

# Probing the Compatibility of Type II Ketosynthase–Carrier Protein Partners

Andrew S. Worthington,<sup>[a]</sup> Gene H. Hur,<sup>[a]</sup> Jordan L. Meier,<sup>[a]</sup> Qian Cheng,<sup>[b]</sup> Bradley S. Moore,<sup>[b, c]</sup> and Michael D. Burkart<sup>\*[a]</sup>

Drug discovery often begins with the screening of large compound libraries to identify lead compounds. Recently, the enzymes that are involved in the biosynthesis of natural products have been investigated for their potential to generate new, diverse compound libraries. There have been several approaches toward this end, including altering the substrate specificities of the enzymes involved in natural product biosynthesis and engineering functional communication between enzymes from different biosynthetic pathways. While there exist assays to assess the substrate specificity of enzymes involved in these pathways, there is no simple method for determining whether enzymes from different synthases will function cooperatively to generate the de-

sired product(s). Herein we report a method that provides insight into both substrate specificity and compatibility of protein–protein interactions between the acyl carrier protein (ACP) and ketosynthase (KS) domains involved in fatty acid and polyketide biosynthesis. Our technique uses a one-pot chemoenzymatic method to generate post-translationally modified ACPs that are capable of covalently interacting with KS domains from different biosynthetic systems. The extent of interaction between ACPs and KSs from different systems is easily detected and quantified by a gel-based method. Our results are consistent with previous studies of substrate specificity and ACP–KS binding interactions and provide new insight into unnatural substrate and protein interactions.

## Introduction

Fatty acid and polyketide biosynthesis are evolutionarily related pathways that use similar mechanisms to synthesize their products.<sup>[1,2]</sup> In both pathways small acyl carrier proteins (ACPs) serve as a central covalent template to which biosynthetic intermediates are tethered while iterative cycles of loading, condensation, and processing are performed by accessory enzymes such as acyltransferase (AT), ketosynthase (KS), and ketoreductase (KR) domains.<sup>[3–5]</sup> In type I fatty acid synthase (FAS) and polyketide synthase (PKS) systems these enzyme activities are all housed on a single multidomain “megasynthase,” whereas in type II FAS and PKS systems each accessory enzyme is a discrete polypeptide that must interact *in trans* with the growing ACP-tethered intermediate.<sup>[1,2,4–8]</sup>

ACP-bound acyl intermediates are tethered through a thioester bond that is formed by the terminal thiol of a post-translationally introduced 4'-phosphopantetheine arm. This 4'-phosphopantetheine prosthetic group is introduced by the action of a phosphopantetheinyltransferase (PPTase) enzyme, which converts *apo*-ACP to *holo*-ACP by the transfer of the 4'-phosphopantetheine from coenzyme A (CoA) to a conserved serine residue of ACP.<sup>[9]</sup> Recently our group<sup>[10–12]</sup> and others<sup>[13,14]</sup> have shown that promiscuous PPTases, most notably the surfactin PPTase Sfp,<sup>[15]</sup> are highly tolerant of modification to the terminal portion of CoA and will transfer a variety of 4'-phosphopantetheine analogues from CoA analogues to produce *crypto*-ACPs. *Crypto*-ACP refers to an ACP in which the terminal thiol moiety is replaced or “hidden” by an unnatural modification.<sup>[11]</sup> This ability to modify ACPs with virtually any substitution at the thiol terminus of the 4'-phosphopantetheine group has been used for the proteomic identification of PKS and FAS sys-

tems in crude cell lysate<sup>[16,17]</sup> and to provide essential insights into the substrate selectivity of these enzymes.<sup>[18]</sup>

One further area in which this technology has proven useful is in probing the structural interactions of carrier-protein-mediated biosynthesis. In type II FAS and PKS biosynthesis, tight yet transient interactions between the loaded ACP and each of its partner enzymes are essential for efficient biosynthetic productivity. The transient and noncovalent nature of these protein–protein interactions has hindered attempts to study the individual steps of these biosynthetic processes by using traditional structural methods. To help alleviate this problem we recently developed a method to trap ACPs during their interaction with KS domains.<sup>[19]</sup> In natural type II FAS and PKS systems, KS domains are responsible for the condensation reaction in which the growing acyl chain of the FAS or PKS intermediate is first transferred to the nucleophilic cysteine of the KS, followed by decarboxylation and Claisen condensation of a malonyl-ACP

