Chemoenzymatic Approaches for Streamlined Detection of Active Site Modifications on Thiotemplate Assembly Lines Using Mass Spectrometry[†]

Shaun M. McLoughlin,[‡] Matthew T. Mazur,[‡] Leah M. Miller,[‡] Jun Yin,[§] Fei Liu,[§] Christopher T. Walsh,[§] and Neil L. Kelleher*,[‡]
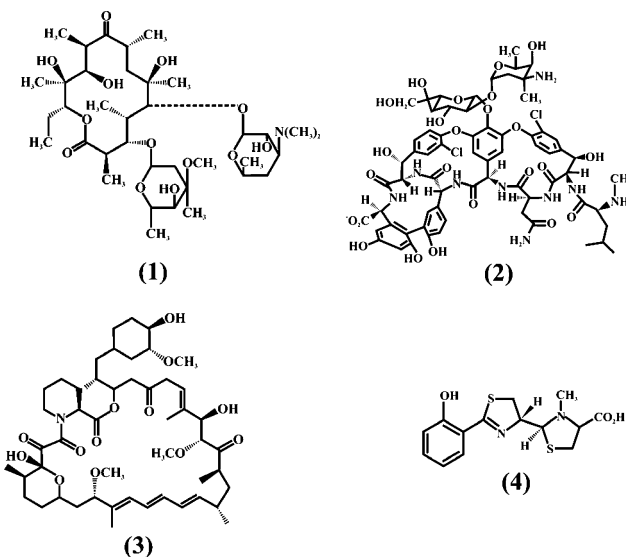
Department of Chemistry, University of Illinois, Urbana, Illinois 61801, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

ABSTRACT: For the direct interrogation of peptides harboring covalently modified serines in nonribosomal peptide synthetases, streamlined methodologies described here employ proteolysis and reporter−coenzyme A analogues of four types. The chromophoric and fluorescent coenzyme A analogues pyrene-maleimidyl-*S*-CoA and BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA were enzymatically loaded onto the active site serines harbored in the ArCP, PCP1, and PCP2 thiolation domains of PchE and PchF, the nonribosomal peptide synthetases responsible for the biosynthesis of the siderophore pyochelin. During the chromatographic separation of cyanogen bromide digests, observation of the absorbance (at 338 and 504 nm) or fluorescence (after irradiation at 365 nm) enabled the selective detection of peptides containing each active site serine. This resulted in quick detection of each active site peptide by Fourier transform mass spectrometry in the fully reconstituted pyochelin system. The loading of short acyl chain reporters in equimolar quantities permitted further insights into digestion heterogeneity and side reactions by virtue of a mass shift signature on each active site peptide. The chromatographic shift of the reporter-loaded peptides relative to peptides carrying on pathway intermediates was 2 min at 7 kDa, providing a general strategy for efficient localization of "carrier" peptides in complex digests of thiotemplate enzymes. Also, the use of the affinity reporter, biotin-maleimidyl-*S*-coenzyme A, permitted the isolation of intact synthetases at high purity via removal of contaminating *Escherichia coli* proteins.

Many therapeutically important natural products and bioactive peptides are assembled nonribosomally on large, multidomain enzyme scaffolds known as nonribosomal peptide synthetases (NRPSs)[1] and polyketide synthases (PKSs) (*1−3*). These bioactive products include several pharmaceuticals that are commonly used in clinical environments today. Erythromycin (**1**) and vancomycin (**2**), for example, are potent antibiotics, while other natural products such as rapamycin (**3**) and epothilone exhibit immunosuppressive and anticancer activities (*4−7*). Additionally, some strains of pathogenic bacteria utilize NRPSs and PKSs to synthesize iron-scavenging compounds (siderophores) such as pyochelin (**4**), which can contribute to bacterial virulence (*8−10*).

With the rising resistance of pathogens to common clinical treatments, the need for novel antibiotic and antiviral

(1)

(2)

(3)

(4)

* To whom correspondence should be addressed: Department of Chemistry, University of Illinois, 53 Roger Adams Laboratory, Box 47-5, 600 S. Mathews Ave., Urbana, IL 61801. Phone: (217) 244-3927. Fax: (217) 244-8068. E-mail: kelleher@scs.uiuc.edu.
[‡] University of Illinois.
[§] Harvard Medical School.

[1] Abbreviations: NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase; PCP, peptidyl carrier protein; ArCP, aryl carrier protein; CNBr, cyanogen bromide; IPTG, isopropyl β-D-thiogalactopyranoside; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; FTMS, Fourier transform mass spectrometry; CoA, coenzyme A.

pharmaceuticals is accelerating rapidly. The "modular" configuration of NRP/PK synthetases has sparked an interest in the reengineering of such enzymes to produce "unnatural" natural products (*11−14*). To successfully generate hybrid enzymes capable of combinatorially producing therapeutically useful compounds, however, it is essential to better understand the covalently driven "assembly line" mechanisms by which these enzymes function.

In several recent studies, Fourier transform mass spectrometry (FTMS) has been employed to directly visualize

Scheme 1: Identification of Active Site Peptides Using Reporter CoA Analogues and Mass Spectrometry
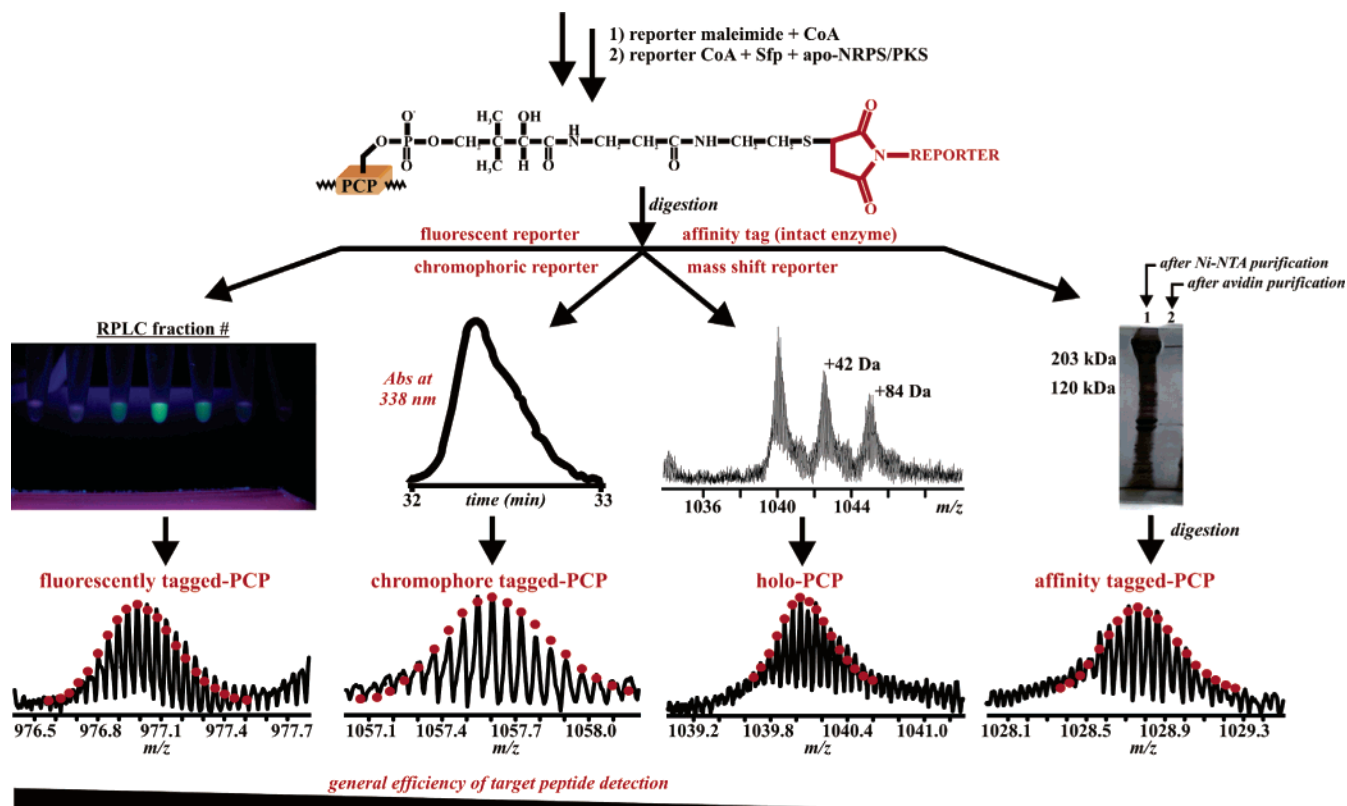
the bioassembly processes on various NRP/PK synthetases (*15−19*). In these works, NRPSs and PKSs were subjected to digestion and fractionation after in vitro reconstitution. Subsequent identification and mass analysis of the peptides harboring the active site serines from the covalently modified thiolation domains allowed for the structural resolution (within <1 Da) and localization of the complex intermediates during the biosynthetic processes. By monitoring changes in the relative occupancies of complex intermediates and monomers over time, we elucidated substrate loading trends and kinetics.

