Site-Specific Observation of Acyl Intermediate Processing in Thiotemplate Biosynthesis by Fourier Transform Mass Spectrometry: The Polyketide Module of Yersiniabactin Synthetase†

Matthew T. Mazur,‡ Christopher T. Walsh,§ and Neil L. Kelleher*,‡

Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Received September 3, 2003; Revised Manuscript Received October 11, 2003

ABSTRACT: Complex arrays of thioester bound intermediates are present on 100−700 kDa enzymes during the biogenesis of diverse types of pharmacophores and natural product drugs. These multidomain enzymes, known as nonribosomal peptide synthetases and polyketide synthases (NRPSs and PKSs, respectively), synthesize from simple, physiologically available substrates bioactive compounds that can be further tailored by a host of modifying domains (e.g., methylation, cyclization, and epimerization) to increase the complexity of the mature final product. Interrogation of such covalent intermediates using mass spectrometry (MS) presents an underutilized method for understanding the covalent catalysis executed by NRPS and PKS enzymes. For the PKS module (205 kDa) from the yersiniabactin (Ybt) gene cluster of *Yersinia pestis*, limited proteolysis afforded a key 11 kDa peptide from the acyl-carrier protein (ACP) domain upon which at least five covalent intermediates could be detected (42, 70, 86, 330, and 358 Da). The isotopic resolution achieved by Fourier transform mass spectrometry (FTMS) allowed for the incorporation of substrates with stable isotopes to confirm the structural assignments of three intermediates (86, 330, and 358 Da) on the Ybt biosynthetic pathway to within 1 Da. Approximately 75% of the enzyme capacity is lost to unproductive decarboxylation of malonyl-S-ACP partly constraining the 1.4 min^{-1} rate of Ybt production *in vitro*. Acyl transfer to the ACP domain (on the Ybt pathway) was promoted by a factor of ∼10 over unproductive CO_2 loss in the presence of the cosubstrate *S*-adenosylmethionine (SAM), with *S*-adenosylhomocysteine unable to restore the condensation yield observed with SAM. The data are consistent with Claisen condensation from KS to the ACP carrier site being reversible, with the absence of downstream methylation providing more opportunity for unproductive CO_2 loss. Extension of such FTMS-based studies will allow the direct visualization of multiple intermediates in determining the catalytic order of events and kinetics of NRPS and PKS systems.

Many medicinally interesting natural products, including clinically important immunosuppressive (e.g., rapamycin) and antitumor drugs (e.g., epothilones), as well as bacterial virulence factors are synthesized on large, multifunctional enzymes through a process of mixed nonribosomal peptide and polyketide synthesis. The general thiotemplate mechanism depicts each synthetase as being equipped with one or more modules, or catalytically active regions responsible for the processing and elongation of one substrate, and it is the linear order in which these modules are arranged that determines the composition of the mature natural product (*1*). To exploit these systems for the combinatorial synthesis of "unnatural" natural products, a better understanding of the mechanisms employed by such enzymes is imperative (*2*). Because the biosynthesis proceeds via covalent acyl-*S*-enzyme intermediates (thioesters) that result in a change in mass, the enzymatic logic involved in NRPS[1] and PKS can be assessed by large molecule mass spectrometry (MS), with Fourier transform instruments providing unique abilities to analyze complex mixtures at a resolving power of $>10^5$ (*3*, *4*).

Yersiniabactin (Ybt) is a bacterial siderophore from the plague-causing *Yersinia* bacterium biosynthesized via a hybrid NRPS/PKS system and is a primary attribute to virulence in mammalian hosts (*5*). In iron-starved environments, Ybt is synthesized by five essential enzymes and exported to strip iron from host proteins (e.g., cytochrome *c*) followed by uptake by the bacterial cell for survival (*6*). The Ybt synthesizing genes are all clustered within a high-pathenogicity island (HPI), and any interruption in Ybt synthesis or regulation in iron deficient environments is

† This work has been supported by grants to N.L.K. (GM 067725) and C.T.W. (AI 42738) from the National Institutes of Health.

* To whom correspondence should be addressed. Phone: (217) 244-3927. Fax: (217) 244-8068. E-mail: kelleher@scs.uiuc.edu.

‡ University of Illinois at Urbana-Champaign.

§ Harvard Medical School.

[1] Abbreviations: NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase; MS, mass spectrometry; FTMS, Fourier transform mass spectrometry; Ybt, yersiniabactin; HPI, high-pathenogicity island; HMWP, high-molecular weight protein; ArCP, aryl carrier protein; HPTT, 2′-(2-hydroxyphenyl)thiazolyl-2,4-thiazolinyl; TE, thioesterase domain; ESI, electrospray ionization; $P_i$pant, 4′-phosphopantetheine; ACP, acyl carrier protein; KS, ketoacylsynthase; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; RPLC, reverse-phase liquid chromatography; CoA, coenzyme A; SAM, *S*-adenosylmethionine; NADPH, nicotinamide adenine dinucleotide phosphate; M_r, relative molecular weight; MS/MS, tandem MS; ECD, electron capture dissociation; KR, ketoacylreductase.

10.1021/bi035585z CCC: $25.00
Published on Web 11/01/2003

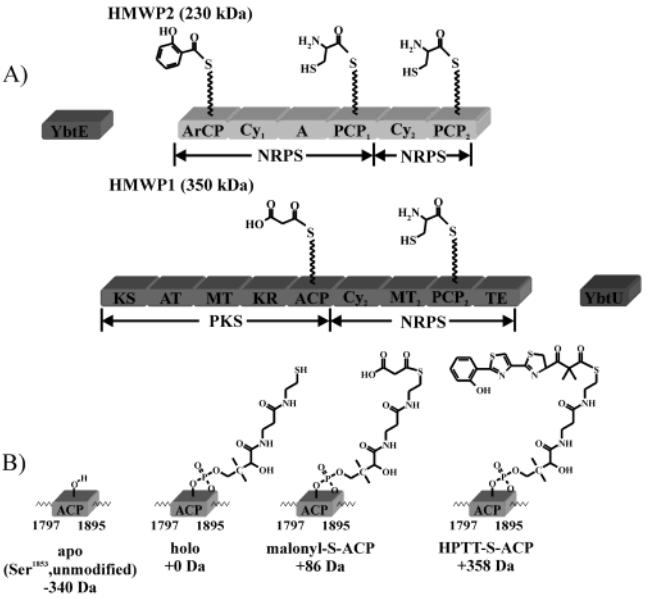


FIGURE 1: (A) View of the domain structure of HMWP2 (230 kDa) and HMWP1 (350 kDa), hybrid NRPS and PKS enzymes responsible for the biogenesis of yersiniabactin. (B) Four different covalent acylthioester states of the ACP domain of HMWP1. The molecular mass differences, in reference to the holo form (HS-enzyme), are listed at the bottom.

