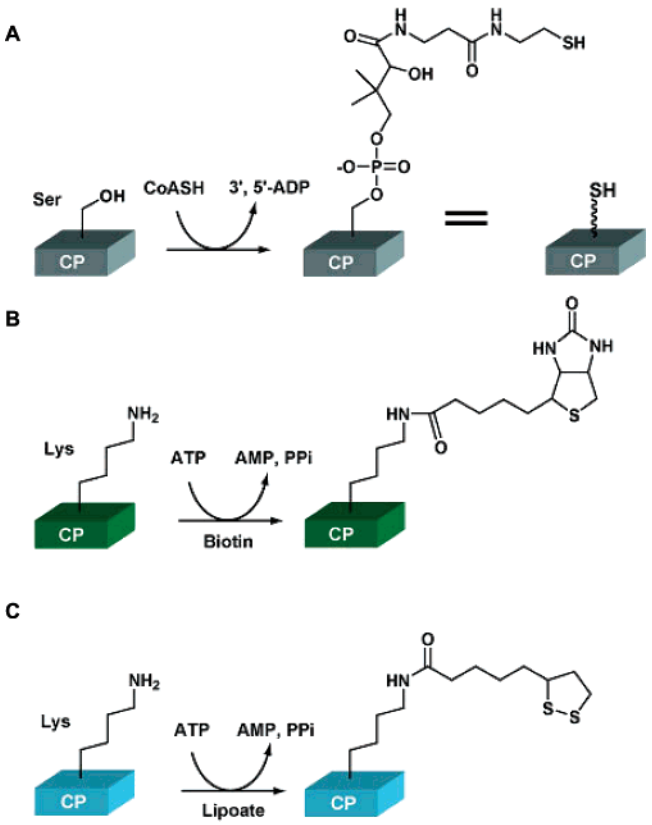

FIGURE 1: Post-translational modification of carrier proteins (CPs). (A) Phosphopantetheinylation of serine in carrier proteins involved in fatty acid, polyketide, and nonribosomal peptide biosynthesis. (B) Biotinylation of lysine in biotin carrier protein, required for CO_2 fixation. (C) Attachment of lipoate to the lysine of enzymes involved in α-keto acid dehydrogenase complexes.

as an extended arm also occurs when biotin and lipoate groups are attached to their respective apoproteins for carboxylation and decarboxylation metabolism, respectively (panels B and C of Figure 1, respectively) (*9*). In all three cases, the prosthetic group is covalently attached to the folded forms of small 8−10 kDa domains that act as acyl/peptidyl carrier proteins (phosphopantetheine), carboxyl carrier proteins (biotin), or lipoamide carrier proteins (lipoate).

Historically, free-standing acyl carrier proteins (ACPs) were first detected in bacterial (e.g., *Escherichia coli*) type II FASs and later as 10 kDa domains integrated into large multidomain FAS proteins (*10*, *11*). ACPs can similarly exist in such contexts in secondary metabolite biosynthesis (*7*, *8*). The type I PKS assembly line of deoxyerythronolide B synthase contains seven ACPs that are embedded in three multimodular synthases (DEBS1−3) (*12*, *13*). In type II PKS systems, the ACP and catalytic domains are separate proteins (*14*, *15*). Most of the NRPSs that have been characterized bioinformatically and biochemically are type I multimodular megasynthases where the carrier proteins, known as peptidyl carrier proteins (PCPs), are embedded in modules along with catalytic domains (*2*, *5*, *10*). Each PCP becomes loaded with an amino acid monomer prior to peptide chain elongation. A third variant of the 8−10 kDa carrier proteins has been delineated in the synthetases of siderophores and other aryl-capped nonribosomal peptides, where aryl acids such as salicylate and 2,3-dihydroxybenzoate are activated and tethered on aryl carrier proteins (ArCPs) (*16*, *17*). Some PCPs and ACPs have been found as free-standing carrier proteins or in incomplete modules, presumably reflecting natural assembly line shuffling in progress (*18*, *19*). In assembly lines that generate PK−NRP hybrid scaffolds such as yersiniabactin (*20*), epothilones (*21*), and rapamycin (*22*), ACPs and PCPs are interspersed. This staggering of ACP and PCP domains reflects usage of an acyl CoA versus an amino acid monomer at that specific stage in the hybrid assembly line (*23*). While ACPs and PCPs share many common features (including the general carrier protein fold, discussed below), one aspect of differentiation is their overall charge: ACPs tend to be acidic while PCPs are neutral (*24*).

An impressive example of the importance of carrier proteins in cell function is demonstrated in *Streptomyces avermitilis*. This bacterium contains at least 70 multimodular FAS, PKS, and NRPS systems, which corresponds to more than 85 total carrier proteins that are involved in the biosynthesis of conditional metabolites and fatty acids (*25*). In *Mycobacterium tuberculosis*, phosphopantetheinylation of carrier proteins in the organism's >18 type I PKSs and NRPSs and two FASs is carried out by two functionally nonredundant PPTases. One PPTase is dedicated to primary (FAS) metabolism, and the other is dedicated to secondary metabolism (*26*). Both PPTases are essential for survival in several species; thus, interference with their ability to activate ACPs and PCPs by post-translational priming represents a potentially novel target for development of antimicrobial agents. The natural antibiotic platensimycin was recently shown to exert its effects by occluding interactions between the type II FAS ketosynthase and its cognate ACP (*27*). These examples show that specific recognition between carrier proteins and their associated catalytic domains can be essential for many cellular functions.

Carrier Proteins Involved in Polyketide and Nonribosomal Peptide Biosynthesis

In PK and NRP biosynthesis, the acyl or peptide backbone is built up from acyl-CoA or amino acid monomers that are tethered to carrier proteins via stepwise condensation reactions (*2*−*5*). In PKSs, carbon−carbon bonds are formed between an upstream acyl unit, presented on an ACP, and a downstream acyl unit, usually presented on a separate ACP, through a Claisen-like condensation that is catalyzed by a ketosynthase (KS) domain (*2*−*4*). This coupling reaction proceeds through a KS-bound intermediate in which the upstream acyl chain is transferred to the active site cysteine of the KS domain prior to formation of the carbon−carbon bond with the malonate extender. In NPRS systems, the condensation (C) domain catalyzes formation of an amide bond between an upstream peptidyl or acyl thioester and the downstream amino acid group, each presented on separate PCPs (*2*, *5*). The monomers to be incorporated are selectively activated by acyl transferase (AT) domains in PKS and amino acyl adenylation (A) domains in NRPS. Typical reaction cascades for elongation in NRPS and PKS biosynthesis are shown in Figure 2. In many cases, a variety of other domains can be found embedded at various positions relative to the carrier protein that can modify the newly incorporated monomer; these have been discussed in detail elsewhere (*18*).

Since the substrates for biosynthetic operations in every elongation cycle are presented as covalently attached thioesters