Despite the potential of using mass spectrometry to analyze these synthetases, the measurement platform is limited by the rate of target peptide identification because of the complexity of the digestion mixtures and the speed of RPLC fraction analysis. The vast number of peptides generated by the digestion of such large enzymes coupled with complicating side reactions, nonexhaustive digestion, contaminating *Escherichia coli* peptides, unexpected modifications, and sequence errors necessitates extensive fractionation, careful interrogation of each chromatographic fraction by tandem mass spectrometry, and lengthy optimization of the digestion protocol. Additionally, mass spectrometer-related phenomena such as signal suppression during ionization can prevent the successful identification of key peptides. Therefore, the development of methodologies for targeting the peptide search to specific chromatographic regions will expedite mass spectral analysis.

In this work, we describe the implementation of a cyanogen bromide digestion protocol and the use of several coenzyme A analogues with spectroscopic, Δ mass, and affinity properties for mapping the modified active site peptides from the pyochelin synthetase biosynthetic enzymes,

PchE (159 kDa) and PchF (198 kDa) (Figures 6 and 7 of the Supporting Information) (*20−22*). These "reporter"−coenzyme A analogues were exclusively loaded onto the conserved active site serines within the thiolation domains of PchE and PchF using the indiscriminate 4′-phosphopantetheinyl transferase, Sfp (*23*). After digestion and fractionation, the active site peptides from the aryl carrier protein (ArCP) and peptidyl carrier protein (PCP) domains were rapidly identified using UV and mass spectral detection by virtue of the signature fluorescence, absorbance, and mass shift profiles of the reporter molecules (Scheme 1). Purification of intact PchE modified with an affinity CoA analogue was also effective in removing contaminating *E. coli* proteins, resulting in less complex digestion mixtures for active site peptide mapping (*24, 25*). These approaches permitted directed detection of biosynthetic intermediates from pyochelin synthetase, and initial kinetic data were also obtained.

EXPERIMENTAL PROCEDURES

Materials. Plasmids for PchD, PchE, and PchF are the same as in prior reports (*26−28*). Competent BL21(DE3) *E. coli* was purchased from Invitrogen (Carlsbad, CA). Ampicillin- and kanamycin-treated LB agar plates were purchased from the cell and media facility at the University of Illinois. Isopropyl β-D-thiogalactopyranoside (IPTG) was purchased from Promega (Madison, WI). Nickel affinity (Ni−NTA) chromatography resin and accessories were purchased from Qiagen (Valencia, CA). DEAE medium was purchased from Amersham Biosciences (Piscataway, NJ). The Bradford protein assay was purchased from Bio-Rad (Hercules, CA). Centriplus protein concentrators and C18 ZipTips were purchased from Millipore (Billerica, MA). Jupiter C4 columns and accessories were purchased from

Phenomenex (Torrance, CA). BODIPY-FL-*N*-(2-aminoethyl)maleimide was purchased from Molecular Probes (Eugene, OR). Nanospray emitter nozzles and sample tips were purchased from Advion BioSciences (Ithaca, NY). All other chemicals were purchased from Sigma-Aldrich (Milwaukee, WI). Synthesis of pyrene-maleimidyl-*S*-CoA and biotin-maleimidyl-*S*-CoA was accomplished by the Walsh Laboratory at Harvard Medical School (*24*, *25*).

Expression and Purification of PchD, PchE, and PchF. Plasmids harboring the pyochelin synthetase genes for PchD (pET29b), PchE (pET28b), and PchF (pET29b) were transformed into the BL21(DE3) strain of *E. coli*. PchD was cultured 1×1 L at 37 °C with agitation at 200 rpm in LB medium supplemented with 25 μg/mL kanamycin, while PchE and PchF were cultured similarly, only at 30 °C. The OD_{600} was monitored periodically, and when it reached 0.4, the PchD culture was induced with 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) while the temperature of the PchE and PchF constructs was lowered to 24 °C. At an OD_{600} of 0.6, the PchE and PchF cultures were induced with 0.5 mM IPTG. Post induction growth continued for 4 h for PchD and 12 h for PchE and PchF before the cultures were harvested for 10 min at 6000*g*. The cell pellets were resuspended in 30 mL of 25 mM Tris-HCl (pH 8), 500 mM NaCl, and 5 mM imidazole before lysis by two passages through a French pressure cell. The lysates were clarified for 20 min at 99000*g* before incubation with Ni−NTA resin (2 mL for PchD and 8 mL for PchE and PchF) at 4 °C using slow rotation for 2 h. The resin from each construct was packed into polypropylene columns and washed extensively with lysis buffer. The bound enzymes were eluted from the resin by treatment with 2 mL aliquots of lysis buffer and increasing concentrations of imidazole (5, 10, 20, 30, 40, 50, 100, and 500 mM). The eluents were subjected to SDS−PAGE; those fractions containing PchD were pooled and dialyzed against 25 mM Tris-HCl (pH 8), 0.5 mM TCEP, 10 mM $MgCl_2$, and 10% glycerol, while those containing PchE and PchF were pooled and dialyzed against 25 mM Tris-HCl (pH 8), 0.5 mM TCEP, and 10% glycerol. The enzymes were concentrated using YM50 Centriplus concentrators before being flash-frozen in 50 μL aliquots and stored at −80 °C. Concentrations were determined by the Bradford assay using IgG as a standard.

Synthesis and Purification of Coenzyme A Analogues. Biotin-maleimidyl-*S*-coenzyme A (**1a**) and pyrene-maleimidyl-*S*-coenzyme A (**2a**) were obtained from the Walsh Laboratory (see the Supporting Information). BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-coenzyme A (**3a**) was synthesized according to previously reported protocols (*22*).

Briefly, coenzyme A (2.5 mM) and BODIPY-FL-*N*-(2-aminoethyl)maleimide (2.5 mM) were added to 100 mM MES (pH 6), 100 mM MgOAc (pH 6), and 15% DMSO to a final volume of 2 mL. The mixture was vortexed, and then placed at 4 °C in the dark for 30 min. After incubation for an additional 10 min at room temperature, the reaction was assessed by TLC. Unreacted dye was removed by extracting three times into EtOAc. The remaining material was divided into 20 μL aliquots and frozen at −80 °C.

The λ_{max} values for BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-coenzyme A and pyrene-maleimidyl-*S*-coenzyme A were measured on a Genesys model 10UV UV−vis spectrophotometer (ThermoSpectronic) between 300 and 700
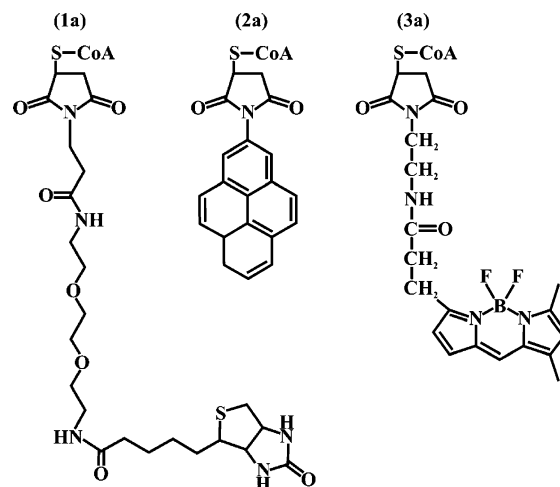
nm in a 50:50 H_2O/CH_3CN mixture in 0.1% TFA. The analogues exhibited absorbance maxima at 504 and 338 nm, respectively (data not shown).