[a] A. S. Worthington, G. H. Hur, J. L. Meier, Prof. M. D. Burkart

Department of Chemistry and Biochemistry

University of California–San Diego

9500 Gilman Drive, La Jolla, California 92093-0358 (USA)

E-mail: mburkart@ucsd.edu

[b] Dr. Q. Cheng, Prof. B. S. Moore

Scripps Institution of Oceanography, University of California–San Diego

9500 Gilman Drive, La Jolla, California 92038-0204 (USA)

[c] Prof. B. S. Moore

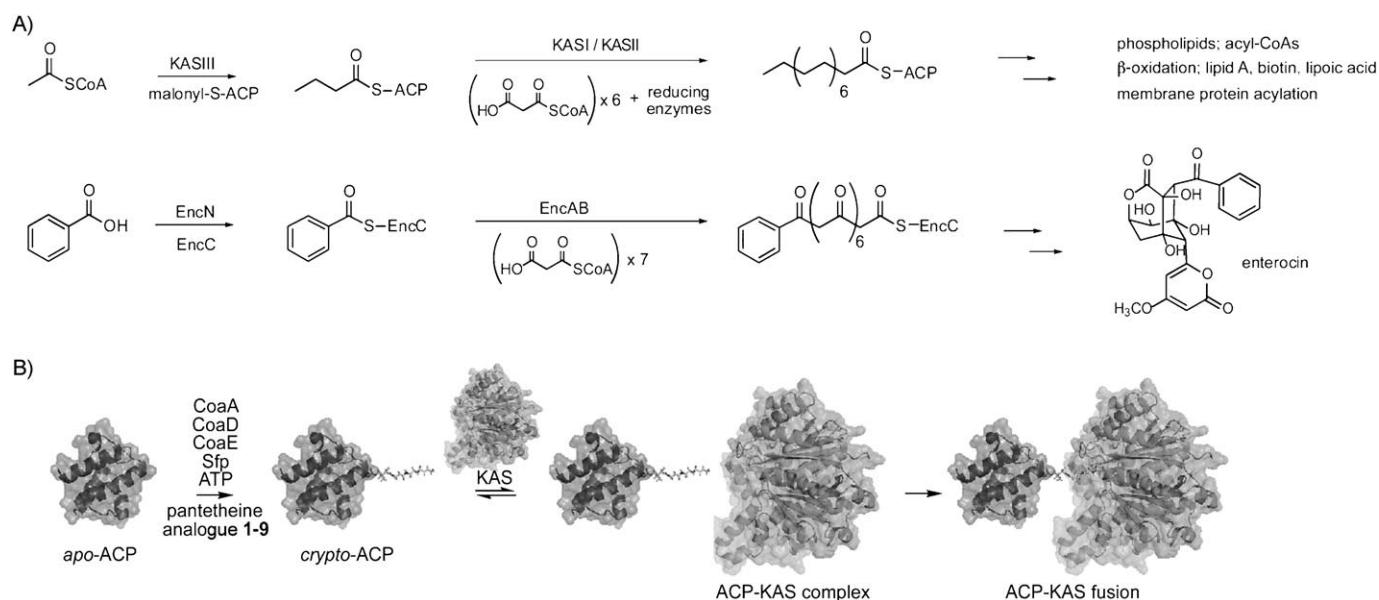
Skaggs School of Pharmacy and Pharmaceutical Sciences

University of California–San Diego

9500 Gilman Drive, La Jolla, California 92038-0657 (USA)



Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author.



**Figure 1.** A method to probe interactions between modified ACPs and KSs of type II FAS and PKS pathways. A) A comparison of type II biosyntheses between FAS in *E. coli* and enterocin biosynthesis in *S. maritimus*. B) Apo-ACP is incubated with CoaA, CoaD, CoaE, and Sfp in the presence of ATP and pantetheine analogues **1–9** to generate crypto-ACP. Crypto-ACP reacts with the active-site cysteine of the ketosynthase to form ACP–KAS fusion product.

intermediate with the KS-bound acyl-enzyme intermediate.<sup>[20–22]</sup> To study this transient interaction we used a chemoenzymatic synthesis of CoA analogues combined with post-translational modification by the PPTase Sfp to produce *crypto*-ACPs in which nonhydrolyzable electrophilic probes are tethered in place of the natural thiol terminus of *holo*-ACP. This reactive moiety can then interact with the nucleophilic active-site cysteine of a KS domain, which results in the formation of a covalent crosslink. This interaction can be detected by running the one-pot reaction on SDS-PAGE and observing a gel shift. Tighter interactions between the ACP and KS yield a greater signal, while a diminished or negligible signal is an indication that either the two proteins are failing to associate, or that the reactive CoA analogue is effectively competing with the crosslinking reaction for modification of the KS active site.

In addition to its utility in isolating covalently crosslinked ACP–KS complexes for crystallographic studies,<sup>[23]</sup> this crosslinking procedure presents a novel method for studying the step-wise compatibility of ACP and partner enzymes that originate from alternate biosynthetic systems. Currently there is no simple procedure for determining whether an ACP and a KS from two different biosynthetic systems will possess the requisite protein–protein interactions to insure tight binding and efficient processing of biosynthetic intermediates; this shortcoming has presented an obstacle in the rational manipulation of modular synthases for the purposes of generating novel combinatorial compound libraries.<sup>[24–26]</sup> By comparing the extent of crosslinking between an ACP and KS we can demonstrate how well the KS will accept alternate carrier proteins from other organisms and/or biosynthetic pathways.

Here we present a full study of ACP–KS crosslinking in which we apply nine new electrophilic 4'-phosphopantetheine analogues towards crosslinking studies of three systems: KASI–

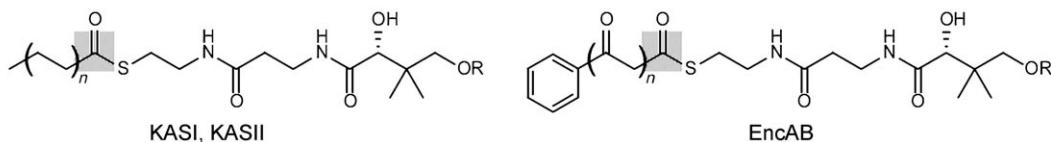
ACP and KASII–ACP from *E. coli* type II FAS, and EncAB–EncC, a KS–ACP pair from the *S. maritimus* type II enterocin PKS (Figure 1).<sup>[27–29]</sup> A gel-based assay is used to probe and quantify the interactions between both natural and unnatural ACP–KS partners by using our panel of reactive pantetheine analogues; the results reflect the specificity of a KS for its natural ACP partner. In addition we verify the active-site specificity of our crosslinking experiments by performing MALDI and LC–MS/MS analyses on in-gel digests of crosslinked proteins. Finally we use our panel of electrophilic pantetheine analogues to correlate substrate specificity with the structure of the electrophilic 4'-phosphopantetheine arm of the *crypto*-ACP (Scheme 1). This study represents the first application of this technology to a medicinally relevant type II PKS biosynthetic system, and our findings accurately reflect the known differences in both binding affinities and substrate specificity among KS domains and provide new insight into the interaction between ACP and KS domains in type II FAS and PKS systems.