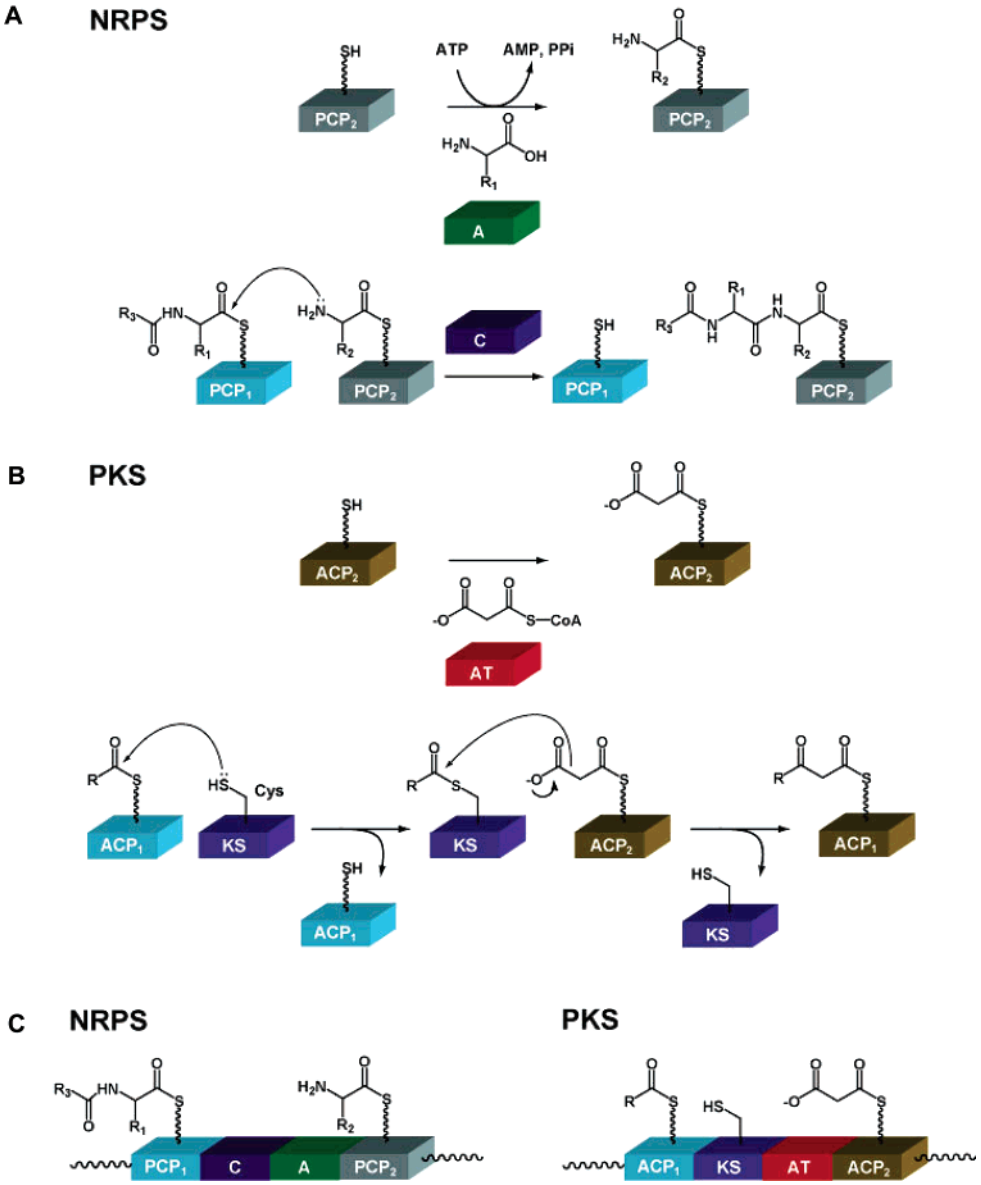

FIGURE 2: Reaction scheme for NRPS and PKS biosynthesis. (A) During production of an NRP, an amino acid is activated as the amino acyl-*O*-AMP ester by an adenylation (A) domain and subsequently thiolated onto the pantetheine arm of a peptidyl carrier protein (PCP). The condensation (C) domain mediates formation of an amide bond between two thioesters (peptidyl or amino acyl) tethered to PCP domains. (B) A similar strategy is employed in PKS biosynthesis, except that acyl-CoA monomers are incorporated onto the pantetheine of the acyl carrier protein (ACP) by acyl transferase (AT)-mediated transthioesterification. The acyl chain is transferred to the active site cysteine of the ketosynthase (KS) domain, where it is coupled with a malonate unit that is presented on a separate ACP via decarboxylation and Claisen-like condensation. (C) Typical type I domain arrangement giving rise to the coupling reactions shown in panels A and B.

to carrier proteins, the specific protein−protein interactions between carrier proteins and other PKS/NRPS catalytic domains are critical for conversion to the final product (*28*, *29*). Thus, efforts to reengineer assembly lines by domain or module swapping will require a detailed understanding of interdomain interactions (*30*−*35*). Much progress has been made in deciphering the protein recognition determinants for phosphopantetheinylation of carrier proteins (*34*, *35*). However, many other key domain−domain interactions involving carrier proteins are required during production of the PKS or NRPS scaffold. For a typical cycle of PKS elongation, an ACP must form interdomain interactions with (i) the AT domain that loads the monomer, (ii) the upstream KS domain that incorporates the monomer into the acyl chain, (iii) any other domain that performs tailoring of the resulting β-ketone, and (iv) the downstream KS domain that incorporates the next monomer into the acyl chain during the next elongation cycle. Similarly, incorporation of an amino acid in NRPS biosynthesis requires communication between the PCP and the A domain, and between the PCP and the upstream and downstream C domains. Until recently, little was known about the molecular details for how the small (80−100-residue) carrier protein is able to coordinate with several much larger domains in a series of well-timed events. For example, how much of a carrier protein's surface is involved in each of these interdomain interactions? Does each domain contact a separate face of the carrier domain? Is the swinging of the pantetheine arm directed by the catalytic domains or through the carrier protein? Several recent reports (discussed below) have provided some insight into these questions.

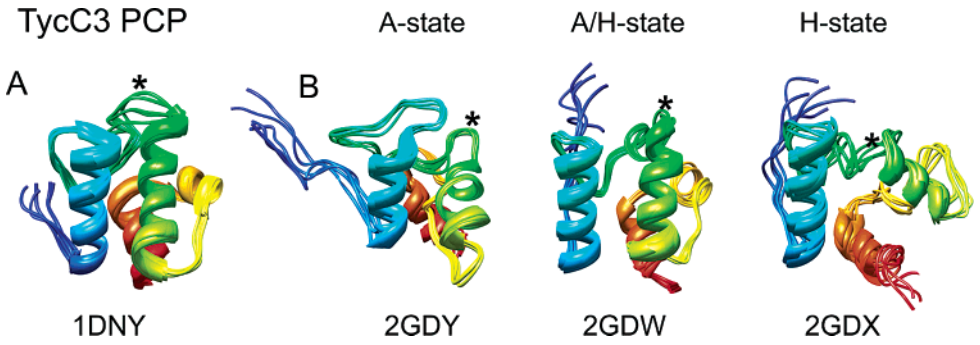

FIGURE 3: Structural mobility in the PCP of the third module of tyrocidine synthetase (TycC3-PCP). (A) Initial four-helix structure of TycC3-PCP reported by Weber et al. (*36*). (B) Conformational heterogeneity revealed upon reexamination of the TycC3-PCP structure recently reported by Koglin et al. (*42*). The PCP adopts three distinct conformations depending on the phosphopantetheinylation state. The A and A/H states are accessed by apo-PCP; holo-PCP is able to adopt the A/H and H states. The A/H state is similar to the canonical four-helix carrier protein fold (shown in panel A). The location of the conserved serine, which is phosphopantetheinylated in the holo form, is denoted with an asterisk.

Conformational Plasticity in Carrier Proteins Modulates Domain Interactions

The prototypic structure of the carrier protein is a three- or four-helix bundle such as that shown in Figure 3A for TycC3-PCP, the carrier protein from the third module of the tyrocidine synthetase of *Bacillus subtilis* (PDB entry 1DNY) (*36*). This general fold motif is similar for a number of carrier proteins from FAS, PKS, and NRPS systems, which together form a distinct superfamily of structurally well conserved proteins (*24*, *36*−*39*). The highly conserved serine residue, onto which the 4′-phosphopantetheine prosthetic group is attached, is located at the N-terminal end of the second helix (helix II).

The role of the carrier protein in PKS and NRPS biosynthesis is to shuttle substrates between upstream and downstream domains throughout the cycles of activation and condensation (*2*−*5*, *8*, *9*). To do this, carrier proteins must coordinate several adjacent domains that are much larger (in NRPS, for example, typical condensation and activation domains are ∼50 kDa each whereas carrier proteins are only 8−10 kDa in size). A recently released high-resolution crystal structure of a KS−AT fragment from the type I PKS DEBS demonstrated that a high degree of rigidity can exist in didomain fragments (*40*). The phosphopantetheine arm provides the carrier protein with a 20 Å reach; however, this distance is likely not sufficient to provide the full range of motion required to shuttle substrates between domains, especially if there is little inherent flexibility in the "linker" regions that connect the catalytic domains (such as between the KS and AT in DEBS). Intrinsic conformational motions were postulated over a decade ago for the *E. coli* acyl carrier protein (*41*). Such dynamic changes in the carrier protein structure could potentially be involved in directing the movements of the phosphopantetheine arm and of adjacent domains. However, no high-resolution information about these structural changes was available until recently.

Structural insights into the conformational motions that mediate the timing of domain interactions involving a PCP were provided by a recent reexamination of the TycC3-PCP solution structure (*42*). Initially, the structure of TycC3-PCP had been revealed to be the typical four-helix bundle shown in Figure 3A (*36*). However, upon closer examination, Koglin et al. reported that in fact both the apo and holo forms of TycC3-PCP each adopted two distinct conformational states that were in slow exchange on the NMR time scale (Figure 3B). These multiple structural states were detected in the form of two sets of $^{15}N-^{1}H^{N}$ signals each for each form of the protein (*42*). The apo and holo forms shared a common structural state, termed the "A/H" state, that was similar to the canonical four-helix carrier protein structure. The apo form adopted another conformation, the "A" state, in which helices I and II were largely unraveled. The holo form's alternative state, the "H" state, contained many elements of helices I, II, and IV but with much different relative positioning of the helices. In addition, backbone rearrangements in going from the A/H to H state resulted in the unfolding of helix III. When the conformational motions of holo-PCP were considered in relation to the cofactor, the transition resulted in an approximately 100° movement of the phosphopantetheine arm across the face of the PCP. This observation provided the first atomic-detail insight into the directed phosphopantethine arm movements required for shuttling of the cofactor between upstream and downstream domains (*43*).