Loading Coenzyme A Analogues onto Pyochelin Biosynthetic Enzymes. PchE (21 μM) and PchF (10 μM) were separately combined with Sfp (2 μM) and each coenzyme A analogue (300 μM) to the final concentrations parenthetically indicated. The reaction mixtures were incubated in the dark at 30 °C for 4 h (30 min for acetyl-*S*-CoA and acetoacetyl-*S*-CoA) before the reactions were quenched into 30% formic acid and 1 M CNBr. The digestions were incubated at 4 °C for 12 h before lyophilization. The pellets were resuspended in NH_4OAc (pH 4), 6 M urea, 10% CH_3CN, and 10 mM TCEP prior to injection onto a Jupiter C4 reversed phase column. After being washed in a 90% H_2O/10% CH_3CN mixture (with 0.1% TFA) for 10 min, the peptides were eluted in CH_3CN from 10 to 30% over the course of 5 min and from 30 to 55% over the course of 25 min at a rate of 1 mL/min. The absorbance was monitored at 220 nm for all fractionations, at 338 nm for pyrene-maleimidyl-*S*-CoA-loaded peptides, and at 504 nm for BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA-loaded peptides. Fractions were collected every 1 min and lyophilized prior to mass spectral analysis.

Affinity Purification of Biotin-Maleimidyl-S-CoA-Loaded PchE. PchE (56 μM) was combined with Sfp (2 μM) and biotin-maleimidyl-*S*-CoA (53.2 μM) to the final concentrations parenthetically indicated. After incubation at 30 °C for 4 h, the solution was diluted to a total volume of 20 mL in 50 mM Tris-HCl (pH 8). The reaction mixture was mixed with 2 mL of activated Softlink avidin resin and allowed to stir at room temperature for 2 h. The resin was packed into a polypropylene column and extensively washed with 50 mM Tris-HCl (pH 8). The affinity-tagged PchE was eluted from the resin with successive 1 mL treatments of 50 mM Tris-HCl (pH 8) and 5 mM biotin. The fractions were subjected to polyacrylamide gel electrophoresis and silver staining. The fractions containing PchE were pooled and dialyzed against 50 mM Tris-HCl (pH 8) for 12 h. The eluent was concentrated using Centriplus YM 50 concentrators before cooling to 4 °C. Serial dilutions of the concentrated PchE were run on gels against the unpurified material and subjected to silver staining.

Full Reconstitution of Pyochelin Synthetase. PchE (10 μM) was combined with PchF (10 μM), PchD (10 μM), Sfp (1
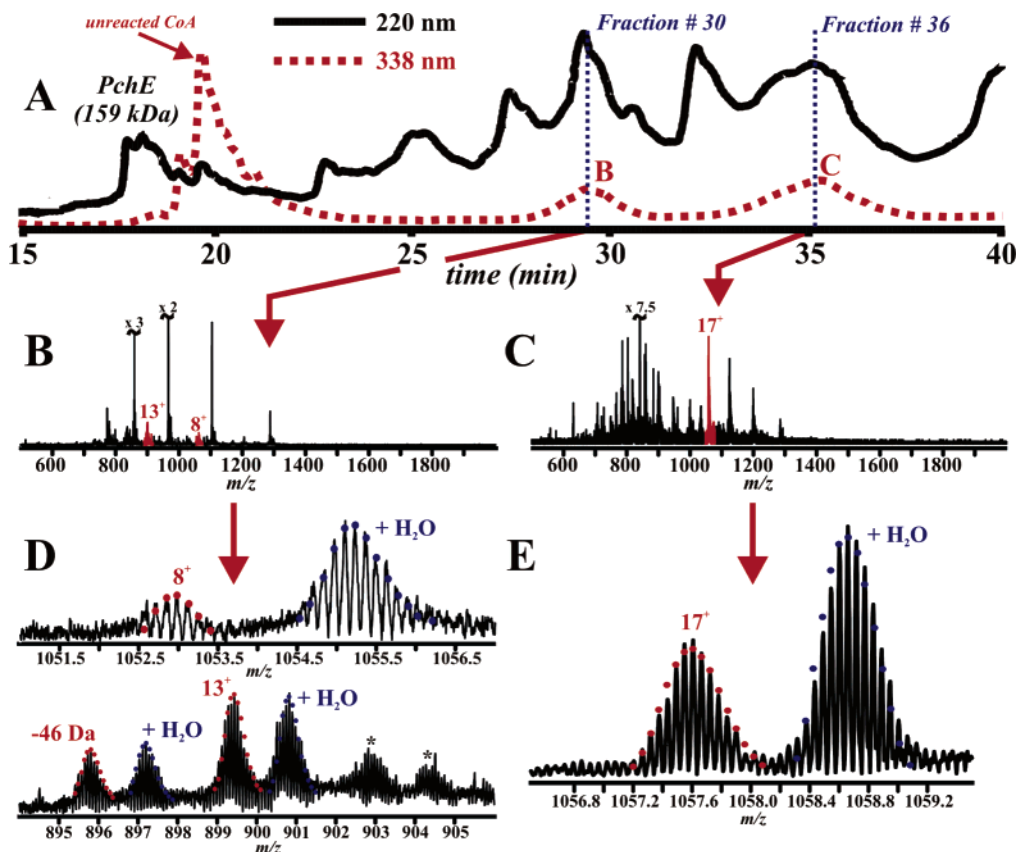
FIGURE 1: Loading PchE with pyrene-maleimidyl-*S*-CoA, followed by digestion, reduction, fractionation, and mass spectral analysis. (A) UV trace at 220 and 338 nm of the digestion mixture over a C4 reversed phase column. The fractions of interest are denoted. (B) Broadband mass spectrum of chromatographic fraction 30. Key peptides are denoted by their charge state, and some peaks were truncated by 2−3-fold for spectral clarity. (C) Broadband mass spectrum of chromatographic fraction 36. Again, the charge state indicates a key peptide, and a peak was truncated approximately 8-fold for spectral clarity. (D) The 8^+ and 13^+ charge states of the exhaustive (top) and unexhaustive (bottom) PCP1 active site peptides. Modified species are denoted. (E) The 17^+ charge state of the unexhaustive ArCP active site peptide.

μM), coenzyme A (200 μM), salicylic acid (1 mM), L-cysteine (5 mM), NADPH (0.75 mM), and SAM (0.75 mM) and diluted to the parenthetically indicated final concentrations in 100 mM Tris-HCl (pH 8), 10 mM MgCl$_2$, and 1 mM TCEP. After a 4 h incubation at 30 °C in the dark, synthesis of pyochelin was initiated by adding 5 mM ATP. The reaction continued at 30 °C in the dark for 30 min before being quenched into 30% cold formic acid and 1 M cold CNBr. The mixture was digested for 12 h at 4 °C in the dark before being lyophilized. The pellet was resuspended in 100 mM NH$_4$OAc (pH 4), 6 M urea, 10% CH$_3$CN, and 10 mM TCEP for 30 min at room temperature before the aforementioned fractionation and mass spectral analysis protocols were performed.

FTMS Detection of the Active Site Peptides. The lyophilized RPLC fractions were resuspended in 20−40 μL of a 50:50 H$_2$O/CH$_3$CN mixture with 1% formic acid. The samples were infused into a custom-built quadrupole FTMS hybrid instrument through nanoelectrospray nozzles on a Nanomate 100 robot (Advion BioSciences). After the ions had been passed through a resistively heated metal capillary, they were externally accumulated for 500 ms in an accumulation octupole (*29*). The ions were then shuttled to the analyzer cell through a quadrupole that can function as an ion guide or a selective mass filter. External calibration of the spectra was based on bovine ubiquitin (8559.64 Da). The isotopic distributions of the active site peptides were fit using a least-squares algorithm provided by the FTMS data

analysis program THRASH (*30*). Upon comparison of the least-squares fit to a theoretical fit of the distribution, the m/z values of both the monoisotopic and most abundant isotopic peaks of the peptides were determined. All masses reported in this study are monoisotopic molecular weights (M_r) unless otherwise specified.

