Published in final edited form as:
J Am Chem Soc. 2004 October 20; 126(41): 13265–13275.

Kinetic and Regiospecific Interrogation of Covalent Intermediates in the Nonribosomal Peptide Synthesis of Yersiniabactin

Shaun M. McLoughlin and **Neil L. Kelleher***
Contribution from the Department of Chemistry, University of Illinois, Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

Abstract

For interrogation of enzyme-bound intermediates in nonribosomal peptide synthetases (NRPSs), mass spectrometry is used to read out the kinetics and substrate specificity of this medicinally important class of enzymes. The protein HMWP2 (230 kDa) catalyzes 11 chemical reactions, four of which could be resolved by fast quench approaches combined with mass spectrometry. The rate of complex intermediate accumulation at the PCP1 active site was observed to occur with a rate of 19 s^{-1}, with the rate of cysteine acylation faster than that of intermediate translocation. Use of alternative substrates for salicylic acid (at the ArCP carrier domain) and L-cysteine (at the PCP1 carrier domain) revealed a high penalty for omission of the salicyl alcohol. For some substrates, large discrepancies were found between prior adenylation assays and the current MS-based readouts. Indirect evidence for condensation via a thiolate attack (vs an amino group) was also accumulated. This is the first report to correlate the percent occupancy of multiple active sites in parallel with kinetic and structural resolution of intermediates and provides new evidence of interdomain and intermodule communication within thiotemplate assembly lines.

Introduction

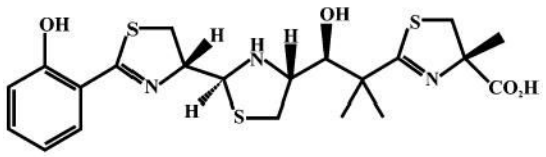
A myriad of natural products are assembled on multienzyme scaffolds using a “thiotemplate” biosynthetic motif.[1–5] These methods of bioassembly are known as nonribosomal peptide synthesis (NRPS) and polyketide synthesis (PKS) and are responsible for the genesis of several structurally heterogeneous, medicinally important compounds.[6–9] For instance, rifamycin, whose precursor is synthesized by the bacterium *Amycolatopsis mediterranei*, is a macrocyclized NRP/PK hybrid product assembled from 11 constituents, and it acts to inhibit bacterial transcription by binding to the β subunit of RNA polymerase.[10,11] This antibiotic is commonly used to combat tuberculosis. Another NRPS product generated by *Amycolatopsis mediterranei*, balhimycin, employs the use of the nonproteinogenic amino acid dihydroxyphenylglycine (assembled using type 3 PKS logic) to create a cross-linked, cup-shaped ring designed to target cell wall receptors of gram positive bacteria, thus inhibiting their growth.[12,13] This NRPS product belongs to a group of vancomycin-type antibiotics, which are commonly referred to as antibiotics of “last resort” due to their effectiveness in preventing the proliferation of many forms of infectious bacteria that have become resistant to other antibiotic treatments, particularly *Staphylococcus aureus*.[14]

In addition to antibiotics, thiotemplate enzymes are commonly employed by various forms of bacteria to create small molecules that serve as virulence factors. A common class of these compounds are siderophores, which are iron-chelating compounds that scavenge ferric cations

E-mail:kelleher@scs.uiuc.edu.

Supporting Information Available: Tables denoting percent occupancies with various substrates withheld, kinetic percent occupancy data, and MS/MS data for the two active site peptides. This material is available free of charge via the Internet at http://pubs.acs.org.

from the host organism.[15–17] An example of a siderophore is yersiniabactin, synthesized by the organism *Yersinia pestis*, the causative bacterium of bubonic plague.[18,19] Yersiniabactin, a heterocyclized moiety generated from salicylic acid, three cysteines cyclized to their thiazoline and thiazolidine forms, malonate, and 3 equivalents of S-adenosylmethionine, possesses six putative chelation sites for the coordination of ferric cations, including the nitrogen bound in the thiazoline and thiazolidine rings as well as the two alcohols and the carboxylate (**1**).[18] Yersiniabactin is synthesized on a NRPS scaffold consisting of four enzymes: YbtE, a 60 kDa external AMP ligase, HMWP2, a 230 kDa multifunctional enzyme responsible for the formation of the hydroxyphenyl-thiazolinyl-thiazolinyl-S-enzyme moiety (HPTT-S-enzyme), HMWP1, another multifunctional enzyme (350 kDa) responsible for generating yersiniabactin from HPTT, and YbtU, a 30 kDa thiazolinyl reductase (Scheme 1). The multifunctional enzymes HMWP1 and HMWP2 possess several domains that are either posttranslationally modified with phosphopantetheinyl cofactors (e.g., aryl, acyl, and peptidyl carrier domains) or contain conserved active site nucleophiles (e.g., acyl transferase, thioesterase, and keto-acyl synthase domains), which serve to tether the constituents and elongating product to the enzyme.[20–23] Because each successive step in the synthesis of yersiniabactin is driven by covalent catalysis, the mass changes induced by acylation, condensation, cyclodehydration, and tailoring of the elongating product can be observed independently using mass spectrometry.[24–26]