Pure A and H states for TycC3-PCP were obtained by controlling the phosphopantetheinylation state or mutating the conserved serine to an alanine. Upon titration of ^{15}N-labeled apo-TycC3-PCP with the phosphopantetheinyl transferase Sfp (*44*), the residues that displayed chemical shift deviations formed a contiguous interaction surface in the A state but not the A/H state. Likewise, titration of holo-TycC3-PCP with the editing thioesterase, TEII (*45*), revealed a protein interaction surface in the H state only. Together, these data indicated that the likely mode by which the PCP is able to control interaction with separate partners is through adoption of several distinct conformational states, each of which is competent for interaction with only one catalytic partner.

While the interactions with both Sfp and TEII are biologically relevant for in vivo production of tyrocidine, the most interesting interactions with TycC3-PCP from the standpoint of NRPS engineering are those that are in cis with the upstream A domain and the downstream C domain (*5*). Recently, the 5 Å resolution crystal structure of the fungal fatty acid synthase was published, revealing the nature of many in cis interdomain interactions (*46*, *47*). However, the electron density corresponding to the ACP [and the C-terminal thioesterase (TE) domain] was not visible, likely

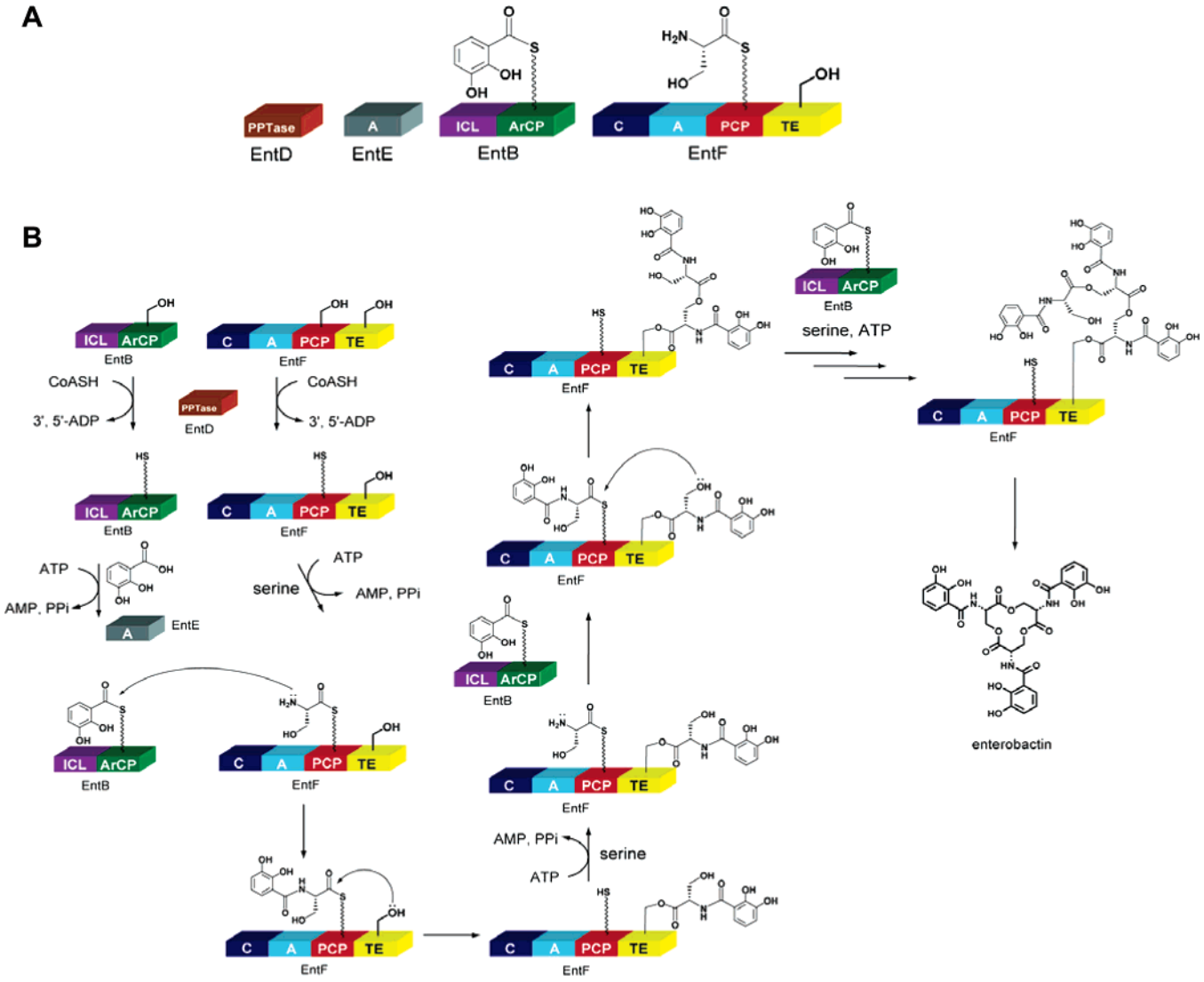

FIGURE 4: Scheme for enterobactin biosynthesis by EntBDEF. (A) The four-component synthetase consists of EntD, a phosphopantetheinyl transferase (PPTase); EntE, a free-standing adenylation (A) domain that is specific for 2,3-dihydroxybenzoate (DHB) activation; EntB, a dual-function isochorismate lyase (ICL)/aryl carrier protein (ArCP) didomain protein onto which DHB is thiolated; and EntF, a four-domain nonribosomal peptide synthetase elongation/termination module. (B) Enterobactin biosynthetic reaction scheme. EntD phosphopantetheinylates the carrier proteins of EntB and EntF. These carrier proteins are then acylated with DHB or serine by the appropriate A domains. The condensation (C) domain on EntF then catalyzes amide bond formation to produce the DHB−Ser molecule, which is subsequently transferred to the active site serine of the EntF TE domain. The activation/condensation cycle is repeated, with the subsequent DHB−Ser molecules coupled through the side chain hydroxyl. The TE domain of EntF then catalyzes lactonization and release of the mature enterobactin molecule.

due to intrinsic carrier protein structural heterogeneity and/or mobility (*46*). In fact, conformational heterogeneity has been detected in several other carrier proteins (including ACPs) (*41*, *48*−*50*), although it is not clear all specific motions will be comparable to those of TycC3-PCP. Nonetheless, future structural studies aimed at gaining insight into in cis interdomain interactions should focus on solution methods of carrier protein-containing didomain fragments to chart the dynamics (*51*, *52*).

Combinatorial Mutagenesis Studies with EntB and EntF Uncover Common Recognition Elements

Recent efforts in our laboratory have focused on using combinatorial mutagenesis coupled with selection for production of the NRP enterobactin in *E. coli* to identify important regions in carrier proteins for interdomain communication both in cis and in trans (*53*−*55*). Enterobactin is an iron-chelating siderophore that is generated in response to iron starvation (*16*, *56*−*58*). The enterobactin synthetase (Figure 4) consists of four protein components, EntBDEF, that use three molecules each of serine and 2,3-dihydroxybenzoate (DHB) to form a trilactone that is able to capture ferric iron with high affinity via hexadentate coordination with the catachols of the three DHB moieties (*56*, *57*). *E. coli* are dependent on production of enterobactin to grow on iron-deficient media (*58*), a feature we used to develop a selection for functional synthetase components by plasmid-based complementation of EntB or EntF knockout cells (*53*−*55*).

The biosynthesis of enterobactin proceeds as shown in Figure 4B. EntE is a free-standing adenylation domain that activates DHB and then acylates the aryl carrier protein

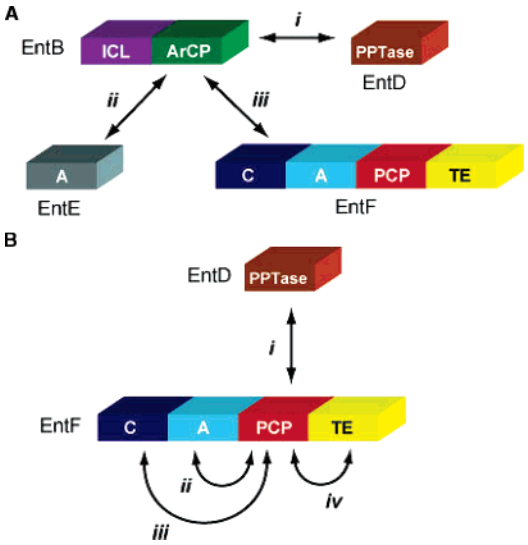

FIGURE 5: Protein−protein interactions involving carrier proteins in enterobactin synthetase. (A) Protein−protein interactions that must occur with EntB-ArCP during enterobactin biosynthesis. The ArCP must interact with (i) EntD during phosphopantetheinylation, (ii) EntE during acylation with DHB, and (iii) EntF during condensation. (B) Interdomain communication involving the PCP domain of EntF. The in trans interaction with EntD (i) must occur as described above. However, the PCP must participate in in cis communication with (ii) the A domain during serine loading, (iii) the C domain during condensation, and (iv) the TE domain during elongation/cyclotrimerization.