detrimental to the propagation of the organism (*7*). The four main proteins involved in Ybt synthesis are YbtE (adenylation domain, 60 kDa), High-Molecular Weight Protein 2 (HMWP2, NRPS−NRPS module, 230 kDa), High-Molecular Weight Protein 1 (HMWP1, PKS−NRPS module, 350 kDa), and YbtU (reductase, 41 kDa) which together consist of 17 catalytically specific domains (Figure 1) (*8*). Biosynthesis is initiated by ATP-directed adenylation and loading of salicylate onto the N-terminal ArCP (aryl carrier protein) domain of HMWP2 by YbtE (*9*), and the next five domains consist of two NRPS modules that utilize and cyclize two cysteines to form 2′-(2-hydroxyphenyl)thiazolyl-2,4-thiazolinyl-enzyme (HPTT-enzyme) (*10*). At an intermolecular NRPS to PKS switch point, a five-domain PKS module loads, bis-methylates, and reduces a malonyl linker group. Finally, an intramodular handoff back to a cysteine-utilizing NRPS module precedes a YbtU-driven reduction and thioesterase (TE)-catalyzed hydrolysis to afford the natural product (seen at the right of Scheme 1).

While this system has been reconstituted *in vitro* (*11*), much of the detailed chemistry and kinetics of the Ybt system remains opaque. The ability to unambiguously identify multiple intermediates present at substoichiometric levels attached directly to the enzymes using mass spectrometry (MS) will allow for a detailed understanding of the mechanism by which such enzymes function. As the natural product is constructed from covalently attached enzyme intermediates, each active site can exist as one of several forms that differ in mass (Figure 1B). Experiments typically require either radioisotope-labeled substrates or base hydrolysis of intermediates for detection, and it is by this iterative, indirect approach that insights into the mechanism have been previously inferred (*9*, *11*−*13*). Via exploitation of its ability to analyze complex samples with high resolution, electrospray ionization (ESI) (*14*) with Fourier transform mass spectrometry (ESI/FTMS) is well suited to interrogation of the enzymatic processing of the polyketide module of Ybt biosynthesis. The mechanism of intermolecular HPTT chain transfer from HMWP2 to HMWP1 is proposed to be initiated by malonyl priming of the 4′-phosphopantetheine ($P_i$pant) tether at the acyl carrier protein (ACP) site, followed by decarboxylation, condensation (facilitated by the ketoacyl-synthase, KS, domain), and bis-C-methylation. The precise order and timing of these events remains to be elucidated because ultrasensitive radioassays provide an incomplete picture of intermediate occupancies at specific carrier sites. This contrasts with this new MS-based approach, which allows a more general readout of covalent intermediates loaded onto thiotemplate carrier sites. This largely untapped detection ability (*15*) is illustrated here with the first detailed interrogation of the catalytic order and substrate dependencies of the enzymatic tailoring occurring at the first NRPS to PKS switch point in yersiniabactin synthetase.

EXPERIMENTAL PROCEDURES

General. PKS protein, residues 1−1895, a recombinant fragment of the enzyme HMWP1, was purified as previously described (*16*). The protein was purified to homogeneity as judged by SDS−PAGE and dialyzed into 50 mM Tris buffer (pH 8.0), 10 mM $MgCl_2$, 1 mM TCEP [tris(2-carboxyethyl)-phosphine hydrochloride], and 10% glycerol. PKS was determined to be post-translationally modified *in vivo* according to mass spectrometric data, and no further incubations were required for enzyme priming. Sfp (*17*), YbtE (*9*), and HMWP2 (*18*) were purified as described elsewhere.

Limited Trypsin Proteolysis. Limited proteolysis was performed by incubating TPCK-treated trypsin (Promega) and PKS (1:10 molar ratio) at 30 °C for 5 min with gentle agitation. All digests were then quenched by the addition of an equal volume of 10% formic acid (Fluka), and mixtures were fractionated by reverse-phase liquid chromatography (RPLC) using a wide-pore (300 Å) C4 column with a linear gradient from 30 to 65% ACN (0.1% TFA) for separation and desalting before FTMS analysis. Samples were lyophilized before resuspension in ESI buffer (49% H_2O, 49% MeOH, and 2% formic acid) for FTMS analysis.

Loading of Intermediates at the AT and ACP Domains. Malonyl loading experiments were carried out by incubating 2.5 μM holo-PKS with 1.5 mM malonyl-CoA (Sigma) for 20 min at 30 °C and stopped by spinning in a Microcon Y-30 (Amicon) according to the manufacturer's protocol. The concentrated enzyme was resuspended in 100 μL of 50 mM NH_4HCO_3 (pH 7.8) and digested with trypsin as described above. For the determination of the apparent nonenzymatic hydrolysis of malonyl from the acyltransferase (AT) and ACP covalent sites, 2.5 μM holo-PKS was incubated with 50 μM malonyl-CoA in 100 mM Tris (pH 8.0), 10 mM $MgCl_2$, and 1 mM TCEP for various times between 20 min and 24 h, the reaction stopped, and the holo-PKS digested as described above.

In experiments in which the HPTT−β-ketoacyl complex intermediate was being loaded, 2.5 μM HMWP2, 2.5 μM YbtE, 2.5 μM holo-PKS, 0.15 μM Sfp, 0.5 mM CoA, 1 mM L-cysteine, 0.375 mM *S*-adenosylmethionine (SAM), 0.375 mM nicotinamide adenine dinucleotide phosphate (NADPH), and 1 mM salicylate in 100 mM Tris (pH 8.0), 10 mM $MgCl_2$, and 1 mM TCEP were incubated for 3 h to ensure

$P_i$pant priming of all enzymes before addition of 1 mM malonyl-CoA and 5 mM ATP to initiate the reaction. Samples were incubated at 30 °C for 20 min before reactions were halted and the samples digested as described above. Stable-isotope substrates were incorporated as described above using, separately, d_6-salicylate (Sigma), 3,3-d_2-cysteine (Cambridge Isotope Laboratories), or CD_3-*S*-adenosylmethionine (CDN Isotopes).