## Results and Discussion

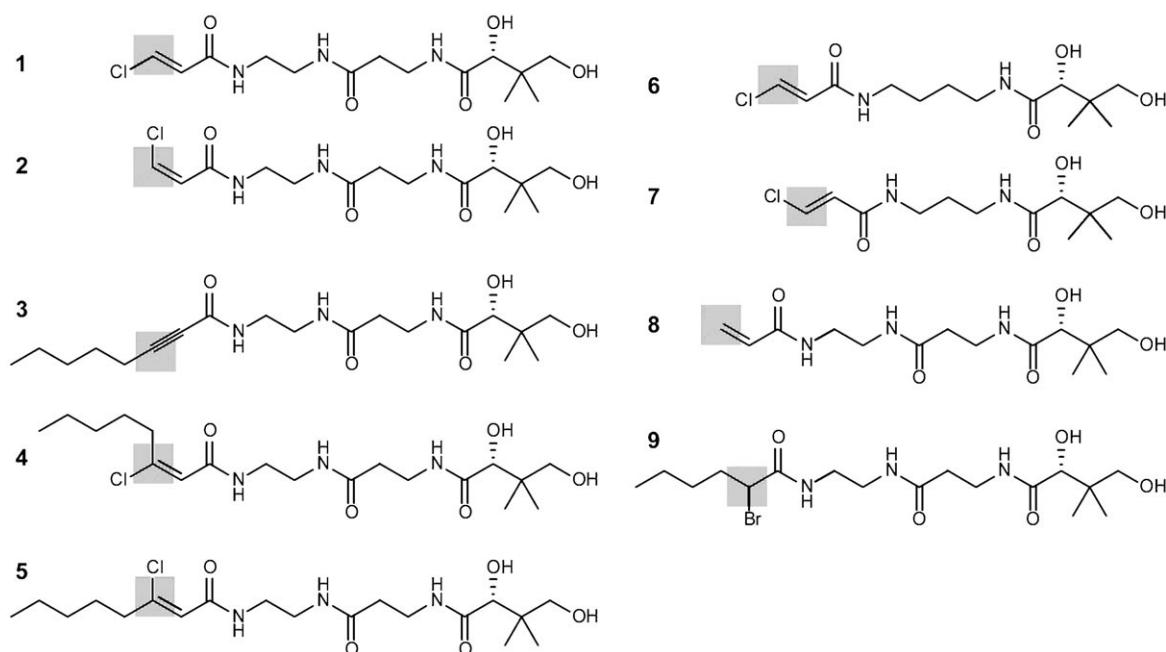
### Design and synthesis of reactive pantetheine analogues for crosslinking

This study utilizes a previously reported chemoenzymatic method in which CoaA, CoaD, and CoaE are used to convert pantetheine analogues **1–9** into CoA analogues in one pot, before transfer of the 4'-phosphopantetheine group to the ACP by the PPTase Sfp (Figure 1B).<sup>[10,19]</sup> In our initial report, we noticed different reactivities of our crosslinking pantetheine analogues with regards to the stereochemistry of the leaving group.<sup>[19]</sup> These differences are likely a reflection of binding-pocket–residue interactions. The enzymes responsible for the

## A) Natural pantetheine substrates



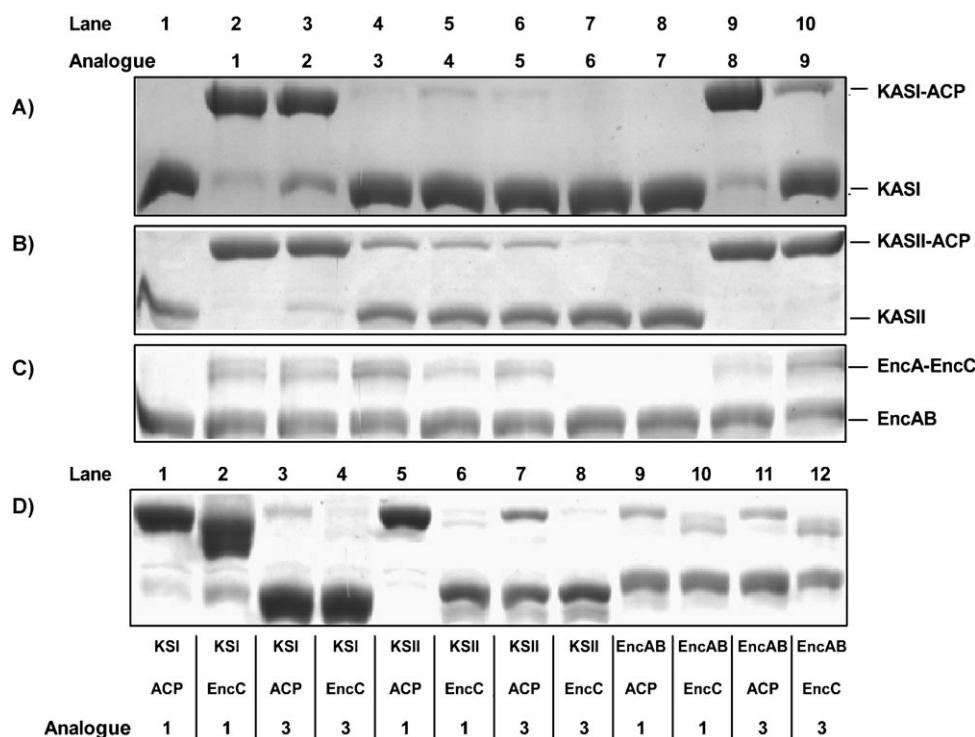
## B) Pantetheine analogues



**Scheme 1.** Pantetheine analogues as probes for the substrate specificity of KS domains in FAS and PKS. The site of attack by the active-site cysteine of KS is highlighted in yellow. A) Natural pantetheine substrates for KS domains KASI and KASII of *E. coli* FAS and EncAB of *S. maritimus* PKS. Substrates of length  $n=0-6$  are bound to the ACP at R. B) Pantetheine analogues 1–9 were used to mimic substrates of varying lengths and degrees of oxidation.

biosynthesis of the *S. maritimus* polyketide enterocin belong to the type II polyketide synthase (PKS) family.<sup>[28,29]</sup> The ketosynthase chain length factor (KS-CLF) EncAB must recognize not only benzoate and malonate that is bound to the carrier protein EncC, but also the growing product bound to EncC; this growing product is transacylated to the active-site cysteine prior to condensation with malonyl-EncC. *E. coli* FAS ketosynthases KASI and KASII also recognize their growing fatty acyl chains in addition to recognizing malonate that is bound to ACP (Figure 1A). Although they have overlapping substrate specificities, KASI and KASII have shown slightly different activities towards a panel of fatty acid substrates. KASII shows greater specific activity with all substrates from hexanoyl-ACP (C6) to hexadecanoyl-ACP (C16). Most notably, KASII shows significantly more activity than KASI with myristoyl(14:0) and palmitoleyl-ACP(16:1) as substrates.<sup>[30]</sup> The binding pocket of KASI is slightly larger and thus more receptive to the longer substrates of the latter stages of fatty acid synthesis; structural and modeling studies have shown that the binding pocket of KASI is too shallow to hold the C18:1 product of the elongation of C16:1.<sup>[31]</sup> Other differences in substrate specificity between the two ketosynthases are thought to be due to the identities of residues lining the substrate binding pockets.<sup>[32]</sup>