RESULTS

Identification of Active Site Peptides from PchE and PchF Using Chromophoric Coenzyme A Analogues. PchE and PchF were incubated separately with Sfp and pyrene-maleimidyl-*S*-CoA before cyanogen bromide digestion and fractionation. The pyrene absorbance at 338 nm revealed two HPLC peaks, cresting at 30 and 36 min for the PchE digestion (Figure 1A), and one major peak, cresting at 35 min, for the PchF digestion (Figure 2A). The fractions containing these peaks were collected, lyophilized, and infused into the mass spectrometer (Figures 1B,C and 2B). Fraction 30 from the PchE digestion revealed a species with a mass of 8410.57 Da, which matched that of the pyrene-maleimidyl-labeled exhaustive CNBr digestion product containing the active site serine of the PCP1 thiolation domain within 0.38 Da (45 ppm error, Figure 1D, top). The peptide also displayed a satellite peak of 18 Da, consistent with a degree of hydration. Another species with a mass of 11 671.5 Da also displayed this 18 Da adduct peak, suggesting loading with pyrene-maleimidyl-*S*-CoA (Figure 1D, bottom). Tandem mass spectrometry identified the peptide as the unexhaustive
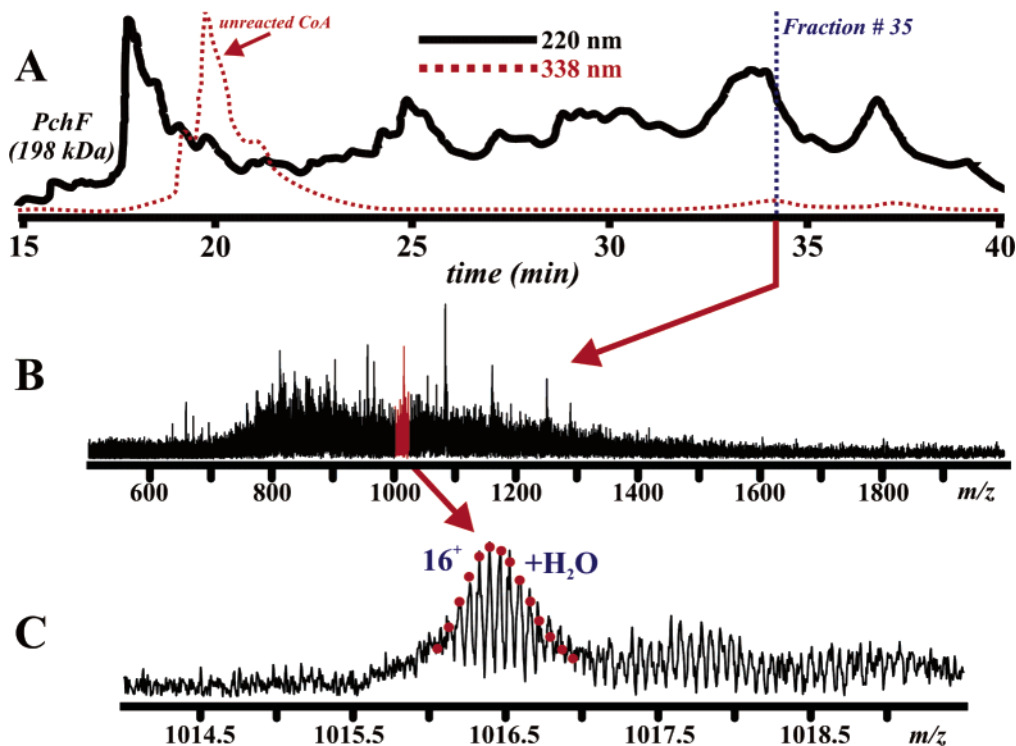
FIGURE 2: Loading PchF with pyrene-maleimidyl-*S*-CoA, followed by digestion, reduction, fractionation, and mass spectral analysis. (A) UV trace at 220 and 338 nm of the digestion mixture over a C4 reversed phase column. (B) Broadband mass spectra of chromatographic fraction 35. The key peptide is denoted by its charge state. (C) The 16+ charge state of the exhaustive PCP2 active site peptide. This peptide was only identified in its hydrated form.

CNBr product ($Arg^{1328}-Met^{1425}$) with an oxidized Met^{1335}, matching within 0.6 Da (51 ppm error, Figure 3 of the Supporting Information). The peptide also displayed another satellite peak of −46 Da (11 625.4 Da, Figure 1D, bottom). Fraction 36 of the PchE digestion revealed a peptide with a mass of 17 951.1 Da, which correlated with the mass of the nonexhaustive CNBr product containing the ArCP active site serine, where Met^{50} was oxidized, within 0.2 Da (11 ppm error, Figure 1E). Mass spectral analysis of fraction 35 from the PchF reaction mixture revealed a peptide with a mass of 16 237.2 Da, which matched that of the predicted, exhaustive CNBr peptide harboring the PCP2 active site serine within −0.4 Da (25 ppm error, Figure 2C). The identities of all peptides were verified using tandem mass spectrometry (Figures 1−4 of the Supporting Information).

Separate incubations of PchE and PchF with BODIPY-FL-*N*-(2-aminoethyl)-maleimidyl-*S*-CoA followed by digestion and fractionation revealed chromatograms similar to those of the pyrene-maleimidyl-*S*-CoA incubations, where RPLC peaks with absorptions at 504 nm were observed at 30 and 37 min in the PchE digestion (Figure 3A) and at 35 min in the PchF digestion (Figure 4A). After lyophilization and resuspension of these fractions in ESI buffer, fractions 30 and 37 of the PchE digestion exhibited the strongest fluorescence when irradiated at 365 nm (Figure 3B). Targeted Q-FTMS analysis of fraction 30 revealed two species with masses of 8527.45 and 11 788.6 Da, which corresponded to the previously observed exhaustive and unexhaustive CNBr products possessing the PCP1 active site serine within 0.17 (20 ppm error) and −0.1 Da (5 ppm error), respectively (Figure 3C,D). Infusion of fraction 37 from the PchE digestion into the mass spectrometer resulted in the observation of a species with a mass of 18 068.0 Da, which also

correlated with that of the previously observed CNBr product containing the ArCP active site serine within 0.1 Da (6 ppm error, Figure 3E). Analysis of fraction 35 from the PchF digestion revealed a peptide with a mass of 16 336.1 Da, which correlated with that of the previously observed PCP2 active site peptide within −0.6 Da (37 ppm error, Figure 4B). The identities of all peptides were verified using tandem mass spectrometry.