Assaying the Order of Decarboxylation, Methylation, and Reduction. Decarboxylation experiments were performed by incubating 2.5 μM holo-PKS with 1 mM malonyl-CoA, in 100 mM Tris (pH 8.0), 10 mM $MgCl_2$, and 1 mM TCEP, with 2.5 μM cysteine-loaded HMWP2, apo-HMWP2, or no HMWP2 and in the complete absence of YbtE and salicylate, for 20 min at 30 °C before concentration and digestion as described above. Methylation experiments were carried out by incubating 2.5 μM holo-PKS with 1.5 mM malonyl-CoA (or 6 mM acetyl-CoA), 1.5 mM SAM, and NADPH in 100 mM Tris (pH 8.0), 10 mM $MgCl_2$, and 1 mM TCEP in the presence or absence of holo-HMWP2 for 20 min at 30 °C. The reaction was allowed to proceed for 20 min before concentration and digestion as described above. As an extension of the above methylation reaction to assay the enzyme's ability to bis-methylate the +42 Da acetyl-*S*-ACP decarboxylation product, 2.5 μM holo-PKS was reconstituted with 2.5 μM HMWP2, 2.5 μM YbtE, 0.15 μM Sfp, 0.5 mM CoA, 1 mM L-cysteine, 0.375 mM SAM, 0.375 mM NADPH, and 1 mM salicylate in 100 mM Tris (pH 8.0), 10 mM $MgCl_2$, and 1 mM TCEP. After 3 h, 6 mM acetyl-CoA and 5 mM ATP were added to initiate the reaction and were processed as described above.

Determining the Dependence of Substrate S-Adenosylmethionine on Complex Intermediate Formation. Experiments were performed by incubating 2.5 μM HMWP2, 2.5 μM YbtE, 2.5 μM holo-PKS, 0.15 μM Sfp, 0.5 mM CoA, 1 mM L-cysteine, 0.375 mM NADPH, and 1 mM salicylate in 100 mM Tris (pH 8.0), 10 mM $MgCl_2$, and 1 mM TCEP in the complete absence of SAM for 3 h to ensure $P_i$pant priming of all enzymes before addition of 1 mM mal-CoA and 5 mM ATP to initiate the reaction. Reaction mixtures were incubated for an additional 20 min before the reactions were halted and the mixtures digested as described above. To further probe the dependence of SAM on the condensation yield, a separate reaction mixture but including 0.375 mM *S*-adenosylhomocysteine was incubated and processed as described above.

FTMS Analysis. The system used in this study was a custom 8.5 T quadrupole-FTMS hybrid described elsewhere (*19*). In general, ions from ESI were directed through a heated metal capillary, skimmer, and multiple ion guides into the ion cell ($\sim 10^{-9}$ Torr) of the FTMS system. Theoretical isotopic distributions were generated using Isopro version 3.0 and fit to experimental data by least squares to assign the most abundant isotopic peak (*20*). The mass difference (in units of 1.0024 Da) between the most abundant isotopic peak and the monoisotopic peak is denoted in italics after relative molecular weight (M_r) values (*21*). Spectra were calibrated externally using bovine ubiquitin (8564.63-*5* Da).

Determination of Intermediate Relative Ratios. Experiments for measuring the precision of the semiquantitative results reported for intermediate ratios were carried out by repeating the malonyl loading experiments (above) five times; ion abundances were summed over all charge states for the determination of relative ratios of intermediates. Errors reported from those experiments (in percent occupancy) were taken to be 2σ from the average determined by their relative ion abundances.

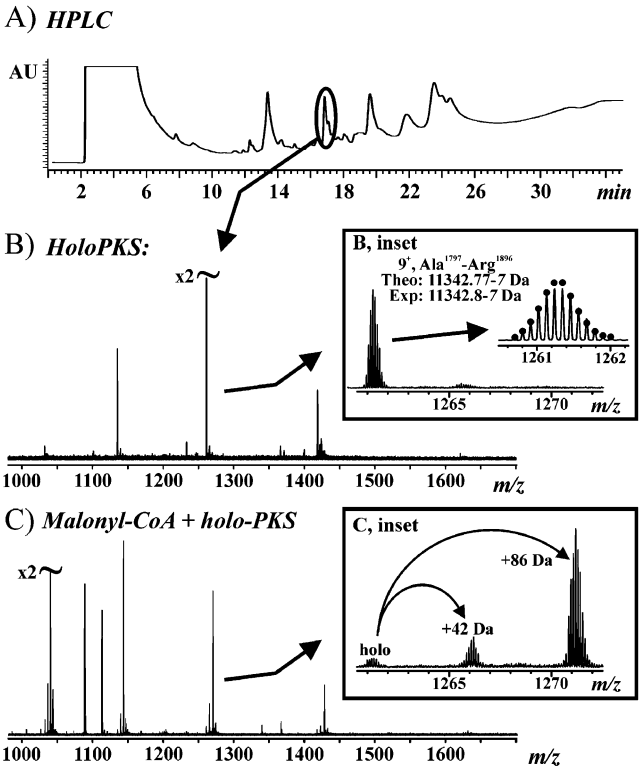


FIGURE 2: (A) Reverse-phase HPLC separation of the PKS module of HMWP1 following a limited proteolysis using trypsin. Acyl carrier protein ($Ala^{1797}-Arg^{1896}$) elutes at 16−17 min. (B) Electrospray ionization Fourier transform mass spectrum (ESI/FTMS) of tryptic ACP peptide from panel A. The inset shows an expanded view of the mass-to-charge (*m*/*z*) region containing the 9^+-ion isotopic distribution. (C) ESI/FTMS of the ACP peptide from panel A following treatment of PKS with malonyl-CoA. The inset shows an expanded view of the *m*/*z* region illustrating the mass shift observed from the covalent attachment of the malonyl moiety (+86 Da). The peak at +42 Da results from the decarboxylation of malonyl-*S*-ACP (see Results and Discussion).

RESULTS

FTMS Analysis of the ACP Domain of Holo-PKS. To determine the covalent state of the ACP domain of the 205 kDa PKS module (residues 1−1895) after overexpression in *Escherichia coli*, a limited tryptic digest was fractionated by RPLC (Figure 2A). ESI/FTMS analysis of the 16−17 min fraction afforded an abundant species with an M_r of 11342.8-*7* Da (Figure 2B), consistent to within 0.06 Da of that of the C-terminal ACP fragment ($Ala^{1797}-Arg^{1896}$) excised at a domain junction and harboring the $P_i$pant cofactor on Ser^{1853}. This 11.3 kDa species was subjected to tandem MS (MS/MS) using electron capture dissociation (ECD) (*22*), generating 11 and 26 *c* and $z^{\cdot}$ ions, respectively, and confirms the putative assignment. The apo form of the enzyme (340.3 Da lower) was observed in a later fraction (18−19 min) with a far lower relative abundance (data not shown). Repeated experiments where RPLC fractions containing apo- and holo-ACP were mixed gave an apparent ratio of 2:98. Also, it was found that a species with an M_r of