Pantetheine analogues 1–9 can be thought of as incorporating two important components for the purpose of this study: 1) an affinity tag of varying stereochemistry and electrophilicity and 2) an alkyl chain that is designed to mimic a variety of substrate sizes (Scheme 1). Because terminal chloroacrylamides<sup>[33]</sup> proved to be the most efficient ACP–KS crosslinkers in our original study, we first examined the subtle differences of this reactive electrophile in depth. In order to examine the influence of electrophile stereochemistry on crosslinking we synthesized both terminal (1–2), and long-chain (4–5) chloroacrylamides of varying Z/E stereochemistry. We also investigated the effect of varying the placement of the Michaeli-accepting double bond of the chloroacrylamide by synthesizing compound 6, which places it at the same chain length as the naturally occurring 4'-phosphopantetheine thioester, and compound 7, which places the electrophilic carbon one carbon shorter than the naturally occurring thioester. To understand the effect of the chlorine group on crosslinking better, we synthesized 8, a non-chloro-containing terminal acrylamide. Finally we examined the reactivity of two other long-chain electrophile-containing pantetheine analogues,  $\alpha,\beta$ -unsaturated alkyne 3 and racemic  $\alpha$ -bromoacetamide 9.<sup>[34]</sup> In terms of acyl chain length, these analogues can be grouped into three sub-



**Figure 2.** Application of the method for KS–ACP formation to *E. coli* FAS and *S. maritimus* PKS proteins by using nine reactive pantetheine analogues. SDS-PAGE analysis of one-pot reactions to form KS–ACP complexes by using the panel of analogues and three combinations of KS, ACP proteins: A) *E. coli* KASI and ACP B) *E. coli* KASII and ACP C) *S. maritimus* EncAB and EncC. D) SDS-PAGE analysis of one-pot reactions to form KS–ACP complexes by using the shorter analogue 1 and longer analogue 3 to compare substrate specificity and binding compatibility between the KSs of type II biosynthesis and both natural and unnatural ACP partners.

groups: those with acyl chains that are shorter than acylated 4'-phosphopantetheine (6–7), those with acyl-chains that roughly correspond to a short acyl chain (1–2, 8), and those that incorporate a longer acyl chain (3–5, 9).

After overproduction and purification of the *apo*-forms of the carrier proteins *E. coli* AcpP (from type II FAS) and *S. maritimus* EncC (from type II enterocin PKS) the one-pot chemoenzymatic method was applied with analogues 1–9 to post-translationally modify each ACP and produce *crypto*-ACPs. Each *crypto*-ACP then was reacted with its natural partner KS. Two

of the more potent *crypto*-ACP forms, which were derived from compounds 1 and 3 were also incubated with a KS from a different organism and biosynthetic system. In almost all of the cases tested, there was some interaction between *crypto*-ACP and KS (Figure 2). This interaction was detected as a gel shift on SDS-PAGE; the degree of crosslinking was quantified by using the gel-analyzing function of the program ImageJ (Table 1).<sup>[35]</sup> For all natural systems, crosslinking was confirmed by MALDI MS/MS analyses (Figure 3).<sup>[19]</sup>

#### Functional characteristics of linker region and length of 4'-phosphopantetheine arms for ACP-KS crosslinking

Inspection of the results of the gel-shift assay (Figure 2) indicate several important crosslinking characteristics for the pantetheine analogues in our panel. An initial observation is that our results reaffirm the importance

of the second amide bond of pantetheine for enzyme recognition.<sup>[36]</sup> Analogues 6 and 7, which lack this important recognition element allowed only minimal interaction with a ketosynthase domain in the KASII-ACP system and showed no detectable interaction in other systems, despite the optimal placement of the electrophilic chloroacrylamide in 6. This is not due simply to a lack of ACP loading by 6 and 7 because MALDI-MS

#### Enc C

MKHHHHHHHHH GGLVPRGSH GSEFLSSDQL LQILSECSDG EQALAQGDDI LDREFEVLYG DSVLLETSA QLKHYDIDI PEDVIRDLKT PRAVLEYVNV PASRN

#### Enc A

MGSHHHHHHHH SSGLVPRGSH MMSRHVVITG LSVLAPGSTD TEGFWKMITA GRSAIRRITA FDPTPFRSQV AGEVLDLPLA CGFSRREVRR LDRASLLAVA CARRAVAAEG ITSGTSIDPG RIGVSVGNAV GSATSIENEY VVLSDEGRQW LVDQKYQSPH LFDYFVPGSI ARETAWAAGA EGPVSVISAG CTSGIDALGH ACQLIQEGSA DVMVAGATDA PVTPIAVACF DAIKATTRRN DDPEHAPRPF DRTDRGFAIA EGAAMLVLED SEHARRRGAD IYGTITGYAA HSNAYHMTGL SRDGNEMARA ITTALDEART DPTAVGYINA HGSGTKQNDL HETEAFKKAL GEHAYNVPVS SIKSMIGHSL GAIGAIEVAT CALAIRNGVV PPTANLHEPD PDCDLVDYVPL EARDVLDVV LTVGSGFGGF QSAMIVEKWR GGHV

**Figure 3.** Compiled MS/MS data of gel-excised EncA–EncC complex that was digested with pepsin, trypsin, or chymotrypsin. Residues that are highlighted in yellow and blue represent regions of the proteins that were identified by their peptide fragments. Peptides that contain the active site of EncC (S62, highlighted in green) and the active site of EncA (C191, highlighted in purple) were not isolated.

**Table 1.** Quantification of the degree of crosslinking between KS and ACP using the program ImageJ. A) Comparison of crosslinking using the panel of analogues with natural KS-ACP partners. B) Comparison of natural versus unnatural KS-ACP partners by using analogues 1 and 3. The pantetheine analogue used in each experiment is designated in bold.