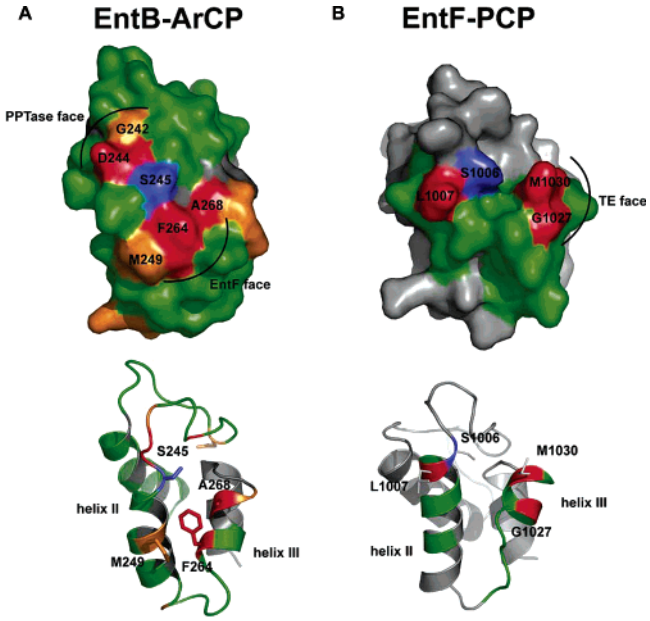

FIGURE 6: Regions of EntB-ArCP and EntF-PCP that were found to be highly conserved by combinatorial mutagenesis and selection. (A) Surface and ribbon representations of EntB-ArCP color-coded for the degree of conservation where red indicates high conservation, orange intermediate conservation, and green low conservation. The phosphopantetheinylated serine is colored blue. By biochemical characterization of point mutations, we were able to assign one cluster of residues as the interaction surface for PPTases such as EntD and Sfp (G242 and D244) and another cluster as the interaction surface for EntF (M249, F264, and A268). The side chains for these residues are shown in the ribbon representation. (B) Surface and ribbon representations of EntF-PCP, color-coded as in panel A. L1007 was involved in recognition by EntD, while G1027 and M1030 are involved in interdomain communication with the EntF TE domain.

(ArCP) of EntB (*56*, *57*). EntB also contains an N-terminal isochorismate lyase (ICL) domain that is involved in diversion of the shikimate pathway to produce DHB (*57*, *59*, *60*). However, our laboratory has shown in vitro that the ICL domain is not functionally involved in the NRPS logic of the synthetase (*56*, *57*). EntF is a serine-incorporating NRPS module. The C domain of EntF catalyzes formation of the amide bond between DHB (presented on EntB) and serine (presented on the PCP domain of EntF) (*58*, *61*). The nascent DHB−serine condensation product is then shuttled to the active site serine of the adjacent thioesterase (TE) domain. Three elongation cycles result in buildup of the linear trilactone on the TE domain, which catalyzes ring closing and release of the mature enterobactin molecule (*62*).

The aryl carrier protein of EntB (EntB-ArCP) and the PCP of EntF (EntF-PCP) represent two different scenarios for interdomain interactions in NRPS biosynthesis (Figure 5). EntB-ArCP must participate in protein−protein interactions (i.e., in trans interdomain interactions) with three different proteins: (i) with EntD during phosphopantetheinylation, (ii) with EntE during acylation with DHB, and (iii) with EntF during condensation of DHB with serine. While EntF-PCP must also participate in the in trans interaction with EntD (point i above), the biosynthetic reaction cascade requires a set of in cis interdomain communications: (ii) with the A domain during serine loading, (iii) with the C domain during condensation, and (iv) with the TE domain during formation of the trilactone and subsequent ring closure.

Combinatorial mutagenesis was performed on ~80% of the EntB-ArCP surface (a total of 44 residues), and key positions were identified by selection on iron-deficient media (*53*, *54*). The combinatorial mutagenesis scheme utilized partial codon variation such that side chains were permitted to vary among WT, Ala, and sometimes a third or fourth residue (a technique known as "shotgun alanine scanning") (*63*−*65*). From sequence analysis of a representative number of surviving colonies, the relative conservation of WT side chain identity relative to Ala was determined and categorized as high (WT preferred over Ala by more than 20-fold), intermediate (WT preferred over Ala between 6- and 20-fold), or low conservation (WT preferred over Ala by ≤6-fold) (Figure 6A). The majority of the EntB-ArCP surface residues (36 of 44 examined) were tolerant to mutation while supporting enterobactin production in vivo. However, two small clusters of conserved residues that surround the phosphopantetheinylated serine (S245) became evident. By biochemical analysis of point mutations, one cluster was assigned as the protein interaction face for EntF and the other cluster as the interface for PPTases such as EntD and Sfp (*53*, *54*). Two key conclusions emerged from this work. (a) The functional protein interaction surfaces on the ArCP are relatively localized, meaning that there are only a few key residues that participate in the protein interaction surface for EntD/Sfp or EntF; (b) in the case of EntD/Sfp and EntF, the most critical portions of the protein interaction surfaces are distinct (i.e., nonoverlapping). Both conclusions may be good news to the biosynthetic engineer; designing noncognate interactions between a carrier protein and a catalytic domain should be possible with only a few targeted mutations if these results hold true in other cases. Furthermore, if all carrier proteins have distinct recognition regions for separate catalytic domains, then altering specificity for one catalytic

partner should be possible while leaving interactions with the others intact.

The interaction surface on EntB-ArCP for EntF consisted of one residue on helix II (M249) and two residues on helix III (F264 and A268). Together, these side chains formed a contiguous hydrophobic surface adjacent to S245. The discovery that helix III in EntB-ArCP contained important recognition residues is significant in light of other findings. This portion of TycC3-PCP had the most dramatic structural changes during interconversion of the A/H and H states (*42*). Furthermore, residues in this region of holo-TycC3-PCP had significant chemical shift perturbations upon titration with TEII, indicating that this region likely also plays a role in recognition for that system. Furthermore, in the structural analysis of the ACP from frenolicin biosynthesis (FrnN), Li et al. found that the helix III region of this ACP also equilibrated between two conformations, although they were only able to assign the major conformer (*48*). Helix III conformational mobility has also been recently described in the holo-acyl carrier protein of *Plasmodium falciparum* (PfACP) (*50*). The flexibility observed for helix III of FrnN and PfACP suggests that it also may play a role in interactions in PKS assemblies. Johnson et al. recently described the solution structure of an acyl carrier protein from *Anabaena* in which the helix III region adopts a 3_{10}-helix rather than an $\alpha\alpha$-helix (*66*). Comparison of the 3_{10}-helix in the *Anabaena* ACP with helix III from other carrier proteins led these authors to postulate that the 3_{10}-helix was important for interactions in this system, too. However, no structural mobility was detected in the 3_{10}-helical region.

The putative helix II and helix III regions of EntF-PCP were analogously targeted (a structure of EntF-PCP is not yet available) for combinatorial mutagenesis to examine if common recognition elements were observed for the in cis and in trans scenarios (*55*). Two residues (G1027 and M1030) were identified to be important for communication between the PCP domain and the downstream TE domain (Figure 6B). On a structural model of EntF-PCP (based on the A/H state of TycC3-PCP), these residues are located on surface-exposed portions of the helix III region (again, proximal to the phosphopantetheinylated serine). The common thread among the helix III-mediated interactions in TycC3-PCP, EntB-ArCP, and EntF-PCP is that in all three cases, it is the acylated form of the pantetheine that is utilized by the interacting catalytic domain (*47*, *53*, *55*). These observations support a model in which the downstream recognition (be it condensation in EntF or transfer to a TE domain in TycC3-PCP and EntF-PCP) is mediated by the helix III region which may act as a conformational switch. It remains to be determined if in cis PCP−C domain interactions (and analogous ACP−KS interactions in PKS) are also mediated by a dynamic helix III. Furthermore, studies aimed at determining how the EntB-ArCP and EntF-PCP mutations result in communication defects will provide insight into the mechanism of interdomain interactions. For example, does mutation of the "hot spot" residues result in a defect in protein−protein interactions? Or does it slow the conformational dynamics such that one of the states (e.g., the H state) is no longer accessible? Studies toward this end are currently underway.