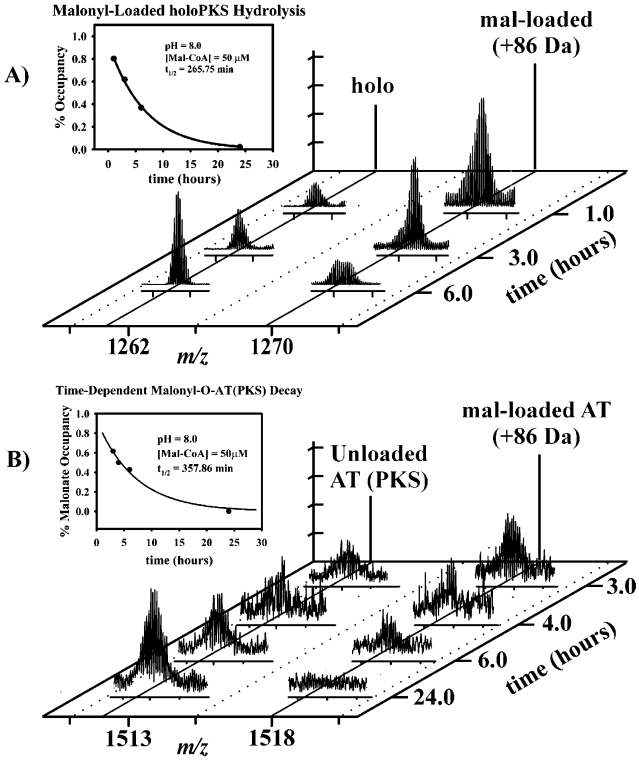


FIGURE 3: Hydrolysis kinetics of (A) malonyl-*S*-ACP and (B) malonyl-*O*-AT domains of the PKS module of yersiniabactin synthetase determined by ESI/FTMS.

22682.5-*14* Da matched within 0.5 Da that of the AT active site-containing peptide of residues $Cys^{582}-Arg^{790}$.

Having a robust protocol for the excision of the ACP domain allowed testing of the stability of thioester-bound intermediates. Treatment of holo-PKS with malonyl-CoA for 20 min followed by limited proteolysis and FTMS analysis gave the spectrum in Figure 2C with an 11428.8-*7* Da species, consistent within 0.1 Da with the 86.0 Da increase attributed to the malonyl-*S*-ACP intermediate (confirmed by ECD, data not shown). The apparent ratio of holo-(HS-ACP)to malonyl-*S*-ACP was 1:9, giving a value of 90 ± 3% for malonyl loading. Further, in the same 16−17 min fraction was detected a 22.7 kDa species (from the AT domain) producing a mass shift of +86 Da, consistent with oxoester formation at Ser^{641} in the AT active site (malonyl-*O*-AT) prior to priming of the ACP site. To explore the rates of thioester (and oxoester) loss during incubation at pH 8.0 that result not from enzymatic processing but from experimental design, time points of malonyl attachment at the ACP and AT domains were obtained (Figure 3). FTMS analysis yielded an apparent hydrolysis rate of 0.0014 min^{-1} for the mal-*O*-AT oxoester (Figure 3B), and a higher rate of 0.0019 min^{-1} for the mal-*S*-ACP thioester observed at the ACP domain (Figure 3A).

With the ability to achieve and retain ~90% occupancy of simple acyl-thioester intermediates on the PKS module (Figure 2C), experiments to produce the authentic growing acyl intermediate accumulating on the ACP domain were undertaken. After reconstitution of HMWP2 with the PKS module of HMWP1 (Figure 1), digestion, and fractionation, FTMS analysis of the 18−19 min fraction generated six new species (visible in Figure 5) at 11384.8-*7* Da (+42 Da), 11412.8-*7* Da (+70 Da), 11540.0-*7* Da (+198 Da), 11567.9-*7* Da (+225 Da), 11700.8-*7* Da (+358 Da), and 11720.3-*7* Da (+378 Da). Each of these species has been verified as the ACP carrier peptide by MS/MS (spectra not shown), and their proposed structures are given in Table 1. The +42 Da species is attributed to the (uncoupled) decarboxylation of malonyl-*S*-PKS to acetyl-*S*-PKS (**3**, Figure 4C). The observed mass shift of +358 Da (±2 Da; see below) is consistent with the buildup of the natural 3-(hydroxyphenylthiazolylthiazolinyl)-*β*-keto-2,2-dimethyl acyl-enzyme (HPTT-*β*-ketoacyl-enzyme, **5a**, Figure 5A) with the first thiazoline ring air-oxidized to the thiazole and the *β*-keto group remaining unreduced. In some cases, a broadened peak (i.e., isotopic distribution) centered at +360 Da was observed and is attributed to the mixture of oxidation states (±2 Da), occurring at either the *β*-keto group or first thiazole ring. The second thiazoline ring remained unreduced because of the absence of YbtU. Clear evidence for bis-methylation on the acyl linker group of the complex intermediate was observed in the presence of SAM, and the results were confirmed by introduction of stable-isotope substrates. For example, introduction of CD_3-SAM yielded a 6 Da shift consistent with the formation and trapping of HPTT-*β*-ketoacyl-enzyme (panel A of Figure 5 vs panel B). Using