	1	2	3	4	5	6	7	8	9
KASI-ACP	87	74	4	5	3	0	0	92	19
KASII-ACP	99	92	27	20	17	4	2	100	100
EncAB-EncC	42	43	56	23	35	0	0	24	59
B)									
KASI-ACP	95		7						
KASI-EncC	89		1						
KASII-ACP	99		26						
KASII-EncC	1		1						
EncAB-ACP	34		28						
EncAB-EncC	32		49						

analysis of the one-pot chemoenzymatic reactions before the addition of KS shows the expected mass change for formation of *crypto*-ACPs in both cases (Figure S1 in the Supporting Information). Among these two poor crosslinkers, placement of the reactive electrophile does appear to play a role in ACP–KS crosslinking activity, as evidenced by the slight preference for compound **6** (same length as the natural substrate) over compound **7** (one carbon shorter than the natural substrate) in the KASII–ACP system. Overall placement of the electrophilic trap at a distance greater than that of the position of the natural thioester from the ACP was most effective for crosslinking; this is an indication of the importance of the length of the 4'-phosphopantetheine arm in modulating carrier protein interactions with accessory enzymes. The identity of the electrophilic functionality at the site of attack of the pantetheine analogue also proved very important for ACP–KS crosslinking efficiency, with the most reactive electrophile,  $\alpha$ -bromoacetamide **9**, showing extreme reactivity with KASII and EncAB systems.

#### Correlation of 4'-phosphopantetheine structure with ACP–KS crosslinking efficiency and substrate specificity

Crosslinking between a *crypto*-ACP and KS domain is dependent upon the ability of the KS to both recognize the surface residues of the carrier protein and accommodate and react with the reactive substrate analogue of the modified 4'-phosphopantetheine arm. A greater crosslinking signal, which is detected as a gel shift on SDS-PAGE, indicates a tighter interaction between the ACP and KS. A weaker signal is indicative of either a lack of favorable binding interactions between the partner ACP and KS, a lack of binding interactions between the 4'-phosphopantetheine arm of the *crypto*-ACP and the active site of the KS, or an extremely reactive electrophilic pantetheine analogue that blocks the KS active site before transfer to the ACP. For the purposes of this study, we chose electrophilic warheads of tempered reactivity (acrylamides and acetyl-enamides being some of the least reactive Michael acceptors, haloacetamides being the least reactive  $\alpha$ -halocarbonyl compounds) to avoid the latter problem. Furthermore, because the surface residues of our *crypto*-ACPs are invariable through the full spectrum of our pantetheine analogues, crosslinking efficiency of an ACP and KS from the same system (e.g., *E. coli* type II FAS) should be mostly a function of the ability of the KS active site to interact favorably with the structure of the *crypto*-ACP's unnatural 4'-phosphopantetheine arm.

For the functional assessment of these interactions we performed thirty one-pot ACP–KS crosslinking reactions by applying each pantetheine analogue **1–9**, along with a DMSO control to each of the three systems that were tested (Figure 2A–C). Our results correlate well with previous studies of substrate specificity in ACP–KS systems, and also allow us to draw new conclusions that are yet to be verified. For example, previous studies of ACP–KS dissociation constants in *E. coli* type II FAS have estimated the  $K_d$  between KASII and C14-ACP as 70 nm, versus a weaker interaction of 180 nm between KASI and C14-ACP. Both interactions are stronger than that of the ACP–KS that is responsible for the biosynthesis of the type II PKS prod-

uct actinorhodin, which is estimated to lie in the 1–10  $\mu$ m range.<sup>[37]</sup> The dissociation constant between EncAB and EncC is expected to be similar to that of the actinorhodin system because both biosyntheses are carried out by related *Streptomyces* bacteria. This overall weaker interaction of ACP and KS partners in type II PK biosynthesis is reflected in the relative crosslinking efficiencies of EncC–EncAB when compared to the ACP and KS from *E. coli* type II FAS. Of the five most active compounds, **1–3**, **8**, and **9**, only two crosslink EncAB with an efficiency of greater than 50%, whereas three of five crosslink AcpP–KASI with more than 70% efficiency, and four of five crosslink AcpP–KASII with more than 90% efficiency; there is a direct correlation between crosslinking efficacy and binding of ACP and KS (Table 1A).

In addition to such broad intersystem insights, careful analysis of the crosslinking results provides knowledge of the active sites of the KS enzymes themselves. For example, comparing the results of testing our panel of analogues on the natural KASI–ACP and KASII–ACP systems shows that in addition to its tighter binding, KASII acts more efficiently than KASI on larger, longer substrate mimics. The data generated from the alkyl-chain-incorporating substrate mimics (compounds **3**, **4**, **5**, and **9**) indicate a clear preference (greater than fourfold) by KASII over KASI for larger substrates. This supports the previously noted study by Edwards et al., which showed that KASII acts on larger substrates C14:0 and C16:1 at least twofold more efficiently than KASI.<sup>[38]</sup> In comparison, the two KS domains had similar activity (> 70% crosslinking) towards the smaller substrates, compounds **1**, **2**, and **8**. All three of these compounds are similar in size to natural FAS substrates at the beginning of elongation, but compounds **3–5** and **9** are more similar to intermediates in FAS elongation. The slight preference for compound **8** might be a result of a narrow active-site channel, which has trouble accepting the bulky halide incorporated by **1** and **2**; this hypothesis is supported by the preference for compound **3** over compounds **4** and **5** in the KASII–ACP system.