Fast Observation of Acylated Active Site Peptides from PchE and PchF. Having targeted the fractions containing the active site peptides using chromophoric CoA analogues, we loaded a series of low-mass acyl-CoA's onto each active site to visualize holo- and thioester-loaded forms of the active site peptides. Each enzyme was incubated in equimolar amounts of coenzyme A, acetyl-*S*-coenzyme A, and acetoacetyl-*S*-coenzyme A for 30 min at 30 °C prior to quenching, digestion, and fractionation. Fractions 25−40 of the PchE fractionation were profiled, and a total of five peptides displayed the distinctive pattern of 42 Da mass shifts (Figure 5). Of these peptides, three displayed masses that correlated with the respective holo form masses of the ArCP (17 653.5 Da, Figure 5E) and PCP1 (8113.45 Da, Figure 5B; 11 373.3 Da, Figure 5C) peptides previously observed using the chromophoric CoA analogues. Another peptide in fraction 35 with a mass of 25 417.2 Da matched another unexhaustive product harboring the ArCP domain, possessing oxidized Met^{50} and Met^{158} (Figure 5D). Fraction 29 also revealed another peptide with a mass of 15 097.8 Da, putatively matching another PCP1 unexhaustive product, where Met^{1335} and Met^{1405} had been oxidized (Figure 5A). Both peptides were substantially lower in abundance than the peptides previously identified with the chromophoric CoA analogues.
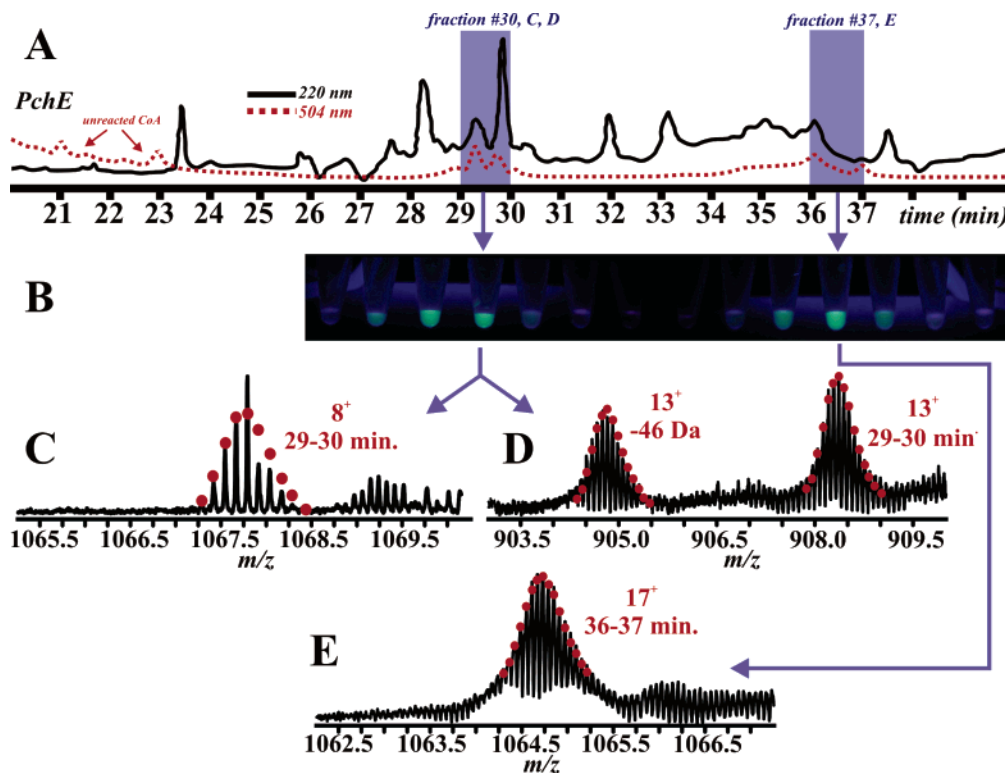
FIGURE 3: Loading PchE with BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA, followed by digestion, reduction, fractionation, and mass spectral analysis. (A) UV trace at 220 and 504 nm of the digestion mixture over a C4 reversed phase column. (B) Fluorescence of selected chromatographic fractions. The most intense fluorescence was confined to fractions 29, 30, 36, and 37. (C) The 8^+ charge state of the exhaustively digested PCP1 active site peptide. (D) The 13^+ charge state of the unexhaustively digested PCP1 active site peptide. (E) The 17^+ charge state of the unexhaustive ArCP active site peptide.
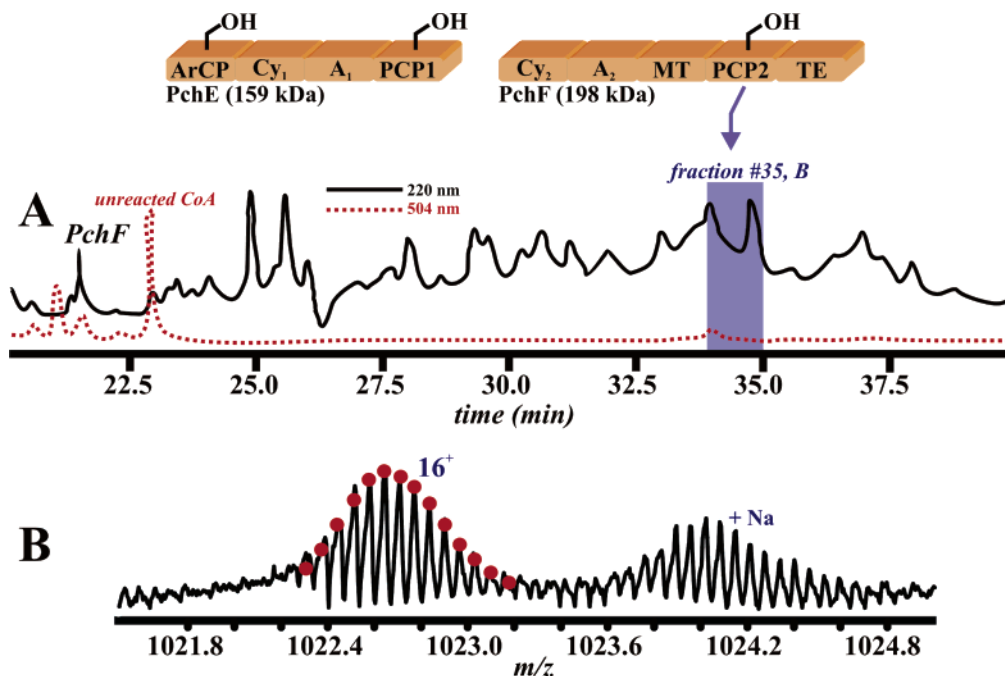

FIGURE 4: Loading PchF with BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA, followed by digestion, reduction, fractionation, and mass spectral analysis. (A) UV trace at 220 and 504 nm of the digestion mixture over a C4 reversed phase column. (B) The 16^+ charge state of the exhaustive PCP2 active site peptide.

A survey of chromatographic fractions from a CNBr digestion of PchF revealed only one peptide that harbored a single 42 Da mass increase (Figure 5 of the Supporting Information). This peptide's mass, 15 921.7 Da, corresponded to the exhaustive holo form PCP2 active site peptide within -0.8 Da (50 ppm error). No other peptides possessing the PCP2 active site serine were observed.

Purification of Recombinant PchE Using Biotin-Maleimidyl-S-CoA. The biotin-maleimidyl-*S*-CoA analogue was loaded onto PchE and batch purified from avidin resin over
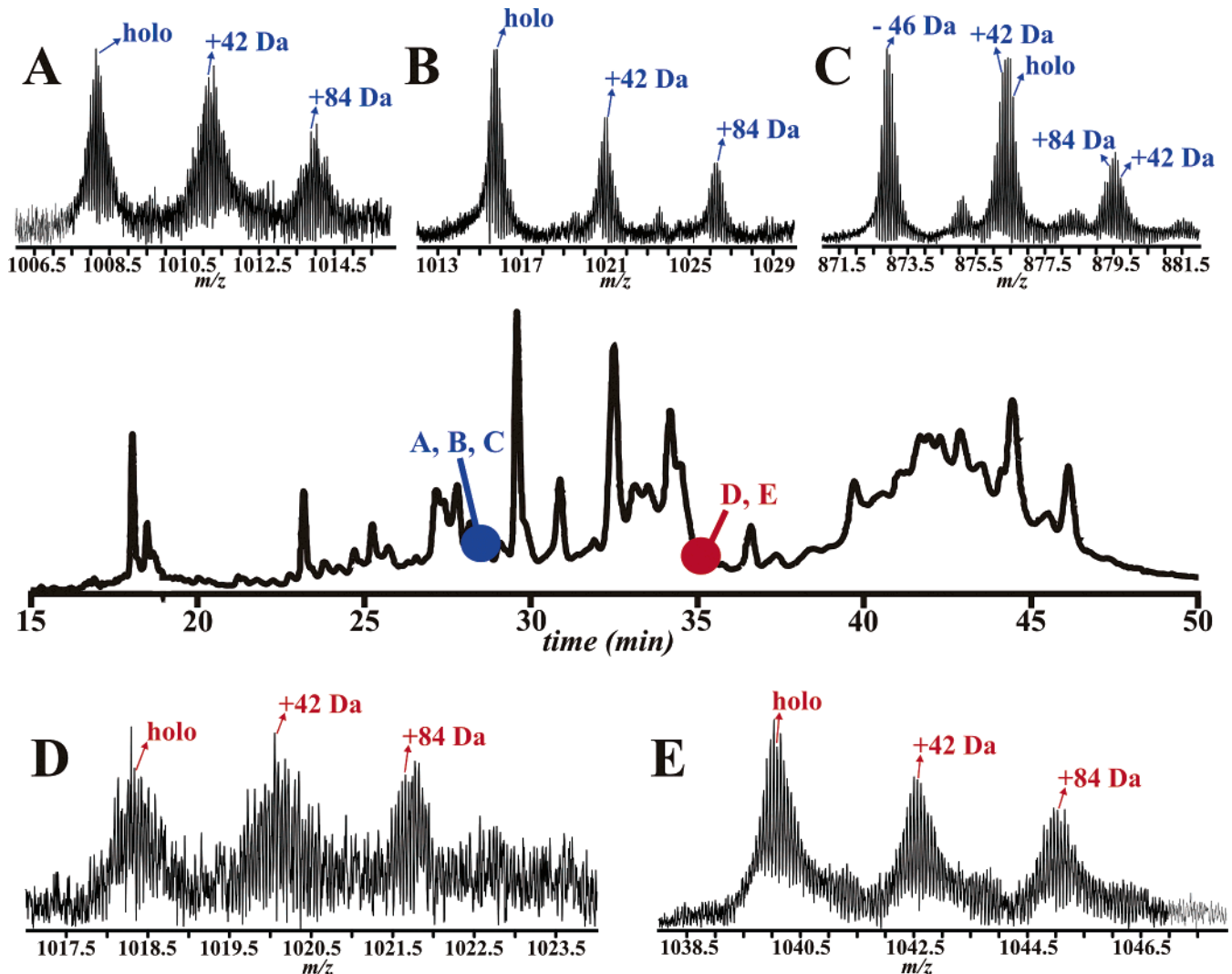
FIGURE 5: Loading PchE with CoA, acetyl-*S*-CoA, and acetoacetyl-*S*-CoA at equimolar concentrations. (A) Mass shift profile of a previously undiscovered unexhaustive PCP1 active site peptide, shown to incorporate acetyl and acetoacetyl moieties. This peptide represented two missed methionyl cleavages. (B) Exhaustive peptide containing the PCP1 active site serine. (C) Mass shift profile for the unexhaustive PCP1 active site peptide. Because of the −46 Da form, each mass shift results in a broadened isotopic distribution of 42 and 46 Da. (D) Previously undiscovered unexhaustive peptide harboring the ArCP active site serine and representing two missed cleavages. (E) Mass shift profile of the unexhaustively digested peptide harboring the ArCP active site serine, representing only one missed cleavage.
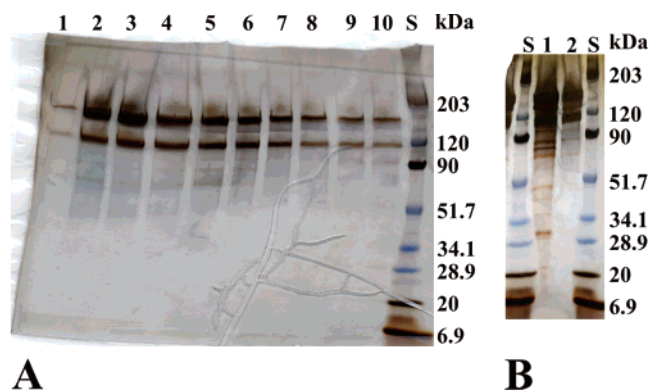