(1)

Electrospray ionization (ESI) combined with Quadrupole/Fourier-Transform mass spectrometry (ESI-Q-FTMS) is well suited for interrogation of mass changes to thiolation domains because its sensitivity permits detection of low abundance intermediates and it can resolve covalent changes to enzyme-bound intermediates that are <1 Da at ~30 kDa.[27] However, while ESI-Q-FTMS is an optimal technique to analyze complex proteolytic mixtures, it is ill-suited to analyze proteins above 120 kDa with isotopic resolution.[28] As a result, the enzymes must first be digested into smaller peptides after in vitro reconstitution.[24,25] By monitoring the covalent states of the peptides harboring thioester intermediates, mechanistic insight into the biogenesis of yersiniabactin becomes possible.

Established biochemical assays, including $^{32}PP_i$ exchange to assess adenylation and radio-HPLC to detect radioactive product formation, have been previously used to study HMWP2.[21,29] While rates of acylation and product formation were successfully calculated and unnatural substrates accepted by HMWP2 were identified, the studies were limited by the specific radioactivity of the substrates used, the lack of synthesized radioactive unnatural substrates/products, and a lack of regiospecific analyses because the intermediates are hydrolyzed off of the enzyme.[21,29] Additionally, the $^{32}PP_i$ exchange assay was originally designed to probe the activity of amino acid-tRNA synthetases and assesses adenylation activity, not covalent loading in NRPS/PKS systems directly.[29] In this study, ESI-Q-FTMS was used to interrogate the timing and substrate specificity of the initiating reactions in yersiniabactin biosynthesis. Several acceptable unnatural substrates were identified (including some that had been previously thought to be inactive), and kinetic insights into the formation of the complex intermediate hydroxyphenyl-thiazolinyl-S-PCP1 (HPT-S-PCP1) were elucidated with a temporal resolution down to 10 ms, although intermediate formation was not

observed until 250 ms. This presents the first study that correlates the percent occupancy of multiple active sites in parallel with the kinetic and structural resolution of NRPS intermediates.

Experimental Procedures

Materials

Plasmids for HMWP2, Sfp, and YbtE were obtained from the Walsh laboratory at the Harvard Medical School (Boston, MA). Competent BL21(DE3) *E. coli* was purchased from Invitrogen (Carlsbad, CA). Competent DH5α *E. coli* and ampicillin-treated LB agar plates were purchased from the cell and media facility at the University of Illinois (Urbana, IL). Isopropyl β-D-thiogalactopyranoside (IPTG) and sequencing grade trypsin were purchased from Promega (Madison, WI). Nickel affinity chromatograpy resin and accessories were purchased from Qiagen (Valencia, CA). DEAE media was purchased from Amersham Biosciences (Piscataway, NJ). Bradford protein assay was purchased from Biorad (Hercules, CA). Centriplus protein concentrators were purchased from Millipore (Billerica, MA). Electrospray emitter tips and assembly were purchased from New Objective (Woburn, MA). Nanospray emitter nozzles and sample tips were purchased from Advion Biosciences (Ithaca, NY). All other chemicals were purchased from Sigma-Aldrich (Milwaukee, WI).

Overexpression and Purification of HMWP2 and YbtE

Overexpression and purification of HMWP2 and YbtE were performed as described previously.[30,31] A pET16b plasmid bearing a N-terminally decahistidine tagged HMWP2 construct and a pPROEX1 plasmid bearing a N-terminally hexahistidine tagged YbtE construct were both transformed into the BL21(DE3) expression strain of *E. coli* and plated on LB agar with 100 μg/mL ampicillin. Cultures of the HMWP2 *E. coli* construct were grown (4 × 4 L) at 30 °C with shaking (250 rpm) in 2xYT media with 2 mM $MgCl_2$, and the YbtE *E. coli* construct was grown (1 × 1.5 L) at 37 °C with shaking (250 rpm) in LB media. The OD_{600} was measured periodically, and when it reached 0.4, the temperature of the HMWP2 growth was lowered to 18 °C. When the OD_{600} reached 0.8, the cultures from both constructs were induced with IPTG (50 μM for HMWP2, 1 mM for YbtE). Growth was allowed to continue (12 h for HMWP2, 2 h for YbtE) before harvesting by centrifugation (10 min at 9000*g*). The cell pellets from the YbtE and HMWP2 growths were each independently resuspended in 25 mM Tris–HCl pH 8/500 mM NaCl, and then lysed separately by two passages through a French press. The lysates were clarified by ultracentrifugation (30 min at 99 000*g*). The supernatant from both cultures was stirred in parallel on Qiagen nickel affinity resin for 2 h before being packed into gravity columns. After extensive washing in lysis buffer containing 5 mM imidazole, the protein was eluted in an imidazole gradient from 10 to 500 mM lysis buffer + imidazole over 17–21 mL. The fractions from both purifications were analyzed by SDS-PAGE, and those containing HMWP2 were pooled and dialyzed against 25 mM Tris–HCl pH 8/0.5 mM TCEP, and those containing YbtE were pooled and dialyzed against 25 mM Tris–HCl pH 8/0.5 mM TCEP/10 mM $MgCl_2$/10% glycerol. The HMWP2 was further purified by loading onto a 6 × 3 cm DEAE anion exchange column and eluting in a linear gradient in lysis buffer containing 0–500 mM NaCl over 110 mL at 1 mL/min. The fractions were analyzed by SDS-PAGE (8% T gel), and those containing HMWP2 were pooled and dialyzed against 25 mM Tris–HCl pH 8/0.5 mM TCEP/10% glycerol. Both YbtE and HMWP2 were concentrated separately using Centriplus concentrators before Bradford analysis using BioRad protein assay (IgG being the standard). The YbtE and HMWP2 solutions were divided into 50 μL aliquots and stored at −80 °C.