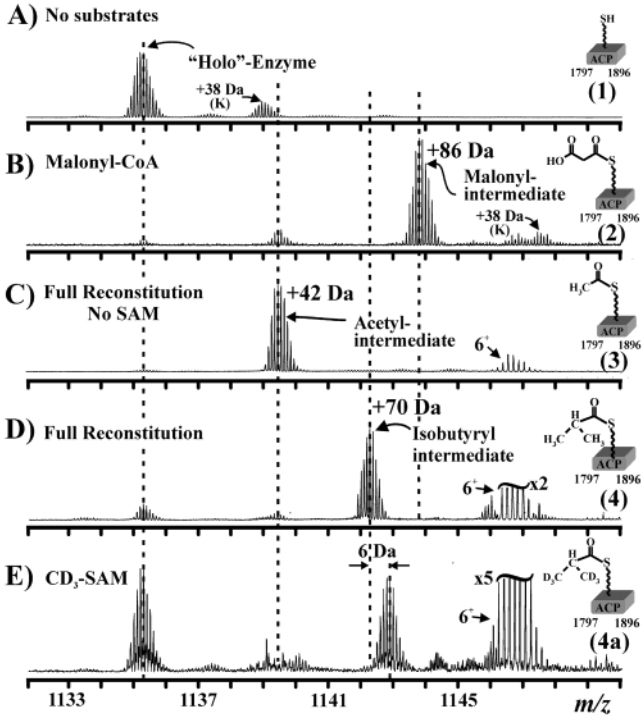


FIGURE 4: Isotopic distributions of thioester intermediates present on the ACP carrier site from the indicated conditions. The scale is set to *m/z*, and is expanded to illustrate only "simple" thioester intermediates. All ACP peptide ions are 10^+; the 6^+ ion at *m/z* ~1147 is an unrelated peptide. The phosphopantetheine cofactor is drawn as a wavy line. (A) Exclusively holo form, 340 Da above the naked protein sequence (with potassium adduct, +38 Da). (B) Malonyl-*S*-ACP (+86 Da) produced after treatment with malonyl-CoA. (C) Acetyl-*S*-ACP produced by stimulated decarboxylation of malonyl-*S*-ACP in the presence of the upstream complex intermediate without SAM. (D) Conditions of full reconstitution, including enzymes Sfp, YbtE, and HMWP2, the PKS module of HMWP1, and all requisite substrates (CoA, salicylate, cysteine, malonyl-CoA, SAM, NADPH, and ATP). (E) Conditions identical to those for panel D but with the cosubstrate SAM replaced with isotopically labeled CD_3-SAM. For details, see Experimental Procedures.

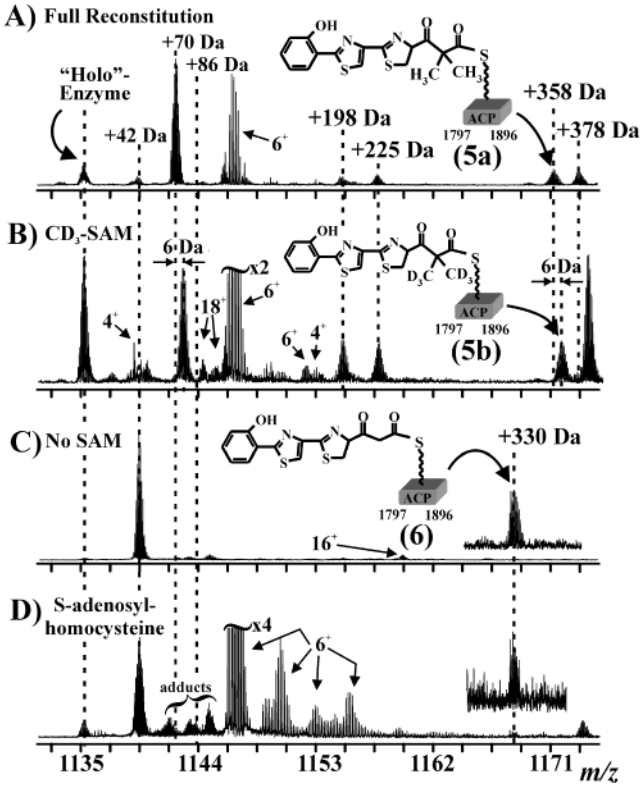


FIGURE 5: High-resolution mass spectra depicting in detail all complex acyl intermediates loaded onto PKS1 at the ACP peptide under the indicated conditions. The scale is set to *m*/*z* and is expanded to illustrate all enzyme intermediate mass shifts. All ACP peptide ions are 10^+, and unrelated peptide ions are labeled with their respective charge states. (A) Fully reconstituted conditions. (B) Same as panel A, but using CD_3-labeled SAM. (C) Same as panel A, but in the absence of SAM. (D) Same as panel C, but with SAM replaced with *S*-adenosylhomocysteine. For details, see Experimental Procedures.

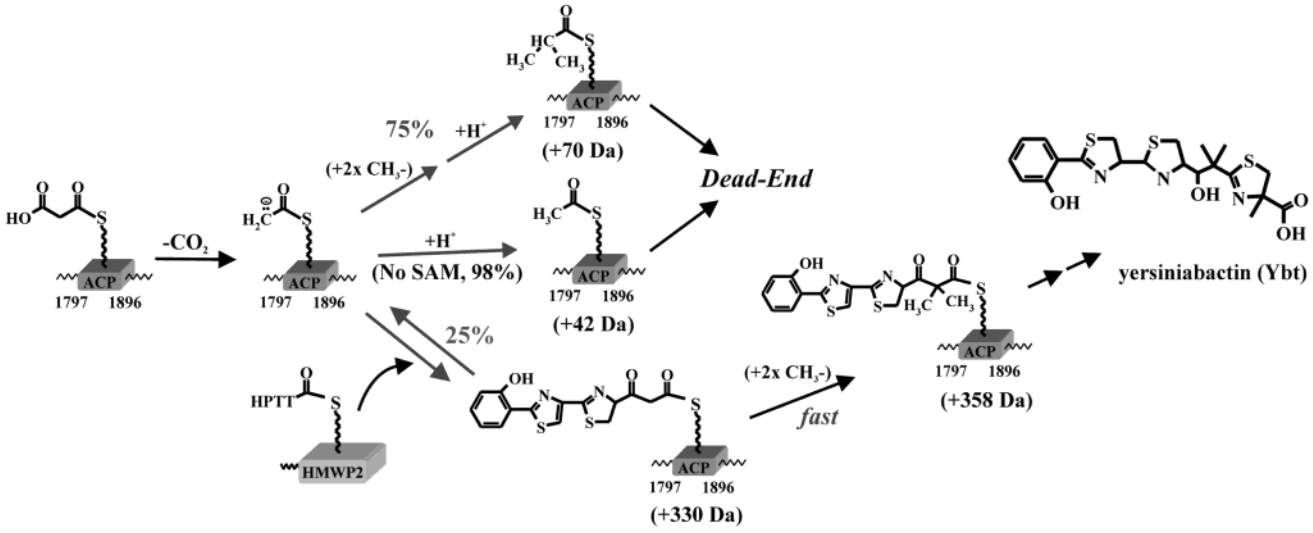
Table 1: Chemical Structures of Observed Mass Shifts

Mass Shift (Da)	Structure[a]
+0	SH
+42	
+70	
+86	
+198	?
+225	?
+358	
+378	

[a] Assignments confirmed by stable isotope labeling; phosphopantetheine arm illustrated as a wavy line (left).

d_6-salicylate and d_2-cysteine separately yielded mass shifts of 5 and 3 Da, respectively, consistent with the structure in Figure 5A. The +198 and +225 Da species produce no detectable mass shift upon inclusion of stable-isotope substrates, and because these species appear to contain no cysteine, salicylate, or methyl groups (produced from SAM), the identities of these products (reproducibly observed) remain to be elucidated. The largest addition observed at +378 Da undergoes mass shifts analogous to the +358 Da species (i.e., +5, +3, and +6 Da) upon addition of d_6-salicylate, d_2-cysteine, and CD_3-SAM, respectively. The identity of the +70 Da species, predicted to be the bis-