The substrate-binding pocket of the ketosynthase EncAB needs to be larger and bulkier than that of either *E. coli* FAS ketosynthase. The active site of EncAB holds substrates that contain up to seven unreduced ketide units followed by the benzoate starter unit and produces a final product of the condensation reactions that is larger than those produced by either KASI or KASII. The results of our crosslinking studies with compounds **3–5** and **9** reflect this difference. Of the three systems that were tested in our panel, EncAB showed the greatest activity towards octyl-chain-incorporating compounds **3–5**. EncAB did not have the greatest reactivity of the three systems with compound **9** (KASII was extremely reactive with this analogue), but hexanoyl bromoacetamide **9** show the greatest crosslinking of EncC–EncAB compared with the rest of the panel. These results are likely due to the inherent volume of the substrate-binding pocket. In addition, EncAB does not interact as well with EncC-bound compound **8** as it does with EncC-bound compounds **1** or **2**. Even among these starter unit analogues, having a large, bulky leaving group on the pantetheine analogue seems to render it more effective at binding

in the expansive KS active site. However, unlike in the smaller active sites of the *E. coli* ketosynthases, there is no preference for the *E* isomer **1** over the *Z* isomer **2** in these less bulky crosslinking agents. This lack of a clear stereochemical preference for EncAB crosslinking with **1** and **2**, along with the complete crosslinking that is observed in KASII, suggests that the lack of complete crosslinking by racemic bromoacetamide **9** in the Enc system is due more to the weak protein–protein interactions of EncC–EncAB than to an intrinsic preference for one stereoisomer of **9**. To verify this will require the synthesis of enantiopure pantetheine derivatives. Alternatively, the lack of complete crosslinking in the Enc system could be due to competition for the active site between non-carrier-protein-bound pantetheine analogues and *crypto*-ACP, a phenomenon that was not observed in the tight-binding type II FAS systems.

#### Verification of ACP modification and ACP–KS crosslinking site selectivity

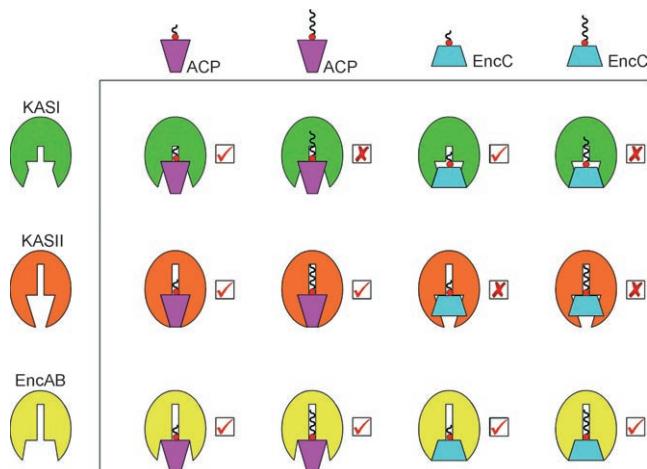
To verify the compatibility of analogues **1–9** with our one-pot chemoenzymatic carrier protein modification protocol we used MALDI-MS to monitor the mass increase upon post-translational modification of *apo* to *crypto*-ACPs (Figure S1). In our initial report the site-specific nature of our crosslinking technique on *E. coli* AcP and both KASI and KASII was demonstrated by using data from MALDI MS/MS analyses of in-gel-digested complexes.<sup>[19]</sup> We extended this approach to the *S. maritimus* ACP EncC and KS EncAB, which are involved in the biosynthesis of the polyketide enterocin (Figure 3).<sup>[27,29]</sup> In this experiment, EncC was post-translationally modified by a CoA entity that was derived from compound **3**. *Crypto*-EncC was then allowed to react with EncAB. The one-pot reaction was run on multiple lanes of SDS-PAGE; the subunits EncA (KS) and EncB (CLF) of the EncAB were dissociated under these denaturing conditions. Bands that correspond to crosslinked EncA–EncC were then excised and digested with either trypsin, chymotrypsin, or pepsin. MALDI MS/MS analyses resulted in the identification of peptides that correspond to 71% sequence coverage of EncC and 96% of EncA. Importantly, the active-site residues, Ser72 of our EncC construct and Cys191 of our EncA construct, were not identified in the analysis. This is consistent with post-translational modification of Ser72 and subsequent reaction of *crypto*-EncC with Cys191, and results in the site-directed crosslink between EncAB and EncC. While it would have been desirable to obtain peptide fragments representative of the active-site regions in the absence of crosslinking conditions, this would have required considerably more experimentation. Instead we chose to address this issue by using protein crystallographic studies, which are currently underway. In addition, all of these active-site-specific crosslinking reactions failed when the KS was subjected to heat denaturation prior to reaction (unpublished results). Thus, from these experiments and those of our initial report, the site-specific nature of the crosslinking has been verified for all ACP and KS species used to date.

#### Investigation of intersystem ACP–KS partner protein interactions by crosslinking analyses

With each of the compounds tested, KASII from *E. coli* FAS crosslinks its cognate carrier protein AcP to a greater degree than KASI. As stated above, this suggests that the AcP–KASII interaction is stronger than that of AcP–KASI, a finding that is supported by previous kinetic analyses of these enzymes.<sup>[30]</sup> What we intended to examine next was whether the strength of these interactions were mainly functions of the ACP or KS component, and to what extent these protein–protein interactions were compatible across different biosynthetic systems. To study this, we performed combinatorial crosslinking experiments by using two compounds: chloroacrylamide **1**, our model “starter unit” electrophilic 4'-phosphopantetheine analogue, and  $\alpha,\beta$ -unsaturated acetylene amide **3**, a long-chain 4'-phosphopantetheine precursor, which had shown good crosslinking of EncC–EncAB enzymes. By using our one-pot chemoenzymatic method, each analogue was loaded onto the carrier proteins AcP or EncC before incubation with one of the three ketosynthase domains, KASI, KASII, or EncAB (Figure 2D).

Analysis of the crosslinking reactions by gel-shift assay showed that, as expected, each KS domain showed a preference for interacting with its natural carrier protein (Table 1B). In addition, KASII showed negligible crosslinking to *crypto*-EncC regardless of whether it was modified with analogue **1** or **3**. The correlation between the crosslinking efficiency and compatibility of ACP–KS binding surfaces suggests that the binding interactions of KASII are mainly dependent on a tight surface residue interaction rather than interactions that result from the loading of substrate on the 4'-phosphopantetheine arm. In contrast KASI showed a high degree of crosslinking (89%) with EncC that was modified with starter unit analogue **1**, but negligible reactivity (1%) with *crypto*-EncC loaded with **3**; this indicates the presence of a less permissive active site that is coupled to more permissive, less-specific binding residues at the protein surface. Together these results suggest that KASI and KASII utilize subtly different yet complementary binding modes in *E. coli* type II FAS (Figure 4).