Differential Carrier Protein Recognition by PKS and NRPS Catalytic Domains in Nature

Nature utilizes orthogonal carrier protein recognition as a strategy to separate primary metabolism from secondary metabolism. For example, the phosphopantetheinyl transferase (PPTase) for lipid biosynthesis in *B. subtilis* (ACP synthase, ACPS) is not able to post-translationally modify TycC3-PCP (*34*−*36*). Likewise, our laboratory reported that ACPS from *E. coli* and EntD are largely specific to their cognate carrier proteins (ACPS recognizes the fatty acid ACP, and EntD recognizes EntB and EntF with virtually no cross-talk) (*67*, *68*). Lee et al. (*69*) described a fundamental incompatibility between ACPs involved in spore pigment and PK production. Very recently, two nonredundant but essential PPTases have been characterized in *M. tuberculosis* (*26*): one for fatty acid synthase priming and the other for production of PK and NRP virulence factors. These orthogonal recognition events are beneficial to the organism. Since expression of the biosynthetic machinery for secondary metabolites is upregulated in response to stress or starvation, it makes good biological sense to separate these systems from the primary metabolism that is required in the moment-to-moment life cycle.

In several cases, bacteria have evolved orthogonal carrier protein recognition within a single biosynthetic gene cluster (*70*−*72*). An example of this strategy is found in the type II PKS systems of frenolicin and the R1128 substances (*70*, *71*). The type II PKSs produce aromatic polyketides from a poly-β-ketone precursor that is built up by iterative condensations of malonyl-CoA extender units (*14*). However, unlike the assembly line enzymes discussed so far (the type I synthases and synthetases), type II PKSs contain a single, central ACP, and a heterodimer consisting of a KS domain and a chain length factor (CLF) that alone orchestrate the synthesis of the entire poly-β-ketone. Tailoring enzymes such as aromatases and cyclases convert the ketide into the mature product. This set of enzymes, required for production of the aromatic scaffold, together are known as the minimal PKS. Since all domains are free-standing (i.e., "dissociative") in the type II scenario, biosynthetic production requires a set of very well timed protein−protein interactions between the different catalytic subunits and the ACP.

The default starting unit for the poly-β-ketone in most type II PKSs is an acetyl moiety that is derived from decarboxylation of an initial malonyl-CoA unit (somewhat akin to loading in type II FAS) (*1*, *14*). However, several aromatic PKs (such as frenolicin and R1128) contain non-acetate starter units (*70*, *71*, *73*). The clusters for these PKSs contain an additional ACP and an additional KS domain that are dedicated to the initiation condensing reaction (ACP_i and KSIII, respectively) (*70*, *71*, *73*, *74*). The KSIII displays substrate preference for an alternate acyl-CoA over malonate (for example, proprionyl-CoA in R1128) and condenses this starter unit with a malonate that is specifically presented on the ACP_i (Figure 7A). Khosla and co-workers (*70*) demonstrated that KSIII is not able to perform the condensation with the malonyl unit presented on the ACP from the minimal PKS (here, we term this the elongation ACP or ACP_e). This orthogonality in ACP recognition is striking in light of the relatively high degree of homology between the initiation and elongation ACPs (37% in the frenolicin system

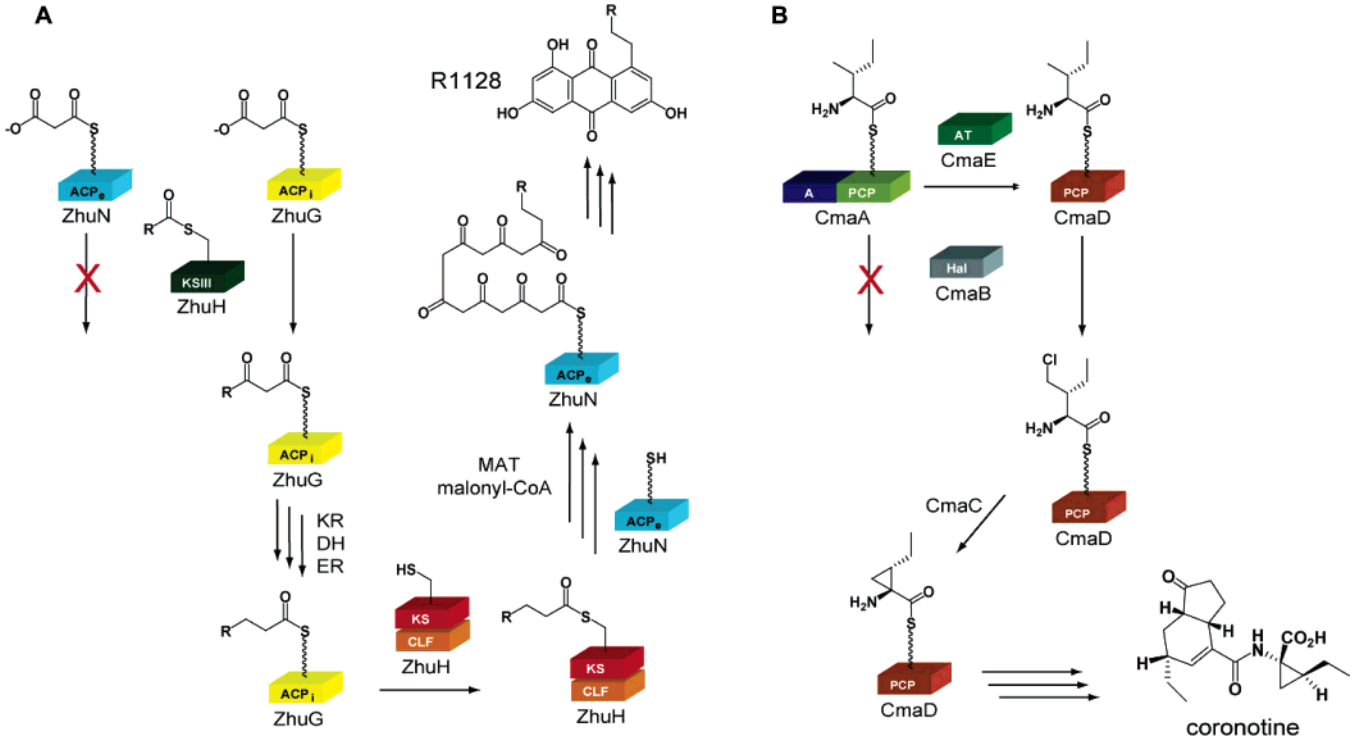

FIGURE 7: Orthogonal carrier protein recognition in natural systems. (A) During the biosynthesis of the R1128 substances [where R = CH_2CH_3 and R = $CH(CH_3)_2$ are the most abundant], initiation and elongation of a poly-β-ketone chain take place on two separate ACPs with orthogonal recognition properties. Initiation is catalyzed by ZhuH, a KSIII homologue, with malonate presented on ZhuG (the initiation ACP, or ACP_i). Malonate that is presented on ZhuN, the elongation ACP (ACP_e), is not utilized as a substrate by ZhuH. This newly formed diketide undergoes tailoring to give an alkyl butyryl group that acts as the priming unit for elongation by the tandem action of ZhuH (KS-CLF) and ZhuN. Subsequent tailoring steps yield the final aromatic product. (B) In coronatine biosynthesis, allo-isoleucine is loaded onto CmaA, an A-PCP didomain. The allo-isoleucine is then transferred to the free-standing PCP CmaD by CmaE, an aminoacyl transferase. The allo-isoleucine presented on CmaD (but not CmaA) is chlorinated by the non-heme halogenase CmaB. Subsequent action of CmaC gives rise to the cyclopropyl group that ultimately is incorporated into the final product.

and 44% in the R1128 system). Interestingly, one residue on the ACPs of the R1128 system that was found to be important for determining the KS recognition specificity aligns with M249 (on helix II) of EntB-ArCP. This observation further suggests that downstream catalysts in PKS and NRPS may share the same interfaces on their respective carrier proteins.

When different initiation KS/ACP pairs were matched with different minimal PKSs in vivo, the result was production of aromatic PKs that had been loaded with non-acetate starter units (*71*). This strategy was used to generate new, regioselectively modified aromatic PKSs, some of which exhibited anti-glucose-6-phosphatase activity. In most cases, it appeared that the first condensation product, presented on ACP_i, could be recognized by many different minimal PKS systems. Further studies will be required to determine if engineering of ACP or PCP interactions from completely exogenous systems will allow for incorporation of different starter units and consequent natural product scaffold diversification (*75*, *76*).