Scheme 1: Partitioning of Malonyl-*S*-ACP during Induced Decarboxylation between on Pathway C−C Bond Formation and off Pathway Protonation to Acetyl-*S*-ACP[a]



[a] Following intramodular loading of the malonyl group by the action of the AT domain onto the ACP domain, the malonyl moiety is decarboxylated when the HPTT donor acyl group is presented on HMWP2. Either unproductive decarboxylation (and protonation, +42 Da) or coupled decarboxylation and condensation (+330 Da) can occur. The Claisen condensation β-ketoacyl-*S*-ACP is trapped in a rapid bis-methylation to the +358 Da acyl-*S*-ACP. In the absence of the methyl donor SAM, the branch point to unproductive decarboxylation is highly favored, with proposed reversal of the C−C bond forming step and irreversible protonation to the acetyl-*S*-ACP.

methylated form of the uncoupled decarboxylation product [42 Da + (2 × 14 Da) = 70 Da], was confirmed by the mass shift of 6 Da observed upon inclusion of CD_3-SAM.

Assaying the Stimulated Decarboxylation of Malonyl-S-ACP by HMWP2. To more fully understand the regulation and necessary stimulus of what triggers the decarboxylation event at the ACP site, malonyl-*S*-PKS was allowed to incubate with holo-HMWP2, with cysteine-loaded HMWP2, and without HMWP2. When malonyl-*S*-ACP was incubated with holo-HMWP2 and without HMWP2, analysis showed only 10% was decarboxylated to acetyl-*S*-ACP, with 90% remaining as malonyl-*S*-ACP and 1% as holo-PKS (Figure 4B). Incubation of malonyl-*S*-PKS with cysteine-loaded HMWP2 produced a significant enhancement of decarboxylation of the malonyl moiety to 60% occupancy by the +42 Da acetyl-*S*-ACP (with 30% as +86 Da malonyl-*S*-ACP and 10% as the holo form).

Timing of Methylation at the ACP Domain. It has been shown that bis-methylation occurs at C_α, but the timing of methylation relative to decarboxylation and condensation and the substrate dependencies of this tailoring process remain unproven. To elucidate the regulation of methyl transfer, malonyl-*S*-PKS was incubated with SAM and NADPH both in the presence and in the absence of HMWP2. In thecomplete absence of HMWP2, there was no indication of either mono- or bis-methylation (<1%) occurring (data not shown) on the malonyl- or acetyl-*S*-ACP acyl groups. Further, the addition of holo-HMWP2 to the reaction mixture produced no detectable amounts of mono- or bis-methylation of the malonyl-*S*-ACP intermediate. To ascertain whether the bis-methylation event can occur following unproductive decarboxylation, acetyl-*S*-PKS (formed directly from acetyl-CoA via the AT domain) was incubated with SAM and NADPH and produced no observable methylation events (mono, 56 Da; bis, 70 Da); further, the presence of holo-HMWP2 yielded no detectable amounts of the +70 Da mass shift. Only under reconstitution-type conditions (with requisite enzymes holo-PKS and holo-HMWP2, and substrates), but with substitution of acetyl-CoA for malonyl-CoA, was C-methylation detected: FTMS analysis detected a peak at +70 Da at a relative occupancy of ~4% to the total enzyme covalent status (data not shown), consistent with isobutyryl-*S*-ACP (+70 Da).

Dependence of Condensation on S-Adenosylmethionine. In the absence of SAM but with all substrates and holo-HMWP2 present, no bis-methylated intermediate (+70 Da) was observed, with ~98% of detectable ACP site occupancy harboring the +42 Da decarboxylation product of malonate (Figures 4C and 5C). A trace amount of a +330 Da component was observed (Figure 5C, inset), consistent with condensation of HPTT and the ACP carbanion. Why the absence of SAM nearly eliminated all formation of the HPTT−β-ketoacyl-*S*-enzyme intermediate remained unclear. To determine if a SAM-induced conformational change improves the condensation yield, a surrogate substrate incompetent for methyl transfer, *S*-adenosylhomocysteine, produced no observable enhancement of the +330 Da species (Figure 5D).

DISCUSSION

ESI-MS is currently used to determine relative ratios of particular species based on the relative abundance of ions of similar structure (e.g., only isotope differences in the new methods for quantitative proteomics). Here, ESI-MS was used to semiquantitatively identify species with subtle mass changes (Figure 1B) on a 12 kDa "carrier peptide". This approach, however, relies on the assumption that all such species (95% identical atoms) ionize with very similar efficiencies. Studies that aim to actually quantitate the differences in the ionization efficiencies of different forms of >10 kDa carrier peptides are underway.

To obtain site-specific information contained within the multiple carrier sites of thiotemplate enzymes, limited proteolysis with trypsin was employed for the generation of <40 kDa peptides that can be managed by ESI/FTMS. First, to address the ability to semiquantitatively measure the relative ratios of intermediates loaded at the ACP site, malonyl loading experiments were repeated several times ($n = 5$) to determine the precision in percent occupancies. By summing the ion abundances over all charge states, we determined the average relative ratio of the malonyl-loaded form to the holo form to be 90 ± 3% (at 2σ). As noted above, this ratio depends on their concentrations in solution *and* their differences in ionization efficiency during ESI; it does not report the absolute amounts of species present. Further, the data of Figure 3 that follow the time-dependent loss of covalent intermediates at pH 8 show no ratio abnormalities at any time points. With this, ratios reported here are precise to within 10% and presumed to be semiquantitative. Systematic errors are still possible (accuracy could be off), but would not obscure insight into enzymatic logic obtained in this study.

The starting point for our approach is the direct observation of covalently loaded intermediates at the active sites of the enzyme under study. Incubation of holo-PKS with malonyl-CoA produced the expected mass shift of +86 Da to both the limited proteolysis products that contained the AT (22.7 kDa) and ACP (11.3 kDa) active sites. Because oxo- and thioesters are known to be labile under alkaline conditions (*23*, *24*), data from a series of time points from 20 min to 24 h were collected to assess, simultaneously, the oxoester product at the AT domain and the thioester product at the ACP domain. As can be seen from Figure 3, at short incubation times both acyl-*O*-AT and acyl-*S*-ACP intermediates are retained at a very high percentage (>90%); therefore, subsequent time analyses were carried out at 20 min to balance accumulation of intermediates versus hydrolysis. A time course for evaluating buildup and decay of tandem acyl-*S*-ACP intermediates will be the subject of future study.