In Vitro Reconstitution of HMWP2 for Alternative Substrate Assays

Fifty μL aliquots of HMWP2 (13.2 μM) were combined with YbtE (13.2 μM), Sfp (1.6 μM), and coenzyme A (362 μM), as well as the desired substrates for reconstitution (1–2 mM) and were diluted to 90 μL in stock reaction buffer to the final concentrations parenthetically

indicated (the stock reaction buffer being 100 mM Tris–HCl pH 8/10 mM $MgCl_2$/1 mM TCEP). The substrates used to assay active site loading efficiency and specificity include salicylic acid, d_6-salicylic acid, 2,3-dihydroxybenzoic acid, 2,4-dihydroxybenzoic acid, 2,6-dihydroxybenzoic acid, benzoic acid, 4-methylsalicylic acid, *p*-toluic acid, 4-hydroxybenzoic acid, L-phenylalanine, L-tyrosine, 3-hydroxypicolinic acid, 2-hydroxynicotinic acid, anthranilic acid, 2-thiosalicylic acid, L-cysteine, L-serine, L-homocysteine, 4-vinyl-L-glycine, 2,3-diaminopropionic acid, and α-amino-L-butyric acid. The reactions were incubated at 30 °C for 4 h to prime the active sites with phosphopantetheinyl cofactors. ATP (10 mM) was subsequently added to a final reaction volume of 100 μL to initiate intermediate formation. After further incubation at 30 °C for 5–30 min, the reactions were quenched either by addition of formic acid or by addition of trypsin followed by formic acid. In both cases, formic acid was added to a final concentration of 30% and did not precipitate the HMWP2.

Kinetic Analysis of the PCP1 Active Site

For 2–20 s time points, 50 μL aliquots of HMWP2 (13.2 μM) were combined with YbtE (13.2 μM), Sfp (1.6 μM), coenzyme A (362 μM), salicylic acid (1 mM), and ATP (5 mM) and were diluted to 95 μL in stock reaction buffer to the parenthetically indicated final concentrations. The reactions were incubated at 30 °C for 4 h to both prime the active sites with their phosphopantetheinyl cofactors and load the ArCP active site with salicylic acid. Five μL of 100 mM L-cysteine was added while mixing to initiate the reaction. The reactions were quenched after the desired time by adding 60 μL of 99% formic acid. For subsecond time points down to 10 ms, 100 μL of the salicyl-loaded and primed HMWP2 was mixed with 100 μL of 10 mM L-cysteine through a Kintek Chemical Fast Quench instrument (model RQF-3), using approximately 100 μL of 99% formic acid to quench the reactions.

Digestion of HMWP2

Cyanogen bromide digestion of HMWP2 was accomplished by addition of 5 M CNBr in CH_3CN to the quenched loading reactions at 4 °C (1 M final concentration), incubation in the dark for 12–18 h, and lyophilization. Trypsinization was carried out by adding trypsin to the 100 μL HMWP2 reactions in a 1:5 wt/wt ratio of protease to substrate. After proteolysis at 30 °C for 5 min, the digestions were quenched as described previously and lyophilized. The lyophilized pellets from both reactions were resuspended in 100 mM NH_4OAc pH 4/6 M urea/5 mM TCEP/10% CH_3CN for 1 h at room temperature prior to being injected onto a Jupiter C4 reverse-phase column (4.6 × 150 mm). The digestions were fractionated in a linear gradient of 30–70% CH_3CN in 0.1% TFA over 40 min at 1 mL/min. One-minute fractions were collected, lyophilized, and stored at −20 °C before MS analysis.