Another example of orthogonal carrier protein recognition occurs in the biosynthesis of the phytotoxin coronatine. A cyclopropyl-containing precursor (coronamic acid) is produced via a cryptic chlorination strategy (Figure 7B) (*72*, *77*, *78*). Allo-isoleucine is halogenated at the γ-position by a non-heme halogenase, CmaB. Then, CmaC, a zinc-dependent enzyme, catalyzes deprotonation at the α-position and subsequent SN_2-like displacement of the newly installed chlorine, yielding the cyclopropyl group. This unusual modification takes place on a carrier protein-bound substrate; however, the halogenase CmaB is specific in its recognition properties (*72*). The allo-isoleucine is activated and thiolated by an A-PCP didomain protein, CmaA, but it is not until the allo-isoleucine is transferred to a free-standing PCP, CmaD, that the halogenation reaction takes place. Thus, halogenase CmaB specifically recognizes CmaD but not the PCP of CmaA. The transfer of the allo-isoleucine between CmaA and CmaD is mediated by an aminoacyl shuttle enzyme, CmaE. The structural rationale for the differential recognition of PCP scaffolds by the CmaB halogenase is not yet clear, but the shuttle enzyme CmaE may gain general utility for moving aminoacyl moieties between holo-PCP domains.

The examples listed here demonstrate instances of specific carrier protein recognition by catalytic domains in nature. Specific recognition among two or more carrier proteins may have evolved for functional reasons such as control of substrate identity or reactivity, or it may simply be a result of gene duplication. Other PKS or NRPS clusters contain an excess of carrier proteins the function of which is not known. For example, in the mixed type I/II assembly lines of jamaicamide (*79*), mupirocin (*80*), and curacin (*81*), multiple carrier protein domains exist in tandem. The purpose of these tandem carrier proteins is not known. Perhaps carrier proteins in tandem may be a natural strategy for exerting multivalency effects, or perhaps they serve as a "waiting room" where builtup intermediates are held before subsequent modification.

Recently, our group and others (*82*−*86*) have exploited orthogonal carrier protein recognition for the development of site-specific tags based on derivatized phosphopantetheine groups. This strategy has been employed for labeling protein fusions containing a carrier protein domain in a variety of applications (*82*−*85*) and could potentially be used for shotgun sequencing of biosynthetic clusters in high-throughput work (*86*). This work shows that deciphering the rules of carrier protein recognition could prove useful not only in biosynthetic engineering efforts but also in the development of site-specific labels based on the phosphopantetheine post-translational modification.

Future Efforts for Engineering of Noncognate Carrier Protein Interactions

The manipulation of PKS or NRPS systems for the production of new natural product variants with a high efficiency is a two-pronged challenge. First, the degree of chemical substrate promiscuity by the domains that catalyze the transformations on the growing substrate chains must be deciphered (*2*). Much headway has been made toward this end, and in altering substrate selectivity of particular domains (*87*−*91*). Second, a detailed understanding of the protein-based interdomain interactions that occur during the biosynthesis is required (*28*−*37*, *91*). Since almost every intermediate for biosynthetic processing is presented on a carrier protein, a major endeavor will be to uncover the "rules" for recognition of carrier proteins by catalytic domains to effect the orderly and intended growth of a given PK, NRP, or PK−NRP scaffold.

The collective work presented here suggests that helix III plays a large role in recognition, perhaps due to its inherent structural flexibility and functional ability to act as a conformational switch. Therefore, engineering of carrier protein interactions should focus on such "hot spots". Further characterization will be required to determine (i) if all carrier proteins share the same hot spots and (ii) if different interacting domains recognize distinct hot spots. While rational site-directed mutagenesis for this engineering is a reasonable first-pass strategy for altering selectivity, the power of combinatorial mutagenesis and selection will likely be required to achieve the goal of nativelike recognition of noncognate carrier proteins.

Conclusions

Carrier protein domains represent architectural scaffolds with information-rich surfaces that can be distinguished by distinct catalytic partner proteins. Presentation of both activated monomers and the growing acyl/peptidyl chains on the phosphopantetheinyl arms of carrier protein domains represents a significant investment in machinery for natural product assembly lines. The prosthetic group requires post-translational installation on every ACP/PCP but brings a reactive thiolate at the end of the phosphopantetheine arm. Capture of the substrate acyl/aminoacyl groups as covalent thioesters provides requisite activation for Claisen and amide bond condensations, while the long arm, in conjunction with interdomain movements, reaches into upstream and downstream active sites. Attachment of the acyl-*S*-pantetheinyl arms to the 10 kDa ACP/PCP protein platform allows for specificity between multiple possible catalytic partners. Logical analogies exist in protein-based signaling cascades in eukaryotes, including the 8 kDa information-rich ubiquitin tag that is recognized by dozens of different partner proteins in controlling the fate of molecules attached to the ubiquitin scaffold (*9*, *92*).

ACKNOWLEDGMENT

We thank previous and current co-workers Michael Fischbach, Zhe Zhou, and Eric Roche (Harvard Medical School), Professor David Liu (Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA), and Professor Volker Doetsch (Frankfurt University, Frankfurt, Germany).

REFERENCES

1. White, S. W., Zheng, J., Zhang, Y.-M., and Rock, C. O. (2005) The Structural Biology of Type II Fatty Acid Biosynthesis, *Annu. Rev. Biochem. 74*, 791−831.
2. Cane, D. E., Walsh, C. T., and Khosla, C. (1998) Harnessing the Biosynthetic Code: Combinations, Permutations, and Mutations, *Science 282*, 63−68.
3. Khosla, C. (2000) Natural Product Biosynthesis: A New Interface between Enzymology and Medicine, *J. Org. Chem. 65*, 8127−8133.
4. Hopwood, D. A. (1997) Genetic Contributions to Understanding Polyketide Synthases, *Chem. Rev. 97*, 2465−2497.
5. Schwarzer, D., Finking, R., and Marahiel, M. A. (2003) Nonribosomal Peptides: From Genes to Products, *Nat. Prod. Rev. 20*, 275−287.
6. Moore, B. S., and Hopke, J. N. (2001) Discovery of a New Bacterial Polyketide Biosynthetic Pathway, *ChemBioChem 2*, 35−38.
7. Walsh, C. T., Gehring, A. M., Weinreb, P. H., Quadri, L. E. N., and Flugel, R. S. (1997) Post-translational Modification of Polyketide and Nonribosomal Peptide Synthases, *Curr. Opin. Chem. Biol. 1*, 309−315.
8. Marahiel, M. A. (1997) Protein Templates for the Biosynthesis of Peptide Antibiotics, *Chem. Biol. 4*, 561−567.
9. Walsh, C. T. (2006) *Post-translational Modification of Proteins: Expanding Nature's Inventory*, Roberts and Company Publishers, Englewood, CA.
10. Cronan, J. E., Jr., and Rock, C. O. (1996) in *Escherichia coli and Salmonella typhimurium: Cellular and molecular biology* (Neidhardt, F. C., Curtis, R., Gross, C. A., Ingraham, J. L., Lin, E. C. C. et al., Eds.) pp 612−636, American Society for Microbiology, Washington, DC.
11. Schweizer, E., and Hofmann, J. (2004) Microbial Type I Fatty Acid Synthases (FAS): Major Players in a Network of Cellular FAS Systems, *Microbiol. Mol. Biol. Rev. 68*, 501−517.
12. Caffrey, P., Bevitt, D. J., Staunton, J., and Leadley, P. F. (1992) Identification of DEBS 1, DEBS 2 and DEBS 3, the Multienzyme Polypeptides of the Erythromycin-Producing Polyketide Synthase from *Saccharopolyspora erythraea*, *FEBS Lett. 304*, 225−228.
13. Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J., and Katz, L. (1991) Modular Organization of Genes Required for Complex Polyketide Biosynthesis, *Science 252*, 675−679.
14. Shen, B. (2000) Biosynthesis of Aromatic Polyketides, *Top. Curr. Chem. 209*, 3−51.
15. Zawada, R. J., and Khosla, C. (1999) Heterologous Expression, Purification, Reconstitution and Kinetic Analysis of an Extended Type II Polyketide Synthase, *Chem. Biol. 6*, 607−615.
16. Crosa, J. H., and Walsh, C. T. (2002) Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria, *Microbiol. Mol. Biol. Rev. 66*, 223−249.
17. Keller, U., and Schauwecker, F. (2001) Nonribosomal Biosynthesis of Microbial Chromopeptides, *Prog. Nucleic Acid Res. Mol. Biol. 70*, 233−289.
18. Fischbach, M. A., and Walsh, C. T. (2006) Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms, *Chem. Rev. 106*, 3468−3496.
19. Garneau, S., Dorrestein, P. C., Kelleher, N. L., and Walsh, C. T. (2005) Characterization of the Formation of the Pyrrole Moeity