Under conditions of full reconstitution of the HMWP2/PKS module of the HMWP1 intersubunit interface, use of a C-terminally truncated form of HMWP1 (PKS module shown in Figure 1) should lead to accumulation of acyl intermediates at the ACP site through prevention of their translocation to the downstream peptidyl carrier site present in intact HMWP1. Under such conditions with PKS, digestion and analysis by FTMS generated six new acyl-ACP species. Incubation with stable-isotope substrates d_6-salicylate, 3,3-d_2-cysteine, and CD_3-SAM produced the predicted mass shifts (+5, +3, and +6 Da, respectively), and the results support the intermediate structures drawn in Table 1. The oxidation states were assigned on the basis of slow air oxidation of the first thiazoline ring to a thiazole, consistent with previous findings on HPTT (*11*), and the β-keto group

remaining unreduced by the KR domain. Although the mass accuracy is limited to ±1 Da (*21*), the data best represent a mass shift of 3 Da upon incubation with 3,3-d_2-cysteine and support the proposed thiazole−thiazoline (d_1-d_2) oxidation states. If the β-keto group had been reduced and both thiazoline rings oxidized, the observed shift due to d_2-cysteine would have been 2 Da. The largest mass shift observed at +378 Da is provisionally attributed to the HPTT−β-ketoacyl-enzyme species containing one uncyclized (18 Da) cysteine and an unoxidized double bond (2 Da), a proposed precursor to the +358 Da acyl enzyme (*8*, *9*). This is supported by the analogous mass shifts observed when using stable isotope substrates, although work to unambiguously assign this intermediate is continuing.

The +42 Da species results from the decarboxylation of malonyl-*S*-PKS and reprotonation of the C_2 carbanion to acetyl-*S*-PKS (Figure 4), and presumably is off the pathway to Ybt formation (Scheme 1, middle). Introduction of CD_3-SAM brought forth a mass shift of +6 Da to the +70 Da species (Figures 4E and 5B) that is attributed to isobutyryl-*S*-ACP, arising by a combination of bis-methylation and decarboxylation of malonyl-*S*-ACP (Figure 4D). No monomethyl species has been detected, suggesting the two methylation events are tightly coupled, and that the second methylation is faster than the first. Interestingly, inclusion of stable-isotope substrates produced no detectable mass shift of the +198 and +225 Da species. The structures of these covalent attachments to *S*-ACP remain unknown, and work to ascertain the identity and source of these components is continuing.

With the exact order of acyl intermediate formation, processing, and substrate dependencies remaining unclear at the ACP site, experiments aimed at shedding light on tailoring processes at the ACP site were carried out. Incubation of malonyl-*S*-PKS with SAM, both in the presence and in the absence of HMWP2, yielded no evidence of mono- or bis-methylation (data not shown), consistent with methylation following the decarboxylation and C−C condensation with HPTT-*S*-HMWP2. To ascertain if bis-methylation can occur following uncoupled decarboxylation or if the presence of the complex intermediate upstream (PCP2 or KS) is requisite for any methylation events, holo-PKS was incubated with acetyl-CoA to form the surrogate decarboxylation product at the ACP site (via acetyl transfer through the AT domain). This acetyl-*S*-PKS was incubated with SAM both in the presence and in the absence of holo-HMWP2, and in both cases, bis-methylation did not produce the +70 Da species, arguing that acetyl-*S*-HMWP1 is not readily deprotonated to the C_2 carbanion and is off the reaction pathway as a dead end. Interestingly, however, when acetyl-*S*-PKS was reconstituted with HMWP2 and substrates, a small yet detectable amount of a +70 Da species was observed. This emergence of a +70 Da species suggests that the PKS module indeed has the ability to bis-methylate the acetyl-*S*-enzyme (Scheme 1, top), albeit at amounts (~5%) that are far smaller than that observed from the methylation and decarboxylation of the authentic malonyl moiety (98% occupancy of +70 Da). The reason for efficient conversion of the 42 Da form to the 70 Da form by bis-methylation only when malonyl-*S*-enzyme decarboxylates is now under further study. The isobutyryl-*S*-enzyme will also be a dead end intermediate. It should be noted that attempts to produce highly loaded (>90%) acetyl-*S*-PKS (via acetyl-CoA through the AT domain) at best resulted in only ~40% of the site occupancy of the ACP active site, whereas malonyl-CoA loading was nearly 100%, reflecting the gatekeeper ability of the AT domain.

In an experiment to see if malonyl-*S*-ACP decarboxylation can be triggered by the presence of upstream carrier protein HMWP2 alone, incubation of malonyl-*S*-PKS with apo-HMWP2 showed evidence of only small amounts of decarboxylation when compared with the amount with malonyl-*S*-PKS in the absence of HMWP2. The significant increase in the level of decarboxylation upon incubation with cysteine-loaded HMWP2 can be explained as either decarboxylation is stimulated by the peptidyl-primed upstream protein via protein−protein interactions that ready the PKS module to make the C_2 carbanion needed for C−C bond formation and chain translocation. Alternatively, some amount of cysteine could be nonspecifically loaded onto the KS domain of the PKS module through slow thiol exchange with cysteine-*S*-HMWP2 and the acyl-*S*-KS signals for malonyl-*S*-ACP decarboxylation. Under full reconstitution conditions (Figure 5A), approximately 75% of the ACP covalent site is occupied by the unproductive decarboxylation product isobutyryl-*S*-ACP. This indicates that decarboxylation of malonyl-*S*-ACP is highly stimulated by the presence of the complex intermediate located upstream (at either the PCP2 domain of HMWP2 or the KS domain of HMWP1), although the exact species stimulating this response remains to be seen, and can be addressed by a KS domain C198A mutant.

Experiments aimed at directly observing the condensation product at the ACP site in the absence of SAM were carried out to evaluate whether decarboxylation or condensation precedes methylation. In the absence of SAM but with all substrates and holo-HMWP2 present, ~98% of detectable ACP site occupancy harbored the +42 Da acetyl-*S*-PKS decarboxylation product of malonyl-*S*-PKS, with trace amounts of the condensation product of +330 Da observed without methylation (Figure 5C). The prevalence of the +42 Da product and the marked lack of a complex intermediate imply that the presence of SAM at this key checkpoint enhances the yield of C−C condensation product versus unproductive protonation (Scheme 1 branch point). The amount of this enhancement can be gleaned from the far right of panels A and C of Figure 5. Use of the surrogate pseudosubstrate *S*-adenosylhomocysteine (incompetent for methyl transfer) produced no observable enhancement of the condensation product of +330 Da (Figure 5D). These data are consistent with the Claisen condensation and HPTT transfer to the ACP covalent site being reversible. In the absence of SAM to quickly trap this +330 Da product by bis-methylation (Scheme 1), C−C bond cleavage via retro-Claisen condensation would transfer the HPTT intermediate back upstream (to the KS domain active site cysteine) and regenerate the C_2 carbanion of the ACP active site (Scheme 1). Off pathway protonation generates the +42 Da acetyl-*S*-ACP which cannot readily be deprotonated to rejoin the catalytic flux. It should also be noted that if the reversibility of the condensation product back upstream to the KS domain is loosely controlled (i.e., KS can accept HPTT from a different precursor, namely, HPTT−β-keto-2,2-dimethyl acyl-*S*-ACP), a second branch point may exist at the ketoreductase reaction which returns the HPTT moiety

upstream to leave the isobutyryl-*S*-ACP (+70 Da) seen in considerable abundance.