FIGURE 6: Batch affinity purification of biotin-maleimidyl-*S*-CoA-loaded PchE. (A) Elution of PchE from the avidin resin over 10 × 1 mL fractions. (B) Avidin-purified PchE vs a PchE control. Lane 1 contained unpurified PchE and lane 2 the concentrated avidin-purified PchE. The S lanes contained a broad range of standards.

10 fractions. SDS−PAGE analysis of the fractions displayed successful elution of PchE from the resin (Figure 6A). The Coomassie stains were faint despite the use of a large amount

of enzyme initially, and silver staining was required. After dialysis and concentration, a silver-stained SDS−PAGE gel of the concentrated PchE versus the unpurified PchE revealed that the enzyme can be removed from its contaminants (Figure 6B). Another band, slightly under the weight of the intact PchE, was also detected.

Full Reconstitution of the Pyochelin Biosynthetic Pathway. With extensive knowledge of the CNBr digestion behavior and elution times of covalently modified "carrier peptides", PchE and PchF were reconstituted with PchD, Sfp, coenzyme A, salicylic acid, L-cysteine, NADPH, *S*-adenosylmethionine (SAM), and ATP in an effort to observe monomer loading and formation of the complex intermediates. After digestion and fractionation (Figure 7A), mass analysis of fraction 35 yielded the identification of a CNBr peptide with a mass of 17 773.7 Da, which matched that of the salicyl-*S*-ArCP intermediate within −0.2 Da (11 ppm error) present at 100% occupancy (Figure 7B). The other nonexhaustive peptide harboring the ArCP domain active site serine described using the mass shift assay was not observed. Analysis of fraction 30 yielded the identification of two peptides, with masses
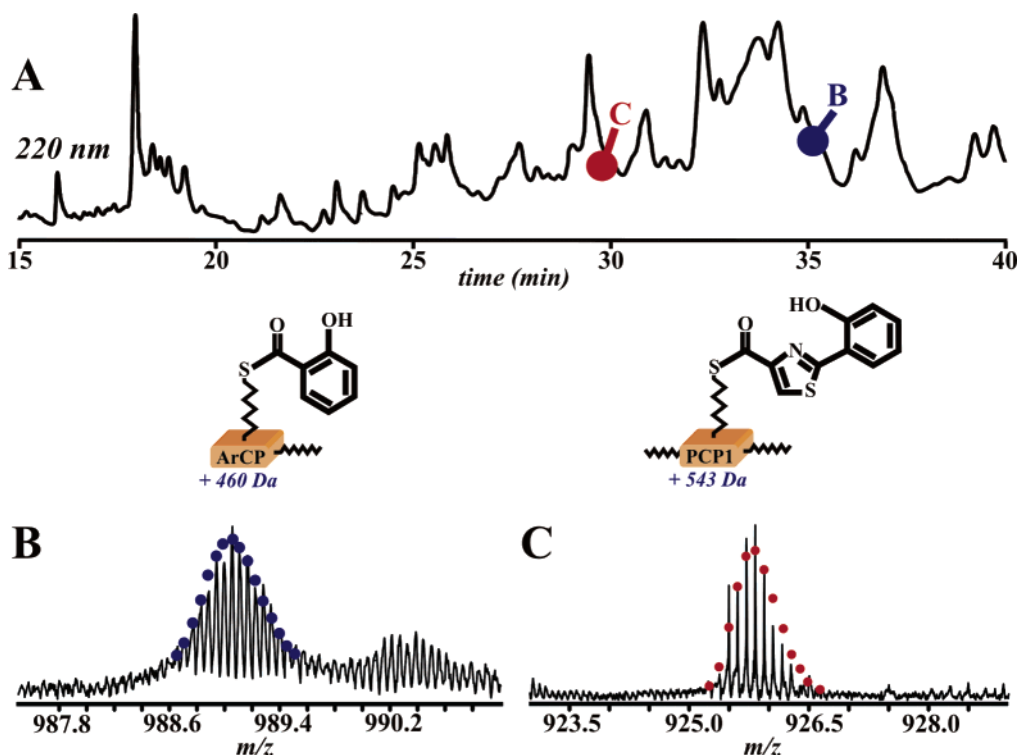
FIGURE 7: Reconstitution of pyochelin synthetase. (A) Chromatographic separation of a CNBr-digested equimolar mixture of PchD, PchE, and PchF under full reconstitution conditions. (B) Structure and mass spectrum of the salicyl-*S*-ArCP complex intermediate. (C) Identification of the hydroxyphenylthiazolyl-*S*-PCP1 intermediate.

of 8318.23 Da (0.1 Da, 12 ppm error) and 11 579.3 Da (0.4 Da, 35 ppm error), correlating with the exhaustive and nonexhaustive CNBr peptides, respectively, of the PCP1 active site containing the hydroxyphenylthiazolyl-*S*-PCP1 intermediate (HPT-*S*-PCP1, Figure 7C). This complex intermediate was present at 100% occupancy, with no evidence of the monomer-loaded or uncyclized forms. Last, analysis of fraction 33 resulted in the identification of a peptide with a mass of 15 921.8 Da, which matched that of the holo form PCP2 peptide described above within −0.7 Da (44 ppm error). The cysteinyl- and hydroxyphenylthiazolylthiazolinyl-*S*-PCP2 (HPTT-*S*-PCP2) complex intermediates could not be identified. To examine the flux of complex intermediates through the synthetase, PchD, PchE, and PchF were reconstituted in a 1:1:1 molar ratio with all appropriate substrates and cofactors. After 20 min, d_6-salicylic acid was introduced before quenching. Observation of the 9^+ charge state of the exhaustively digested CNBr peptide possessing the PCP1 domain active site serine revealed incorporation of four deuteriums over a span of 30 min, consistent with the structure shown in Figure 7C (Figure 8).