during Clorobiocin and Coumermycin A1 Biosynthesis, *Biochemistry 44*, 2770−2780.
20. Miller, D. A., Luo, L., Hillson, N., Keating, T. A., and Walsh, C. T. (2002) Yersinabactin Synthetase: A Four-Protein Assembly Line Producing the Nonribosomal Peptide/Polyketide Hybrid Siderophore of *Yersinia pestis*, *Chem. Biol. 9*, 333−344.
21. Walsh, C. T., O'Connor, S. E., and Schneider, T. L. (2003) Polyketide-nonribosomal Peptide Epothilone Antitumor Agents: The EpoA, B, C Subunits, *J. Ind. Microbiol. Biotechnol. 30*, 448−455.
22. Molnar, I., Aparicio, J. F., Haydock, S. F., Khaw, L. E., Schwecke, T., Konig, A., Staunton, J., and Leadley, P. F. (1996) Organisation of the Biosynthetic Gene Cluster for Rapamycin in *Streptomyces hygroscopicus*: Analysis of Genes Flanking the Polyketide Synthase, *Gene 169*, 1−7.
23. Du, L., Sanchez, C., and Shen, B. (2001) Hybrid Peptide-Polyketide Natural Products: Biosynthesis and Prospects toward Engineering Novel Molecules, *Metab. Eng. 3*, 78−95.
24. Finking, R., and Marahiel, M. A. (2004) Biosynthesis of Nonribosomal Peptides, *Annu. Rev. Microbiol. 58*, 453−488.
25. Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Omura, S. (2003) Complete Genome Sequence and Comparative Analysis of the Industrial Microorganism *Streptomyces avermitilis*, *Nat. Biotechnol. 21*, 526−531.
26. Chalut, C., Botella, L., de Sousa-D'Auria, C., Houssin, C., and Guilhot, C. (2006) The Nonredundant Roles of Two 4′-Phosphopantetheinyl Transferases in Vital Processes of *Mycobacteria*, *Proc. Natl. Acad. Sci. U.S.A. 103*, 8511−8516.
27. Wang, J., Soisson, S. M., et al. (2006) Platensimycin is a Selective FabF Inhibitor with Potent Antibiotic Propoerties, *Nature 441*, 358−361.
28. Gockhale, R. S., and Khosla, C. (2000) Role of Linkers in Communication between Protein Modules, *Curr. Opin. Chem. Biol. 4*, 22−27.
29. Tsuji, S. Y., Wu, N., and Khosla, C. (2001) Intermodular Communication in Polyketide Synthases: Comparing the Role of Protein-Protein Interactions to Those in Other Multidomain Enzymes, *Biochemistry 40*, 2326−2331.
30. Marshall, C. G., Burkart, M. D., Meray, R. K., and Walsh, C. T. (2002) Carrier Protein Recognition in Siderophore-Producing Nonribosomal Peptide Synthetases, *Biochemistry 41*, 8429−8437.
31. Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F., and Weissman, K. J. (2003) The Structure of Docking Domains in Modular Polyketide Synthases, *Chem. Biol. 10*, 723−731.
32. Hahn, M., and Stachelhaus, T. (2006) Harnessing the Potential of Communication-Mediating Domains for the Biocombinatorial Synthesis of Nonribosomal Peptides, *Proc. Natl. Acad. Sci. U.S.A. 103*, 275−280.
33. Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999) Dissecting and Exploiting Intermodular Communication in Polyketide Synthases, *Science 284*, 482−485.
34. 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.
35. Weissman, K. J., Hong, H., Popovic, B., and Meersman, F. (2006) Evidence for a Protein-Protein Interaction Motif on an Acyl Carrier Protein Domain from a Modular Polyketide Synthase, *Chem. Biol. 13*, 626−636.
36. Weber, T., Baumgartner, R., Renner, C., Marahiel, M. A., and Holak, T. A. (2000) Solution Structure of PCP, a Prototype for the Peptidyl Carrier Domains of Modular Peptide Synthetases, *Structure 8*, 407−418.
37. 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.
38. 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 9*, 277−287.
39. Wong, H. C., Liu, G., Zhang, Y.-M., Rock, C. O., and Zheng, J. (2002) The Solution Structure of Acyl Carrier Protein from *Mycobacterium tuberculosis*, *J. Biol. Chem. 277*, 15874−15880.
40. Tang, Y., Kim, C. Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006) The 2.7 Å Crystal Structure of a 194-kDa Homodimeric Fragment of the 6-Deoxyerythronolide B Synthase, *Proc. Natl. Acad. Sci. U.S.A. 103*, 11124−11129.
41. Kim, Y., and Prestegard, J. H. (1989) A Dynamic Model for the Structure of Acyl Carrier Protein in Solution, *Biochemistry 28*, 8792−8797.
42. Koglin, A., Mofid, M. R., Lohr, F., Schafer, B., Rogov, V. V., Blum, M. M., Mittag, T., Marahiel, M. A., Bernhard, F., and Dotsch, V. (2006) Conformational Switches Modulate Protein Interactions in Peptide Antibiotic Synthetases, *Science 312*, 273−276.
43. Walsh, C. T. (1979) *Enzymatic Reaction Mechanisms*, Freeman, San Francisco.
44. 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.
45. Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. M. (2002) Regeneration of Misprimed Nonribosomal Peptide Synthetases by Type II Thioesterases, *Proc. Natl. Acad. Sci. U.S.A. 99*, 14083−14088.
46. Jenni, S., Leibundgut, M., Maier, T., and Ban, N. (2006) Architecture of a Fungal Fatty Acid Synthase at 5 Å Resolution, *Science 311*, 1263−1267.
47. Maier, T., Jenni, S., and Ban, N. (2006) Architecture of Mammalian Fatty Acid Synthase at 4.5 Å Resolution, *Science 311*, 1258−1262.
48. Li, Q., Khosla, C., Puglisi, J. D., and Liu, C. W. (2003) Solution Structure and Backbone Dynamics of the Holo Form of the Frenolicin Acyl Carrier Protein, *Biochemistry 42*, 4648−4657.
49. Zornetzer, G. A., Fox, B. G., and Markley, J. L. (2006) Solution Structures of Spinach Acyl Carrier Protein with Decanoate and Stearate, *Biochemistry 45*, 5217−5227.
50. Sharma, A. K., Sharma, S. K., Surolia, A., Surolia, N., and Sarma, S. P. (2006) Solution Structures of Conformationally Equilibrium Forms of Hol-Acyl Carrier Protein (PfACP) from *Plasmodium falciparum* Provides Insight into the Mechanism of Activation of ACPs, *Biochemistry 45*, 6904−6916.
51. Frueh, D. P., Vosburg, D. A., Walsh, C. T., and Wagner, G. (2006) Determination of All NOEs in ^{1}H-^{13}C-Me-ILV-U-^{2}H-^{15}N Proteins with Two Time-Shared Experiments, *J. Biomol. NMR 34*, 31−40.
52. Frueh, D. P., Sun, Z. Y., Vosburg, D. A., Walsh, C. T., Hoch, J. C., and Wagner, G. (2006) Non-Uniformly Sampled Double-TROSY hNcaHN Experiments for NMR Sequential Assignments of Large Proteins, *J. Am. Chem. Soc. 128*, 5757−5763.
53. Lai, J. R., Fischbach, M. A., Liu, D. R., and Walsh, C. T. (2006) A Protein Interaction Surface in Nonribosomal Peptide Synthesis Mapped by Combinatorial Mutagenesis and Selection, *Proc. Natl. Acad. Sci. U.S.A. 103*, 5314−5319.
54. Lai, J. R., Fischbach, M. A., Liu, D. R., and Walsh, C. T. (2006) Localized Protein Interaction Surfaces on the EntB Carrier Protein Revealed by Combinatorial Mutagenesis and Selection, *J. Am. Chem. Soc. 128*, 11002−11003.
55. Zhou, Z., Lai, J. R., and Walsh, C. T. (2006) Interdomain Communication between the Thiolation and Thioesterase Domains of EntF Explored by Combinatorial Mutagenesis and Selection, *Chem. Biol. 13*, 869−879.
56. Gehring, A., Mori, A., and Walsh, C. T. (1998) Reconstitution and Characterization of the *Escherichia coli* Enterobactin Synthetase from EntB, EntE, and EntF, *Biochemistry 37*, 2648−2659.
57. Gehring, A., Bradley, K. A., and Walsh, C. T. (1997) Enterobactin Biosynthesis in *Escherichia coli*: Isochorismate Lyase (EntB) Is a Bifunctional Enzyme That Is Phosphopantetheinylated by EntD and Then Acylaed by EntE Using ATP and 2,3-Dihydroxybenzoate, *Biochemistry 36*, 8495−8503.
58. Roche, E. D., and Walsh, C. T. (2003) Dissection of the EntF Condensation Domain Boundary and Active Site Residues in Nonribosomal Peptide Synthesis, *Biochemistry 42*, 1334−1344.
59. Ruznak, F., Liu, J., Quinn, N., Berchtold, G. A., and Walsh, C. T. (1990) Subcloning of the Enterobactin Biosynthetic Gene *entB*: Expression, Purification, Characterization, and Substrate Specificity of Isochorismatase, *Biochemistry 29*, 1425−1435.
60. Drake, E. J., Nicolai, D. A., and Gulick, A. M. (2006) Structure of the EntB Multidomain Nonribosomal Peptide Synthetase and Functional Analysis of Its Interaction with the EntE Adenylation Domain, *Chem. Biol. 13*, 409−419.
61. Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000) Aminoacyl-SNACs as Small-Molecule Substrates for the Condensation Domains of Nonribosomal Peptide Synthetases, *Chem. Biol. 7*, 765−772.