ESI-Q-FTMS

The lyophilized fractions were resuspended in 20–40 μL of electrospray solvent (49% CH_3CN, 50% water, 1% formic acid). The peptides were directly infused into a custom-built quadrupole-FTMS hybrid instrument through a microelectrospray assembly terminated with a 50 μm inner diameter ESI tip or through nanoelectrospray nozzles using an Advion nanospray robot. The resulting ions were passed through a resistively heated metal capillary and skimmer before being externally accumulated for a total of 0.5 s in an accumulation octapole.[32] After accumulation, the ions were shuttled to the ICR cell through a quadrupole that can function either as a simple ion guide or as a selective filter. The values reported in this study are monoisotopic molecular weights (M_r) unless otherwise indicated. External calibration of the spectra was based on bovine ubiquitin (8559.64 Da). Increased mass accuracy was achieved for selected spectra using 2 to 3 peaks of known identity as internal calibrants.

FTMS Data Treatment

The isotopic distributions of the active site peptides were fit using a least squares algorithm provided by the FTMS data analysis program THRASH.[33] Upon comparing the least-squares fit to a theoretical fit of the distribution, the *m*/*z* values of both the monoisotopic and the most abundant isotopic peaks of the peptides were determined. The degree of occupancy of a substrate on any active site peptide was determined by comparing the integrated intensity of the three most abundant isotopes of the substrate-loaded versus holo enzyme forms. The accuracy of percent occupancy values between spectra of active sites bearing different intermediates is semiquantitative due to possible differences in ionization efficiencies of the differentially modified peptide ions, but have been shown to be precise to ~5%.[26]

Results

Detection of HMWP2 Active Site Peptides

Cyanogen bromide digestion of HMWP2 followed by reverse-phase fractionation resulted in the detection of over 150 distinct peptides in 21 fractions with 11 matching to the DNA-predicted sequence of HMWP2. This peptide map provided only 23% sequence coverage within 25 ppm primarily due to the low purity and expression yields of HMWP2 and reduced chromatographic resolution late in the HPLC run (Figure 1A). Despite the low sequence coverage, two peptides with M_r values of 6186.19 and 20 704.5 Da in fractions 27 and 34 corresponded to the theoretical CNBr peptides containing the ArCP (Ile33–Met^{87}) and PCP1 (Val^{1396}–Met^{1580}; Figure 1C) active sites within 5 and 15 ppm, respectively. The identity of the PCP1 "carrier" peptide was unambiguously confirmed using tandem mass spectrometry (Supporting Information Figure 1). Due to difficulties in tracking covalent modifications to the ArCP active site peptide produced by CNBr, it is not discussed extensively here.

Incubation of recombinant HMWP2 in trypsin resulted in extensive cleavage of the protein in 5 min. Reverse-phase fractionation (not shown) of the peptides allowed detection of over 100 matching peptides within 25 ppm, yielding 60% sequence coverage of the DNA-predicted sequence over 13 fractions. Two peptides, eluting in fractions 19 and 26, had masses that were consistent with the apo forms of the tryptic peptides containing the PCP1 active site (Glu^{1460}–Arg^{1475}) and the PCP2 active site (Glu^{1989}–Arg^{2020}) with mass values of 1783.85 Da (7 ppm) and 3466.72 Da (26 ppm), respectively (data not shown). Further, a peptide of mass 3071.63 Da eluting in fraction 22 was consistent within 2 ppm with the peptide harboring the ArCP carrier site (Leu^{59}–Arg^{85}) (Figure 1B). The identity of the ArCP active site peptide was confirmed using tandem mass spectrometry (Supporting Information Figure 2). All observed active site peptides in both digestion protocols were primarily in the apo-form, with ~5–10% holo-forms visible, presumably from in vivo addition of the phosphopantetheinyl cofactor (+340 Da) by previously identified *E. coli* enzymes (EntD or ACPS) during HMWP2 overexpression.[34]

Priming the apo-form active sites in vitro by treatment of HMWP2 with Sfp and CoA prior to digestion and fractionation resulted in a 340.1 Da increase in the ArCP and PCP1 peptides (Figure 2A vs 2C and 2B vs 2D, respectively). These data indicate quantitative conversion of apo- to holo-HMWP2, although the holo-PCP2 carrier peptide was not detected despite repeated attempts. Incubation of holo-HMWP2 with YbtE, salicylic acid, and ATP followed by digestion, RPLC fractionation, and MS analysis gave the Figure 2E data, indicating quantitative formation of the salicyl-S-ArCP intermediate. Similar treatment with the addition of L-cysteine resulted in all of the PCP1 peptide occupancy shifting +203.0 and +223.0 Da, indicating a combined 100% occupancy of the condensation products Sal-Cys-S-PCP1 (ring opened) and HPT-S-PCP1 (Figure 2F). The Sal-Cys-S-PCP1 intermediate overlapped with a sodium adduct (+20 versus +22 Da from the HPT-S-PCP1 intermediate), making its

identification and quantitation difficult. The spectrum of the HPT-S-PCP1 intermediate was 2.1 Da underweight, consistent with the oxidation of the thiazoline ring to the thiazole.[10] Replacing d_0-salicylic acid with d_6-salicylic acid caused a 4 Da increase in mass on both the ArCP and the PCP1 active site peptides (Figure 2G and 2H), showing the incorporation of 4 deuteriums into the product. Two of the deuteriums in d_6-salicylic acid are on heteroatoms in phenolic and carboxylic acid positions and were lost to the reaction and proton back exchange, respectively.