The type II PKS ketosynthase EncAB showed less preference for the carrier protein with which it interacted. *Crypto*-EncC that was primed with compounds **1** and **3** crosslinked EncAB reasonably well at values of 32% and 49%, respectively. When the unnatural carrier protein AcP was similarly prepared with **1** and **3** these values were 34% and 28%, respectively. Notably, EncAB showed a significant preference for its natural carrier protein partner when using  $\alpha,\beta$ -unsaturated alkyne **3**, a compound that is more similar in size to the natural substrate of EncAB than compound **1**. This experiment was comparable to the study by Izumikawa et al., which demonstrated that the benzoate:CoA ligase EncN, which is responsible for loading benzoate onto EncC can efficiently load ACPs from type II PKS systems other than enterocin.<sup>[28]</sup> Together with the observed crosslinking results, this apparent promiscuity of the carrier protein that is used in crucial steps of enterocin biosynthesis might be an indication that the enterocin system is capable of utilizing the carrier protein from its type II FAS system for



**Figure 4.** Summary of results from the experiment designed to compare substrate specificity and binding compatibility between the KSSs of type II biosynthesis and both natural and unnatural ACP partners. Wavy lines that are attached to the ACP cartoons correspond to phosphopantetheine analogue 1 (shorter line) or 3 (longer line). Significant binding interactions were detected as a gel-shift on SDS-PAGE in the reactions marked with a “√”, while negligible interactions occurred in the reactions that are marked with a “X”.

enterocin biosynthesis during times of stress. Alternatively, whereas EncC is located in the enterocin biosynthetic gene cluster, we cannot discount the possibility that this organism expresses only a single ACP that is used for both type II FAS and PKS biosynthesis. To date, only 21 proteins from this organism have been sequenced, all but one of which are associated with the enterocin biosynthetic pathway. Based upon the observed results, one could conjecture that when the *S. maritimus* FAS gene cluster is eventually deciphered, the sequence of its associated carrier protein might very likely be strikingly similar to EncC.

## Conclusions

In this study we have developed the ACP-KS crosslinking methods from our preliminary reports, synthesizing a variety of 4'-phosphopantetheine analogue precursors in order to identify the key structural characteristics necessary for ACP-KS crosslinking. In addition we have tested the compatibility of three new electrophilic functionalities with this approach and expanded the structure of our crosslinking reagents to incorporate fatty alkyl substrate mimics. The utilization of this panel allowed the first crosslinking of an ACP-KS pair from a type II PKS system. Such an approach might allow structural analysis of this single step in the biosynthesis of the medicinally relevant enterocin and aid future efforts at rational design of combinatorial biosynthetic enzymes. We demonstrated the site-specificity of our crosslinking agents in the natural KASI-ACP, KASII-ACP, and EncAB-EncC systems by using MS analyses. The site-specificity of crosslinking between unnatural partner proteins was not evaluated in this study. Future studies exploring the functional aspects of these unnatural interactions will address this issue.

Perhaps most interestingly, this study has demonstrated a clear correlation between the structure of the electrophilic 4'-phosphopantetheine arm and the active-site structure, substrate specificity, and crosslinking efficiency. In particular, our results agree with a variety of studies that have investigated binding affinity and substrate specificity of KS domains in type II FAS and PKS systems. This finding points to the information-rich nature of thoroughly studied crosslinking analyses by using our system. Finally, we have used our techniques to investigate the dominant interactions that are responsible for ACP-KS binding in both fatty acid and enterocin biosynthetic systems. Although the results of these intersystem crosslinking studies must be verified by analysis of substrate turnover, we hope to use this approach in the future to aid in the development of unnatural biosynthetic systems by investigating a wide range of ACP-KS pairs to help establish a foundational knowledge of ACP-KS interactions. This approach could be further aided by mutagenic studies in which we use crosslinking studies to determine the amino acid residues that are crucial for protein surface and enzyme–substrate interactions in these systems.

## Experimental Section

Synthetic protocols, compound characterization, and additional experimental procedures can be found in the Supporting Information.

**Crosslinking of recombinant ACP and ketosynthase:** To compare the reactivities of ketosynthases towards a panel of pantetheine analogues (1–9) and to compare the interaction between ketosynthase and natural vs. unnatural carrier protein by SDS-PAGE, the following procedure was used. Carrier protein (4 µg,  $\geq$  fivefold excess), CoaA (0.1 µg), CoaD (0.5 µg), CoaE (1 µg), and *B. subtilis* Sfp (0.5 µg) were added to a 50 mM potassium phosphate, pH 7.0, buffer with 12.5 mM Mg<sub>2</sub>Cl<sub>2</sub> and 8 mM ATP. The pantetheine analogue (0.1 mM) was added, and the mixture was incubated at 37 °C for 5 min, then ketosynthase (2–3.5 µg) was added, and the reaction was allowed to proceed at 37 °C for 1 h. Samples were run on a 9% SDS-PAGE, where cross-linked product was detected by staining with Coomassie. Gels were photographed with a BioRad Fluor-S Multilanner (Hercules, CA, USA), and bands were quantified by using the ImageJ software that was provided by the NIH (Bethesda, MD, USA).<sup>[34]</sup>

**Mass spectrometry analyses:** To verify the post-translational modification of carrier protein by using matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS), the following procedure was used. Carrier protein (4 µg), CoaA (0.1 µg), CoaD (0.5 µg), CoaE (1 µg), and *B. subtilis* Sfp (0.5 µg) were added to a 50 mM potassium phosphate, pH 7.0, buffer with 12.5 mM Mg<sub>2</sub>Cl<sub>2</sub> and 8 mM ATP. Pantetheine analogue 1 or 3, (0.1 mM), was added, and the mixture was incubated at 37 °C for 1 h. The reaction was analyzed by MALDI MS without further purification.