DISCUSSION

The prerequisites for mass spectral analysis of NRP/PK synthetases are the successful digestion of these >100 kDa enzymes, separation of the peptide mixture into fractions of reasonable complexity, and the identification of the peptides containing the covalently modified active site serines from each thiolation domain while preserving the base-labile thioester-bound intermediates. While proteolytic digestions are widely employed and fairly well-behaved for the analysis of complex samples, chemical digestions are less textbook in nature, often inducing side reactions in addition to missed
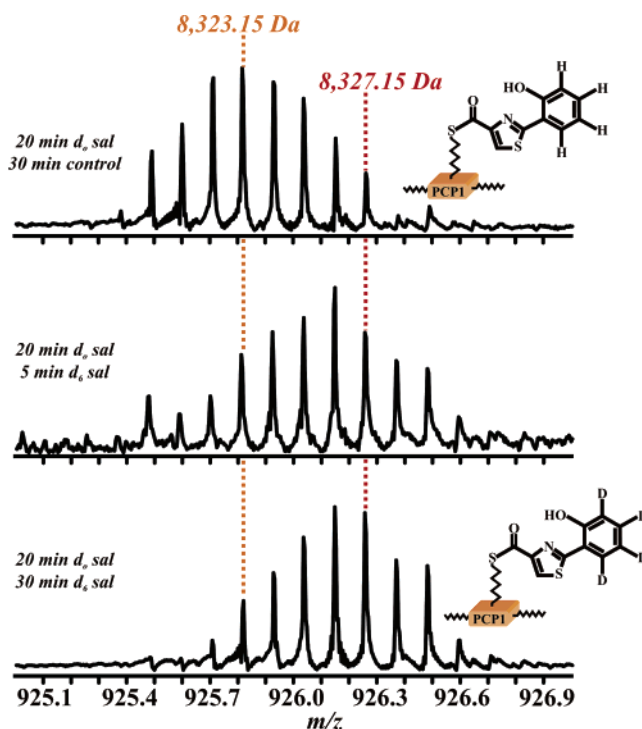

FIGURE 8: Stable isotope pulse into the fully reconstituted pyochelin synthetase. After a 30 min control, the expected isotopic envelope for the exhaustive PCP1 active site peptide was observed. After a 5 min d_6-salicylic acid pulse, four deuteriums were partially incorporated, whereas two deuteriums were lost to the carboxylate and back exchange. After a 30 min d_6-salicylic acid pulse, full incorporation of four deuteriums was observed, indicating complete turnover of the domain.

cleavages. Despite these drawbacks, however, some chemical digestion reagents are compelling for the peptide mapping of NRP/PK synthetases. Cyanogen bromide is particularly

suitable since it operates at low pH and it cleaves proteins into a smaller number of large fragments that are better for semiquantitative determination of multiple intermediates bound to a particular active site. However, even when using chemical digestion agents, the resulting peptide mixture is often too complex for the identification of the target peptides and must therefore be subjected to extensive fractionation. To both expedite and increase the success rate of the peptide identification process, several coenzyme A analogues harboring reporter molecules attached to the terminal thiol were evaluated (Scheme 1).

To localize key peptides within a chromatogram, two chromophoric coenzyme A analogues were evaluated: pyrene-maleimidyl-*S*-CoA (λ_{max} = 338 nm) and BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA (λ_{max} = 504 nm). While both analogues were useful in peptide identification, BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA possesses clear advantages because of its strong fluorescence properties in RPLC solvents. Recent work using BODIPY-FL-*N*-(2-aminoethyl)maleimide has reported low femtomole limits of detection (LOD) when using laser-induced fluorescence (*31*). This decreased LOD will be especially useful when identifying active site peptides from NRP or PK synthetases that are susceptible to ≥95% in vivo phosphopantetheinylation during heterologous expression in *E. coli*. Loading of BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA and pyrene-maleimidyl-*S*-CoA onto the pyochelin biosynthetic enzymes and subjecting the chromatographic fractions possessing the strongest fluorescence and absorbance (at 338 and 504 nm) to mass spectral analysis led to the facile detection of four CNBr peptides containing the covalently modified serines of the ArCP, PCP1, and PCP2 domains (Figures 1−4). Two of the four peptides represented the expected exhaustively digested CNBr products possessing the PCP1 and PCP2 active site serines. The other two peptides, however, contained the ArCP and PCP1 active site serines, but were unexhaustively digested because of the oxidation of Met[50] and Met[1335] to their sulfoxide states.

In addition to the unexhaustively digested species, two moieties of aberrant mass were detected. All pyrene-maleimidyl-*S*-enzyme-loaded peptides exhibited a peak consistent with hydration (18 Da). While it is unclear if this hydration was caused in the synthesis of the reporter or in the subsequent peptide mapping assay, since the absorbance at 338 nm was sufficient to detect the key peptides, it was not investigated further. Additionally, the unexhaustive peptide harboring the PCP1 active site serine displayed a −46 Da species, which has been observed in prior mechanistic studies of cyanogen bromide digestion (*21*). This modification is likely due to a side reaction during digestion, where Thr[1336] is involved in the formation of a second five-membered ring, ultimately permitting ejection of methyl thiocyanate without cleaving the amide bond. Regardless, the successful identification of the modified forms allows us to gain biochemical insights into the active site peptides. The use of coenzyme A analogues coupled with the amazing detection capability of Q-FTMS not only permits localization of these peptides to key chromatographic fractions but also allows for their identification despite the low abundance or alternative covalent states.

A trade-off inherent within these approaches requires the balance of the reporter molecule's size and extinction

Table 1: Mass Shifts of Active Site Peptides Induced by Sfp-Catalyzed Loading of Alternative CoA Substrates

	chromatographic fraction[a]				chromatographic fraction[a]		
	holo	pyrene	BODIPY		holo	pyrene	BODIPY
ArCP	35	36	37	PCP1[c]	28	30	30
PCP1[b]	28	30	30	PCP2	33	35	35

[a] Each fraction was collected at a rate of 1 mL/min for 1 min. [b] Exhaustively digested peptide. [c] Unexhaustively digested peptide.

coefficient. Since a sufficiently large reporter can induce shifting of the target peptide into an undesirable chromatographic fraction, the holo forms of each active site peptide were identified and the RPLC retention times compared. As evidenced in Table 1, the addition of 297.1 Da from pyrene-maleimide and 414.2 Da from BODIPY-FL-*N*-(2-amino-ethyl)maleimide induced a 1−2 min shift in retention on the key peptides, demonstrating that the addition of high-molecular mass reporters can provide the localization of peptides harboring on-pathway intermediates.

To further localize the active site peptides and assess digestion heterogeneity, a mass shift-based assay was employed. Loading PchE and PchF with equimolar mixtures of CoA, acetyl-*S*-CoA, and acetoacetyl-*S*-CoA permitted the detection of the peptides identified using the chromophoric CoA analogues by virtue of 42 and 84 Da mass shifts. The lack of an exact 1:1:1 holo:acetyl:acetoacetyl ratio is likely due to experimental error coupled with unproductive hydrolysis of the acyl-thioester-bound substrates during the loading process. Two other species harboring the ArCP and PCP1 active site serines were also identified. These species were of low abundance and were not observed reproducibly. While this method provides a means of rapidly identifying holo form carrier peptides, only one of the aforementioned methodologies is required for successful active site identification. The combination of multiple strategies, however, can provide valuable insights into chromatographic and proteolytic efficiency and optimization.

Occasionally, poor fractionation during RPLC coupled with low enzyme expression yields (<0.5 mg/L) and the coelution of highly abundant contaminating *E. coli* proteins during the protein purification process can prevent target peptide detection because of signal suppression during ionization. This problem can be partially circumvented, however, by using affinity-tagged CoA analogues. As shown in Figure 6, loading biotin-maleimidyl-*S*-CoA followed by batch purification over avidin resin resulted in the removal of several contaminating proteins from the PchE protein stock. While this approach permits the recovery of enough material for peptide mapping, the low binding capacity and expense of avidin resin make this dimension of purification unreasonable for large quantities. Therefore, while this approach can aid in initial active site mapping attempts, future mass spectrometry-based studies of complex synthetases will entail the use of separations technology online with ESI/FTMS.

Having successfully mapped the three covalently modified active sites from pyochelin synthetase, we reconstituted PchD, holo-PchE, and holo-PchF in a 1:1:1 molar ratio to track complex intermediate formation. Both the salicyl-*S*-ArCP and HPT-*S*-PCP1 complex intermediates were readily identified from the complex mixture of PchD−F. The 100%

occupancy of HPT-*S*-PCP1 indicates that no dead enzyme is detectable. The HPTT-*S*-PCP2 intermediate, however, could not be visualized. Instead, a large amount of holo-PCP2 was identified. The lack of HPTT-*S*-PCP2 could be due to an inactive Cy$_2$ domain resulting in poor transmission efficiency between the proteins, rapid transfer and hydrolysis of the product through the PCP2 and TE domains, signal suppression of the complex intermediate, or a combination thereof. Further optimization of the separation as well as evaluation of monomer loading on PchF will help shed light on these issues. Further, turnover of the PCP1 active site was visualized by employing a d_6-stable isotope pulse after reconstitution. The occupancy of the PCP1 active site exhibited complete incorporation of the labeled salicylic acid within 30 min, presumably due to a combination of non-specific hydrolysis and substrate transfer to PchF (*26−28*). This pulse of isotopes allowed a quick assessment of the kinetics of intermediate movement along the synthetase.