Acylation and Condensation Efficiency of Noncognate Benzoates

As a probe for the aminoacylation selectivity of YbtE and thioester formation at the ArCP active site, several substituted benzoates were assayed for loading efficiency. With the salicyl alcohol (2-OH) involved in ferric ion binding, substrates were chosen to alter this position. Incubation of 2-thiosalicylic acid with holo-HMWP2, YbtE, L-cysteine, and ATP resulted in the formation of a molecular weight corresponding to the 2-thiosalicyl-S-ArCP intermediate (3547.65 Da) at ~25% occupancy (Figure 3A). The identity of this peak was verified using tandem mass spectrometry (data not shown). Similar incubation with anthranilic acid showed no formation of the loaded product (Table 1). Taken together, these results show that the salicyl alcohol plays a key role in efficient recognition of the substrate by the enzyme. As evidenced in Table 1, several unnatural substrates successfully loaded, showing a general lack of YbtE and ArCP domain specificity. However, complete removal of the 2-OH, using benzoic acid as the substrate, yielded a product peak of mass 3515.81 Da, consistent with the benzoyl-S-ArCP intermediate (Table 1). The observed occupancy was ~80%, suggesting that while the salicyl alcohol may contribute to substrate recognition, it is not absolutely necessary. Neither L-phenylalanine nor L-tyrosine successfully loaded onto the ArCP domain above the ~0.1% threshold for FTMS detection, consistent with evolution of a strong bias against more abundant potential substrates present in vivo.

To test the limits of substrate recognition, further structural analogues of salicylic acid were examined and their loading occupancies were compared. Of note, 2,4-dihydroxybenzoic acid and 4-hydroxybenzoic acid were incubated in separate reactions for 30 min and produced ArCP-loaded occupancies of ~99% (Figure 3C) and ~20%, respectively (Table 1). Thus, removal of the 2-OH of 2,4-dihydroxybenzoic acid resulted in a drastic reduction in loading efficiency, further displaying a role of the 2-OH for enzyme–substrate recognition. The same trend was observed with the structural analogues 4-methylsalicylic acid and *p*-toluic acid. Identical incubations of these substrates with holo-HMWP2 resulted in the formation of the ArCP occupancies of ~99% and ~35%, respectively (Figure 3E and 3G).

Downstream Processing of Alternative Substrates

Several competent ArCP substrates, including 2-thiosalicylic acid, 2,4-dihydroxybenzoic acid, 4-methylsalicylic acid, and *p*-toluic acid, were screened for condensation efficiency with a loaded downstream L-cysteine on PCP1 and PCP2. For 2,4-dihydroxy-benzoic acid and 4-methylsalicylic acid, molecular weights were observed that were consistent with the formation of the 2,4-dihydroxyphenylthiazolinyl-S-PCP1 (2,4-DHPT-S-PCP1) (21 265.5 Da) and the 2-hydroxy-4-methylphenylthiazolinyl-S-PCP1 (2-H-4-MPT-S-PCP1) (21 263.5 Da) cyclized product analogues, at ~90% and ~85% occupancy, respectively (Figure 3D and 3F). Treatment of L-cysteine-loaded HMWP2 with *p*-toluic acid resulted in the formation of a product distribution corresponding to the formation of the cyclized *p*-methylphenylthiazolinyl-S-PCP1 product (p-MPT-S-PCP1) (21 247.4 Da), although at only ~35% occupancy (Figure 3H). In contrast to the other substrates, incubation of 2-thiosalicylic acid with L-cysteine-loaded HMWP2 resulted in the accumulation to ~20% occupancy of the 2-thiophenylthiazole-S-PCP1 (2-TPT-S-PCP1) product analogue (21 231.4 Da) that is 34.3 Da under the mass of the theoretical product (thiazoline form; Figure 3B). The mass discrepancy is explained by a CNBr-

induced elimination of the aromatic thiol and oxidation of the thiazoline ring to the thiazole form. The desulfurization of thiols has also been observed on other sulfur-containing intermediates (vide infra). Overall, when loading for 30 min, if the substrate occupancy at the ArCP active site is observed to be poor (e.g., <50% occupancy), the condensation yield is also poor (below 25–35% occupancy). However, it is evident that HMWP2 not only will tolerate the adenylation and acylation of noncognate benzoates, but also their downstream processing.