MALDI MS/MS analyses of in-gel-digested complexes were performed as previously described.<sup>[19]</sup>

## Acknowledgements

Funding was provided by ACS-PRF 42158-G4, NIH R01M075797, NIH R01AI47818, and ACS RSG-01-011-01-CDD. We thank H. Mori from the Japanese *E. coli* consortium for providing genes from the ASKA plasmid ORF library.<sup>[39]</sup>

**Keywords:** acyl carrier protein • acylation • biosynthesis • fatty acids • ketosynthase

- [1] D. A. Hopwood, D. H. Sherman, *Annu. Rev. Genet.* **1990**, *24*, 37–66.
- [2] S. Smith, S. C. Tsai, *Nat. Prod. Rep.* **2007**, *24*, 1041–1072.
- [3] S. W. White, J. Zheng, Y. M. Zhang, C. O. Rock, *Annu. Rev. Biochem.* **2005**, *74*, 791–831.
- [4] C. Hertweck, A. Luzhetskyy, Y. Rebets, A. Bechthold, *Nat. Prod. Rep.* **2007**, *24*, 162–190.
- [5] A. C. Mercer, M. D. Burkart, *Nat. Prod. Rep.* **2007**, *24*, 750–773.
- [6] T. Maier, S. Jenni, N. Ban, *Science* **2006**, *311*, 1258–1267.
- [7] C. Khosla, Y. Tang, A. Y. Chen, N. A. Schnarr, D. E. Cane, *Annu. Rev. Biochem.* **2007**, *76*, 195–221.
- [8] B. Shen, *Curr. Opin. Chem. Biol.* **2003**, *7*, 285–295.
- [9] R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chem. Biol.* **1996**, *3*, 923–936.
- [10] A. S. Worthington, M. D. Burkart, *Org. Biomol. Chem.* **2006**, *4*, 44–46.
- [11] J. J. La Clair, T. L. Foley, T. R. Schegg, C. M. Regan, M. D. Burkart, *Chem. Biol.* **2004**, *11*, 195–201.
- [12] A. C. Mercer, J. J. La Clair, M. D. Burkart, *ChemBioChem* **2005**, *6*, 1335–1337.
- [13] J. Yin, A. J. Lin, D. E. Golan, C. T. Walsh, *Nat. Protoc.* **2006**, *1*, 280–285.
- [14] S. A. Sieber, C. T. Walsh, M. A. Marahiel, *J. Am. Chem. Soc.* **2003**, *125*, 10862–10866.
- [15] L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry* **1998**, *37*, 1585–1595.
- [16] J. L. Meier, A. C. Mercer, H. Rivera Jr., M. D. Burkart, *J. Am. Chem. Soc.* **2006**, *128*, 12174–12184.
- [17] J. L. Meier, A. C. Mercer, M. D. Burkart, *J. Am. Chem. Soc.* **2008**, *130*, 5443–5445.
- [18] P. C. Dorrestein, J. Blackhall, P. D. Straight, M. A. Fischbach, S. Garneau-Tsodikova, D. J. Edwards, S. McLaughlin, M. Lin, W. H. Gerwick, R. Kolter, C. T. Walsh, N. L. Kelleher, *Biochemistry* **2006**, *45*, 1537–1546.
- [19] A. S. Worthington, H. Rivera, J. W. Torpey, M. D. Alexander, M. D. Burkart, *ACS Chem. Biol.* **2006**, *1*, 687–691.
- [20] G. D'Agnolo, I. S. Rosenfeld, P. R. Vagelos, *J. Biol. Chem.* **1975**, *250*, 5283–5288.
- [21] J. G. Olsen, A. Kadziola, P. von Wettstein-Knowles, M. Siggaard-Andersen, S. Larsen, *Structure* **2001**, *9*, 233–243.
- [22] J. Dreier, C. Khosla, *Biochemistry* **2000**, *39*, 2088–2095.
- [23] R. W. Haushalter, A. S. Worthington, G. H. Hur, M. D. Burkart, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3039–3042.
- [24] N. Wu, D. E. Cane, C. Khosla, *Biochemistry* **2002**, *41*, 5056–5066.
- [25] H. G. Menzella, R. Reid, J. R. Carney, S. S. Chandran, S. J. Reisinger, K. G. Patel, D. A. Hopwood, D. V. Santi, *Nat. Biotechnol.* **2005**, *23*, 1171–1176.
- [26] K. J. Weissman, P. F. Leadlay, *Nat. Rev. Microbiol.* **2005**, *3*, 925–936.
- [27] J. Piel, C. Hertweck, P. R. Shipley, D. M. Hunt, M. S. Newman, B. S. Moore, *Chem. Biol.* **2000**, *7*, 943–955.
- [28] M. Izumikawa, Q. Cheng, B. S. Moore, *J. Am. Chem. Soc.* **2006**, *128*, 1428–1429.
- [29] Q. Cheng, L. Xiang, M. Izumikawa, D. Meluzzi, B. S. Moore, *Nat. Chem. Biol.* **2007**, *3*, 557–558.
- [30] K. Arnvig McGuire, J. N. McGuire, P. von Wettstein-Knowles, *Biochem. Soc. Trans.* **2000**, *28*, 607–610.
- [31] J. G. Olsen, A. Kadziola, P. von Wettstein-Knowles, M. Siggaard-Andersen, Y. Lindquist, S. Larsen, *FEBS Lett.* **1999**, *460*, 46–52.
- [32] M. Moche, K. Dehesh, P. Edwards, Y. Lindqvist, *J. Mol. Biol.* **2001**, *305*, 491–503.
- [33] C. P. Govardhan, R. H. Abeles, *Arch. Biochem. Biophys.* **1996**, *330*, 110–114.
- [34] B. V. Plapp, W. S. Chen, *Methods Enzymol.* **1981**, *72*, 587–591.
- [35] M. D. Abramoff, P. J. Magelhaes, S. J. Ram, *Biophoton. Int.* **2004**, *11*, 36–42.
- [36] J. Xun, H. Huang, K. W. Vogel, D. G. Drueckhammer, *Bioorg. Chem.* **2005**, *33*, 90–107.
- [37] J. Dreier, A. N. Shah, C. Khosla, *J. Biol. Chem.* **1999**, *274*, 25108–25112.
- [38] P. Edwards, J. S. Nelsen, J. G. Metz, K. Dehesh, *FEBS Lett.* **1997**, *402*, 62–66.
- [39] M. Kitagawa, T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, H. Mori, *DNA Res.* **2005**, *12*, 291–299.

Received: March 29, 2008

Published online on July 30, 2008