These results provide direct insight into the order of chemical processes that occur at the ACP domain of the PKS module of HMWP1. Malonyl decarboxylation initiates the action, and is greatly stimulated by the presence of the HPTT-*S*-PCP_2 intermediate upstream and not simply the upstream carrier protein HMWP2. The C_2 carbanion of acetyl-*S*-ACP sits at a branch point, at which either on pathway Claisen condensation (to +330 Da) or off pathway protonation to the acetyl-*S*-ACP decarboxylation product (+42 Da) can occur. In the presence of the cosubstrate SAM, there is no detectable +42 and +330 Da species, indicating bis-methylation (+28 Da) occurs rapidly, to produce the +70 Da (via malonyl-*S*-ACP) and +358 Da species (via the +330 Da β-keto-*S*-ACP) seen in appreciable abundance. The condensation product of +330 Da can accumulate in the absence of SAM, but in barely detectable abundance. An observation of this is that HMWP2 and addition of all substrates to malonyl-*S*-PKS gives ~20% of the +358 and +378 Da products thought to be on pathway. There is a trace amount of acetyl-*S*-ACP and 70% as the isobutyryl-*S*-ACP, presumably a dead end bis-methylation/decarboxylation shunt thioester, indicating that with the NRPS−NRPS−PKS three-module protein of the Ybt synthetase assembly line, 75% of the flux generates dead end intermediates. If this ratio persists in the full four-module HMWP2/HMWP1 assembly line with YbtE and YbtU present, two issues arise. One is that the fractional occupancy of carrier sites by acyl-*S*-enzymes that cannot be elongated may explain the low turnover number (1.4 min^{-1}) (*11*) seen on *in vitro* reconstitution. Second, there will be a major role for surveillance thioesterases (*8*) in the hydrolytic removal of the acetyl-*S*- and isobutyryl-*S*-acyl groups from the ACP.

CONCLUSION

This study uses MS to directly observe multiple intermediates bound to a PKS module in a hybrid NRPS/PKS assembly line. By detecting multiple acyl-*S*-ACP intermediates that are both on (+86, +330, and +358 Da) and off (+42 and +70 Da) the Ybt biosynthetic pathway, we can now discern the extent to which unproductive decarboxylation reduces the yield of the natural product. This approach can be extended to deciphering the identity and timing of any ACP domain in any PKS assembly line, as well as growing peptidyl chains on the corresponding PCP waystations of NRPS carrier protein domains. This will enable a far more detailed understanding of NRPS and PKS systems for the improved engineering of "non-natural" synthetases for more efficient creation of new bioactive compounds.

ACKNOWLEDGMENT

We thank Deborah Miller, Lusong Luo, Fei Liu, Leslie Hicks, and Steven Patrie for assistance with this work.

REFERENCES

1. Cane, D. E., Walsh, C. T., and Khosla, C. (1998) *Science 282*, 63−68.
2. Walsh, C. T. (2002) *ChemBioChem 3*, 124−134.
3. Kelleher, N. L. (2000) *Chem. Biol. 7*, R37−R45.
4. McLafferty, F. W. (1994) *Acc. Chem. Res. 27*, 379−386.
5. Schubert, S., Rakin, A., Karch, H., Carniel, E., and Heesemann, J. (1998) *Infect. Immun. 66*, 480−485.
6. Carniel, E. (2001) *Microbes Infect. 3*, 561−569.
7. Bearden, S. W., Fetherston, J. D., and Perry, R. D. (1997) *Infect. Immun. 65*, 1659−1668.
8. Gehring, A. M., DeMoll, E., Fetherston, J. D., Mori, I., Mayhew, G. F., Blattner, F. R., Walsh, C. T., and Perry, R. D. (1998) *Chem. Biol. 5*, 573−586.
9. Gehring, A. M., Mori, I., Perry, R., and Walsh, C. T. (1998) *Biochemistry 37*, 11637−11650.
10. Suo, Z., Walsh, C. T., and Miller, D. A. (1999) *Biochemistry 38*, 14023−14035.
11. Miller, D. A., Luo, L., Hillson, N., Keating, T. A., and Walsh, C. T. (2002) *Chem. Biol. 9*, 333−344.
12. Keating, T. A., Suo, Z., Ehmann, D. E., and Walsh, C. T. (2000) *Biochemistry 39*, 2297−2306.
13. Suo, Z., Tseng, C. C., and Walsh, C. T. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 99−104.
14. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) *Science 246*, 64−71.
15. Leenders, F., Vater, J., Stein, T., and Franke, P. (1998) *J. Biol. Chem. 273*, 18011−18014.
16. Suo, Z., Chen, H., and Walsh, C. T. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 14188−14193.
17. Quadri, L. E. N., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) *Biochemistry 37*, 1585−1595.
18. Keating, T. A., Miller, D. A., and Walsh, C. T. (2000) *Biochemistry 39*, 4729−4739.
19. Patrie, S. M., Quinn, J. P., Whipple, D., Blakney, G. T., Marshall, A. G., Hendrickson, C. L., and Kelleher, N. L. (2001) in *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics*, Chicago.
20. Horn, D. M., Zubarev, R. A., and McLafferty, F. W. (2000) *J. Am. Soc. Mass Spectrom. 11*, 320−332.
21. Senko, M. W., Beu, S. C., and McLafferty, F. W. (1995) *J. Am. Soc. Mass Spectrom. 6*, 229−233.
22. Kelleher, N. L., Zubarev, R. A., Bush, K., Furie, B., Furie, B. C., McLafferty, F. W., and Walsh, C. T. (1999) *Anal. Chem. 71*, 4250−4253.
23. Jaenicke, L., and Lynnen, F. (1960) *Enzymes 3*, 30.
24. Jencks, W. P., and Gilchrist, M. (1964) *J. Am. Chem. Soc. 86*, 4651−4654.

BI035585Z