Dissecting Acylation versus Condensation at the PCP1 Active Site

To examine the specificity of acylation, condensation, and cyclodehydration activities, several alternative substrates and the native substrate, L-cysteine, were examined. Incubation of holo-HMWP2 with Sfp, CoA, YbtE, ATP, 2 mM salicylic acid, and 2 mM L-cysteine resulted in the formation of the HPT-S-PCP1 and Sal-Cys-S-PCP1 intermediates at ~95% total occupancy after 20 min (Table 2, Figure 4D). Similar incubation with the substrate analogue L-homocysteine (which introduces an additional methylene in the side chain of L-cysteine) showed successful acylation to ~40% occupancy, but no condensation with the upstream salicyl-S-ArCP donor (Table 2). Additionally, the loaded homocysteine-S-PCP1 intermediate was shown to be 34.2 Da underweight, a situation similar to the TPT-S-PCP1 intermediate in Figure 3B. Incubations with α-amino-L-butyric acid and L-serine, which have been tested previously using $^{32}PP_i$ exchange and radio-HPLC assays, showed good acylation occupancies (~99% and ~95%; Table 2 and Figure 4A and 4B, respectively) but no evidence of condensation.[29] 4-Vinyl-L-glycine, a substrate previously shown to be inactive by $^{32}PP_i$ exchange, was also shown to successfully acylate to ~60% but not to condense with the upstream donor (Table 2). These combined data show that the specificity of the PCP1 active site and the HMWP2 cysteinyl adenylation domain for acylation of substrates is relatively relaxed. However, the specificity of the Cy_1 domain to condense the upstream salicyl-S-ArCP with an acyl-S-PCP1 intermediate is highly regulated.

Incubation of HMWP2 with 2,3-diaminopropionic acid, another substrate thought to be inactive by $^{32}PP_i$ exchange assay, resulted in successful acylation of the substrate to ~45% occupancy (Table 2). Additionally, the formation of a 21 251.3 Da species, corresponding to the condensed form of the product that has not been cyclodehydrated, at ~5% occupancy was also observed (Table 2, Figure 5H). Further incubation for 2 h resulted in accumulation of the complex intermediate to ~30% (Figure 4C). The successful condensation of this substrate as opposed to other isosteric substrates to L-cysteine such as L-serine (Figure 4B), or α-amino-L-butyric acid (Figure 4A), raises questions regarding the role of the side-chain nucleophile in the mechanism of condensation catalyzed by the Cy_1 domain (see Discussion).

Substrate Acylation at the ArCP and PCP1 Domains as a Function of the Upstream and Downstream States

Having established the relative selectivity of YbtE and the ArCP active site, the influence of the covalent state of the PCP1 domain was investigated. The loading efficiencies of four substrates, salicylic acid, 4-methylsalicylic acid, 2,4-dihydroxybenzoic acid, and 2,3-dihydroxybenzoic acid, were reexamined on the basis of the loading of L-cysteine onto the PCP1 and PCP2 active sites. In each case, the deviation in ArCP loading was ≤5%, indicating little to no penalty in the acylation of the ArCP domain based on the covalent state of the downstream PCP1 active site (Supporting Information Table 1). To investigate the influence of the ArCP domain on the PCP1 active site, each substrate from the previous section was subjected to identical loading experiments both in the presence and in the absence of salicylic acid (Supporting Information Table 2). Incubation of L-cysteine for 10 min at 1 mM showed the formation of a product peak consistent with the cysteine-S-PCP1 intermediate present at ~55% occupancy (Figure 5A). The occupancy in this incubation contrasts sharply with the occupancy observed with bound salicyl-S-ArCP, which results in the formation of the

condensation products (Sal-Cys-S-PCP1 and HPT-S-PCP1) at ~80% combined occupancy (Figure 5A vs 5B). These data indicate that upstream binding of a suitable donor aids considerably in successful acylation of L-cysteine at the PCP1 active site. Further reactions with the other substrate analogues, however, produced mixed results. L-Homocysteine, for instance, displayed a ~15% better acylation without an upstream bound salicylic acid than it did with one (Figure 5C vs 5D). Of the substrates examined, only L-serine displayed a similar dependence on the upstream salicylic acid acylation event, binding ~20% better with salicyl-S-ArCP than without it (Supporting Information Table 2). In contrast, the occupancies of 4-vinyl-L-glycine, 2,3-diaminopropionic acid, and α-amino-L-butyric acid were virtually unchanged (Figure 5E vs 5F, 5G vs 5H and Supporting Information Table 2).