62. Shaw-Reid, C. A., Kelleher, N. L., Losey, H. C., Gehring, A. M., Berg, C., and Walsh, C. T. (1999) Assembly Line Enzymology by Multimodular Nonribosomal Peptide Synthetases: The Thioesterase Domain of *E. coli* EntF Catalyzes both Elongation and Cyclolactonization, *Chem. Biol. 6*, 385−400.
63. Weiss, F. A., Watanabe, C. K., Zhong, A., Goddard, A., and Sidhu, S. S. (2000) Rapid Mapping of Protein Functional Epitopes by Combinatorial Alanine Scanning, *Proc. Natl. Acad. Sci. U.S.A. 97*, 8950−8954.
64. Morrison, K. L., and Weiss, G. A. (2001) Combinatorial Alanine-Scanning, *Curr. Opin. Chem. Biol. 5*, 302−307.
65. Gregoret, L. M., and Sauer, R. T. (1993) Additivity of Mutant Effects Assessed by Binomial Mutagenesis, *Proc. Natl. Acad. Sci. U.S.A. 90*, 4246−4250.
66. Johnson, M. A., Peti, W., Herrmann, T., Wilson, I. A., and Wuthrich, K. (2006) Solution structure of Asl1650, an acyl carrier protein from *Anabaena* sp. PCC 7120 with a variant phosphopantetheinylation-site sequence, *Protein Sci. 15*, 1030−1041.
67. 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.
68. 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*, 923−936.
69. Lee, T. S., Khosla, C., and Tang, Y. (2005) Orthogonal protein interactions in spore pigment producing and antibiotic producing polyketide synthases, *J. Antibiot. 58*, 664−667.
70. Tang, Y., Lee, T. S., Kobayashi, S., and Khosla, C. (2003) Ketosynthases in the Initiation and Elongation Modules of Aromatic Polyketides Synthases Have Orthogonal Acyl Carrier Protein Specificity, *Biochemistry 42*, 6588−6596.
71. Tang, Y., Lee, T. S., and Khosla, C. (2004) Engineered Biosynthesis of Regioselectively Modified Aromatic Polyketides using Modular Polyketide Synthases, *PLoS Biol. 2*, 227−238.
72. Vaillancourt, F. H., Yeh, E., Vosburg, D. A., O'Connor, S. E., and Walsh, C. T. (2005) Cryptic Chlorination by a Non-Haem Iron Enzyme during Cyclopropyl Amino Acid Biosynthesis, *Nature 436*, 1191−1194.
73. Meadows, E. S., and Khosla, C. (2001) In Vitro Reconstitution and Analysis of the Chain Initiation Enzymes of the R1128 Polyketide Synthase, *Biochemistry 40*, 14855−14861.
74. Pan, H., Tsai, S., Meadows, E. S., Miercke, L. J., Keatinge-Clay, A. T., O'Connell, J., Khosla, C., and Stroud, R. M. (2002) Crystal Structure of the Priming β-Ketosynthase from the R1128 Polyketide Biosynthetic Pathway, *Structure 10*, 1559−1568.
75. Lee, T. S., Khosla, C., and Tang, Y. (2005) Engineered Biosynthesis of Aklanonic Acid Analogues, *J. Am. Chem. Soc. 127*, 12254−12262.
76. Izumikawa, M., Cheng, Q., and Moore, B. S. (2006) Priming Type II Polyketide Synthases via a Type II Nonribosomal Peptide Synthetase Mechanism, *J. Am. Chem. Soc. 128*, 1428−1429.
77. Couch, R., O'Connor, S. E., Seidle, H., Walsh, C. T., and Parry, R. (2004) Characterization of CmaA, an Adenylation-Thiolation Didomain Enzyme Involved in the Biosynthesis of Coronatine, *J. Bacteriol. 186*, 35−42.
78. Vaillancourt, F. H., Yeh, E., Vosburg, D. A., Garneau-Tsodikova, S., and Walsh, C. T. (2006) Nature's Inventory of Halogenation Catalysts: Oxidative Strategies Predominate, *Chem. Rev. 106*, 3364−3378.
79. Edwards, D. J., Marquez, B. L., Nogle, L. M., McPhail, K., Goeger, D. E., Roberts, M. A., and Gerwick, W. H. (2004) Structure and Biosynthesis of the Jamaicamides, New Mixed Polyketide-Peptide Neurotoxins from the Marine Cyanobacterium *Lyngbya majuscula*, *Chem. Biol. 11*, 817−833.
80. Rahman, A. S., Hothersall, J., Crosby, J., Simpson, T. J., and Thomas, C. M. (2005) Tandemly Duplicated Acyl Carrier Proteins, Which Increase Polyketide Antibiotic Production, Can Apparently Function Either In Parallel of In Series, *J. Biol. Chem. 280*, 6399−6408.
81. Chang, Z., Sitachitta, N., Rossi, J. V., Roberts, M. A., Flatt, P. M., Jia, J., Sherman, D. H., and Gerwick, W. H. (2004) Biosynthetic Pathway and Gene Cluster Analysis of Curacin A, an Antitubulin Natural Product from the Tropical Marine Cyanobacterium *Lyngbya majuscule*, *J. Nat. Prod. 67*, 1356−1367.
82. Meier, J. L., Mercer, A. C., Rivera, H., Jr., and Burkart, M. D. (2006) Synthesis and Evaluation of Bioorthogonal Pantetheine Analogues for In Vivo Protein Modification, *J. Am. Chem. Soc. 128*, 12174−12184.
83. Clark, K. M., Mercer, A. C., LaClair, J. J., and Burkart, M. D. (2005) In Vivo Reporter Labeling of Proteins via Metabolic Delivery of Coenzyme A, *J. Am. Chem. Soc. 127*, 11234−11235.
84. Zhou, Z., Cironi, P., Lin, A. J., Hrvatin, S., Silver, P. A., Golan, D. E., Walsh, C. T., and Yin, J. (2006) Genetically Encoded Short Peptide Tags for Orthogonal Protein Post-Translational Modification by Sfp and AcpS Phosphopantetheinyl Transferases (manuscript in preparation).
85. Yin, J., Straight, P. D., McLoughlin, S. M., Buckett, P. D., Wessling-Resnick, M., Golan, D. E., and Walsh, C. T. (2005) Single-cell FRET Imaging of Transferrin Receptor Trafficking Dynamics by Sfp-Catalyzed, Site-Specific Protein Labeling, *Chem. Biol. 12*, 999−1006.
86. Yin, J., Straight, P. D., Hrvatin, S., Dorrestein, P. C., Bumpus, S., Jao, C., Kelleher, N. L., Kolter, R., and Walsh, C. T. (2006) Genome-Wide Throughput Profiling of Natural Product Biosynthetic Gene Clusters by Phage Display (submitted for publication at *Chem. Biol.*).
87. Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999) Aminoacyl-CoAs as Probes of Condensation Domain Selectivity in Nonribosomal Peptide Synthesis, *Science 284*, 486−489.
88. Schneider, T. L., Walsh, C. T., and O'Connor, S. E. (2002) Utilization of Alternate Substrates by the First Three Modules of the Epothilone Synthetase Assembly Line, *J. Am. Chem. Soc. 124*, 11272−11273.
89. Cane, D. E., Kudo, F., Kinoshita, K., and Khosla, C. (2002) Precursor-Directed Biosynthesis: Biochemical Basis of the Remarkable Selectivity of the Erythromycin Polyketide Synthase toward Unsaturated Triketides, *Chem. Biol. 9*, 131−142.
90. Yin, Y., Lu, H., Khosla, C., and Cane, D. E. (2003) Expression and Kinetic Analysis of the Substrate Specificity of Modules 5 and 6 of the Picromycin/Methymycin Polyketide Synthase, *J. Am. Chem. Soc. 125*, 5671−5676.
91. Wu, N., Tsuji, S. Y., Cane, D. E., and Khosla, C. (2001) Assessing the Balance between Protein-Protein Interactions and Enzyme-Substrate Interactions in the Channeling of Intermediates between Polyketide Synthase Modules, *J. Am. Chem. Soc. 123*, 6465−6474.
92. Aguilar, R. C., and Wendland, B. (2003) Ubiquitin: Not Just for Proteasomes Anymore, *Curr. Opin. Cell Biol. 15*, 184−190.

BI061979P