CONCLUSION

Direct interrogation of NRPSs/PKSs using mass spectrometry provides a lucrative route toward shedding light on the substrate specificity, kinetics, and mechanism of NRP/PK biosynthesis. This method is hindered, however, by sample complexity induced by the digestion of large, impure enzymes. In this report, we have demonstrated that reporter CoA analogues can successfully localize key peptides to specific chromatographic fractions, reducing analysis time and providing a means of optimization. The use of the fluorescent BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA, among other analogues, allowed for the identification of the target peptides harboring the active site serines of the ArCP, PCP1, and PCP2 thiolation domains of the pyochelin biosynthetic enzymes (PchE and PchF). These identifications permitted very quick and direct observation of the biosynthetic intermediates involved in the "assembly line" production of the siderophore pyochelin. Extension of this methodology to other NRP/PK synthetases will greatly expedite carrier peptide localization and ultimately lead to the improved understanding of diverse thiotemplate systems that defy bioinformatic predictions of substrate utilization or clean categorization of domain architecture.

ACKNOWLEDGMENT

SUPPORTING INFORMATION AVAILABLE

Tandem mass spectrometry data, synthetic details, pyochelin mechanism, and mass shift assays on PchF. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2002) Ways of assembling natural products on modular nonribosomal peptide synthetases, *ChemBioChem 3*, 491−504.
2. Walsh, C. T. (2004) Polyketide and nonribosomal peptide antibiotics: Modularity and versatility, *Science 303*, 1805−1810.
3. Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis, *Chem. Rev. 97*, 2651−2673.
4. Caffrey, P., Bevitt, D. J., Staunton, J., and Leadlay, 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.
5. Schwecke, T., Aparicio, J. F., Molnár, I., König, A., Khaw, L. E., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortés, J., Lester, J. B., Böhm, G. A., Staunton, J., and Leadlay, P. F. (1995) The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin, *Proc. Natl. Acad. Sci. U.S.A. 92*, 7839−7843.
6. 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.
7. Hubbard, B. K., and Walsh, C. T. (2003) Vancomycin assembly: Nature's way, *Angew. Chem., Int. Ed. 42*, 730−765.
8. 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.
9. Bearden, S. W., Fetherston, J. D., and Perry, R. D. (1997) Genetic organization of the yersiniabactin biosynthetic region and construction of avirulent mutants in *Yersinia pestis, Infect. Immun. 65*, 1659−1668.
10. Miller, D. A., Luo, L., Hillson, N., Keating, T. A., and Walsh, C. T. (2002) Yersiniabactin synthetase: A four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*, *Chem. Biol. 9*, 333−344.
11. Cane, D. E., Walsh, C. T., and Khosla, C. (1998) Harnessing the biosynthetic code: Combinations, permutations and mutations, *Science 282*, 63−68.
12. Doekel, S., and Marahiel, M. A. (2000) Dipeptide formation on engineered hybrid peptide synthetases, *Chem. Biol. 7*, 373−384.
13. Walsh, C. T., Fisher, S. L., Park, I. S., Prahalad, M., and Wu, Z. (1996) Bacterial resistance to vancomycin: Five genes and one missing hydrogen bond tell the story, *Chem. Biol. 3*, 21−28.
14. Deresinski, S. (2005) Methicillin-resistant *Staphylococcus aureus*: An evolutionary, epidemiologic, and therapeutic odyssey, *Clin. Infect. Dis. 40*, 562−573.
15. McLoughlin, S. M., and Kelleher, N. L. (2004) Kinetic and regiospecific interrogation of covalent intermediates in the nonribosomal peptide synthesis of yersiniabactin, *J. Am. Chem. Soc. 126*, 13265−13275.
16. Gatto, G. J., Jr., McLoughlin, S. M., Kelleher, N. L., and Walsh, C. T. (2005) Elucidating the substrate specificity and condensation domain activity of FkbP, the FK520 pipecolate-incorporating enzyme, *Biochemistry 44*, 5993−6002.
17. Hicks, L., Weinreb, P., Konz, D., Marahiel, M. A., Walsh, C. T., and Kelleher, N. L. (2003) Fourier-transform mass spectrometry for detection of thioester-bound intermediates in unfractionated proteolytic mixtures of 80 and 191 kDa portions of bacitracin A synthetase, *Anal. Chim. Acta 496*, 217−224.
18. Hicks, L. M., O'Connor, S. E., Mazur, M. T., Walsh, C. T., and Kelleher, N. L. (2004) Mass spectrometric interrogation of thioester-bound intermediates in the initial stages of epothilone biosynthesis, *Chem. Biol. 11*, 327−335.
19. Mazur, M. T., Walsh, C. T., and Kelleher, N. L. (2003) Site-specific observation of acyl intermediate processing in thiotemplate biosynthesis by Fourier transform mass spectrometry: The polyketide module of yersiniabactin synthetase, *Biochemistry 42*, 13393−13400.
20. Zhang, X., Dillen, L., Vanhoutte, K., Dongen, W. V., Esmans, E., and Claeys, M. (1996) Characterization of unstable intermediates and oxidized products formed during cyanogen bromide cleavage of peptide and proteins by electrospray mass spectrometry, *Anal. Chem. 68*, 3422−3430.
21. Kaiser, R., and Metzka, L. (1999) Enhancement of cyanogen bromide cleavage yields for methionyl-serine and methionyl-threonine peptide bonds, *Anal. Biochem. 266*, 1−8.
22. La Clair, J. J., Foley, T. L., Schegg, T. R., Regan, C. M., and Burkhart, M. D. (2004) Manipulation of carrier proteins in antibiotic biosynthesis, *Chem. Biol. 11*, 195−201.
23. Quadri, L. E. N., 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.

24. Yin, J., Liu, F., Schinke, M., Daly, C., and Walsh, C. T. (2004) Phagemid encoded small molecules for high throughput screening of chemical libraries, *J. Am. Chem. Soc. 126*, 13570−13571.

25. Yin, J., Liu, F., Li, X. H., and Walsh, C. T. (2004) Labeling proteins with small molecules by site-specific posttranslational modification, *J. Am. Chem. Soc. 126*, 7754−7755.

26. Quadri, L. E. N., Keating, T. A., Patel, H. M., and Walsh, C. T. (1999) Assembly of the *Pseudomonas aeruginosa* nonribosomal peptide siderophore pyochelin: In vitro reconstitution of aryl-4,2-bisthiazoline synthetase activity from PchD, PchE and PchF, *Biochemistry 38*, 14941−14954.

27. Reimmann, C., Patel, H. M., Serino, L., Barone, M., Walsh, C. T., and Haas, D. (2001) Essential PchG-dependent reduction in pyochelin biosynthesis of *Pseudomonas aeruginosa*, *J. Bacteriol. 183*, 813−820.

28. Patel, H. M., and Walsh, C. T. (2001) In vitro reconstitution of the *Pseudomonas aeruginosa* nonribosomal peptide synthesis of pyochelin: Characterization of backbone tailoring thiazoline reductase and *N*-methyltransferase activities, *Biochemistry 40*, 9023−9031.

29. Senko, M. W., Hendrickson, C. L., Emmett, M. R., Shi, S. D. H., and Marshall, A. G. (1997) External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *J. Am. Soc. Mass Spectrom. 8*, 970−976.

30. Horn, D. M., Zubarev, R. A., and McLafferty, F. W. (2000) Automated reduction and interpretation of high-resolution electrospray mass spectra of large molecules, *J. Am. Soc. Mass Spectrom. 11*, 320−332.

31. Mao, Y., and Zhang, X. (2003) Comprehensive two-dimensional separation system by coupling capillary reverse-phase liquid chromatography to capillary isoelectronic focusing for peptide and protein mapping with laser-induced fluorescence detection, *Electrophoresis 24*, 3289−3295.