Kinetics of Intermediate Formation at PCP1

Kinetic insight into the acylation and condensation events at PCP1 on the pathway of yersiniabactin bioassembly was possible by combining fast quench experiments with MS analysis of the PCP1 active site. A time course displaying the kinetics of intermediate formation at PCP1 was generated using a combination of subsecond and subminute time points (Figure 6). By logarithmically fitting the combined percent occupancy of the Sal-Cys-S-PCP1 and HPT-S-PCP1 intermediates, it was shown that 50% of the PCP1 occupancy consisted of complex intermediates after ~6 s, with an initial condensation velocity of ~19 s^{-1} (5 mM cysteine as the initiating substrate; Figure 6, inset).

In addition to tracking complex intermediate formation, other intermediates were also identified. The cysteine-S-PCP1 intermediate accumulated to ~45% occupancy early within the time course, but gradually declined at later time points to ~20%. The intermediate was also oxidized (+16 Da), which is likely an artifact of analysis (see Discussion). The detection of the cysteine-S-PCP1 intermediate coupled with the rapid formation of condensation products in later time points indicates that the apparent rate of cysteine acylation is faster but comparable to the rate of intermediate condensation at the PCP1 active site.

Discussion

To bolster efforts for reengineering NRP/PK enzymes to produce new “unnatural” natural products, traditional biochemical assays can now be coupled with mass spectrometry to assess the active site microheterogeneity with temporal resolution in the low millisecond regime. To obtain this information using MS, however, the large multifunctional enzymes, which are the hallmark of NRP/PK scaffolds, must first be digested into peptides whose size lies within the tractable range of current MS technology.[24–28,35] The limited digestion of these enzymes has been employed previously by a range of studies, allowing for the extraction of whole domains.[26,35,37] In this study, we employ the use of excess cyanogen bromide and trypsin to extensively digest the enzyme with negligible hydrolysis of the thioester-bound constituents.

Active Site Peptide Mapping

The most challenging and rate-limiting step in the MS-based platform of analysis is the successful mapping of the active site peptides. Impurities associated with protein purification compounded by complex proteolytic digests and unexpected byproducts of chemical digestion often result in poor reverse-phase separations, complex fractions, and, ultimately, uninformative MS spectra. Through use of cyanogen bromide and rapid trypsin digestions, all three active sites in HMWP2 were successfully identified in the apo form. In addition to identifying the exhaustively cleaved tryptic ArCP active site peptide, one unexhaustive product harboring the active site was also identified, representing an N-terminal missed cleavage event between arginine and leucine (3466.72 Da, apo). Similarly, a CNBr peptide correlating to one missed C-terminal cleavage between methionine and proline due to unproductive oxidation of

the thioether side chain was also identified (28 160.3 Da, holo). In both cases, the unexhaustively cleaved peptide represented a minor component in the fraction. For the PCP2 tryptic peptide (3466.72 Da, apo), no missed cleavages were detected and subsequent attempts to visualize the holo- and thioester-loaded forms were unsuccessful for reasons yet unknown. It is likely that the absence of the thioester forms of this small peptide is caused by a shifting of their elution times into complex fractions, where they are not easily visible. For determining ratios of differentially modified forms of the active site peptides, factors such as ionization efficiency, ion charge state, and signal suppression from other species present within the spectrum can affect the accuracy of the measurement. However, the percent occupancies are reproducible within the observed experimental variation, which is ~5%,[26] with larger systematic errors possible due to ESI phenomena, fraction pooling, or protein forms digesting differently (e.g., incomplete digestion).

ArCP Domain Alternative Substrate Incorporation

Upon testing the promiscuity of the salicylate loading reactions with a collection of alternative substrates, three major features regarding the recognition, adenylation, and acylation of substrates were revealed. Modification of the salicyl alcohol using 2-thiosalicylic acid, benzoic acid, and anthranilic acid resulted in loading penalties to give 0–80% occupancy at 30 min (vs ~99% for salicylic acid). Because the relative intensities of the modified forms of the active site peptide are dependent on concentration, the subsequent ratios are dependent on the efficiency and reversibility of the enzyme. As a result, these penalties in loading efficiency when compared to salicylic acid could be the cause of an increase in K_m of the substrate for adenylation or acylation, a decrease in the overall acylation of the enzyme (k_{cat}), or both. These penalties point to an editing function of the enzyme that affects overall efficiency (k_{cat}/K_m), where substrates that either do not possess the alcohol or are otherwise substituted within that position experience a significant kinetic bias at initiation and condensation stages of assembly.

The second major feature from the MS-based assays was the acceptance of alternative substrates that possess the salicyl alcohol, but are also substituted in other places on the aromatic ring. Most notably, the substrate analogues 2,4-dihydroxybenzoic acid and salicylic acid were shown to have comparable loading efficiencies, while 4-hydroxybenzoic acid accrued to only ~20% occupancy. Similar results were obtained with the substrate analogues 4-methylsalicylic acid and *p*-toluic acid. While these results clearly demonstrate the importance of the salicyl alcohol in the enzymatic recognition of salicylic acid by YbtE and the ArCP domain, they also display a promiscuity of the editing function. YbtE and the ArCP domain accept several unnatural substrates, but favor those that possess the 2-OH, regardless of additional aromatic substitution.

The third major insight pertains to the translocation efficiency of unnatural substrates incorporated at the ArCP active site to the L-cysteine-loaded PCP1 active site. Several competent ArCP substrates, including 2-thiosalicylic acid, 2,4-dihydroxybenzoic acid, 4-methylsalicylic acid, and *p*-toluic acid, were screened for condensation efficiency, and all were found to condense and cyclodehydrate to create their respective product analogues. The formation of these unnatural condensed products demonstrates that noncognate substrates are not only accepted for acylation at the ArCP domain, but they can also be processed to downstream active sites.

PCP1 Active Site Alternative Substrate Incorporation and Mechanism of Condensation

After screening a set of six substrate analogues to test for the promiscuity of adenylation and acylation activities of HMWP2 (Table 2), the A_{cys} domain successfully loaded the PCP1 domain with many unnatural substrates from ~40% to ~99% occupancy as compared to ~95% for L-cysteine. For 4-vinyl-L-glycine, L-serine, and 2,3-diaminopropionic acid, all three had been

previously examined using a $^{32}PP_i$ assay to test for successful adenylation. All three were determined to be unsuitable for adenylation (and thus acylation), although reexamination of L-serine in later work showed some incorporation of the radioactive form of the substrate.[21, 29] Reevaluation of these substrates using MS, however, displayed acylation of all three to ~45–95% occupancy. While it is not clear why the $^{32}PP_i$ assay was unable to identify these substrates as candidates for acylation, it is clear from these results that while the $^{32}PP_i$ assay can provide a good assessment of adenylation domain activity, it is not a completely reliable method for discovering alternative substrates.

Despite successful acylation of many substrate analogues, nearly all of the compounds examined proved to be poor substrates for condensation with the upstream salicyl-S-ArCP donor, providing definitive evidence of a "gatekeeper" function exhibited by the Cy_1 domain. However, despite this strong editing function, 2,3-diaminopropionic acid demonstrated condensation with the upstream donor, forming the uncyclized complex intermediate (Figure 4C). The successful condensation of this substrate represents both an important and a curious exception to the editing function demonstrated by the Cy_1 domain. Currently, there are two postulated mechanisms for condensation between the ArCP and PCP1 active sites (Scheme 2). The standard NRPS mechanism entails a nucleophilic attack by the α-amino group on the upstream thioester, while the thiol is responsible for substrate recognition and subsequent cyclodehydration. If this mechanism is true, then the Cy_1 domain of HMWP2 can differentiate completely between the thiol of L-cysteine versus both the hydroxyl of L-serine and the methyl of α-amino-L-butyric acid among other side-chain forms, but the β-amino group of 2,3-diaminopropionic acid is only partially discriminated. An alternative mechanism employs the side chain of cysteine as the condensing nucleophile, with an S → N acyl shift perhaps preceding cyclodehydration (Scheme 2). Because the masses of the condensed products of 2,3-diaminopropionic acid and L-cysteine with salicylic acid, by either the amide or the mixed thioester mechanisms, are equal, this question cannot be answered through direct observation of the intermediate. Loading of the des-amino substrate, 3-mercaptopropionic acid, was attempted to observe the competency of the thiol for condensation, but the substrate was not accepted by the adenylation domain for acylation onto the PCP1 active site (data not shown). While the current data do not resolve the mechanism, successful trapping of the condensed intermediate in the millisecond time regime followed by the alkylation of the free thiol (if present) and FTMS analysis could differentiate between the isomeric condensation products.

Intermodule Communication between the ArCP and PCP1 Domains

While the aforementioned analysis of the specificity and mechanics of acylation and condensation at the ArCP and PCP1 active sites provides insight into each domain's individual characteristics, the capability of one active site to adjust its loading capability with respect to the covalent state of the other was also illuminated. The occupancy of each ArCP substrate analogue was lowered ≤5% when withholding the loaded downstream elements at the PCP1 domain versus an identical control reaction where L-cysteine was present. These differences in percent occupancy are on the fringe of experimental precision and demonstrate that if the ArCP active site possesses a dependence on the covalent state of the PCP1 domain, it is a weak effect (Supporting Information Table 1). Similar experiments on the PCP1 active site revealed a ~25% drop in the acylation of L-cysteine without an upstream salicylS-ArCP intermediate versus an identical control reaction, where salicyl-S-ArCP was present. While these data demonstrate some intermodule communication between the ArCP and PCP1 active sites, the results from the PCP1 substrate analogues were variable (Supporting Information Table 2, Figure 5). It is not clear why the other alternative substrates display such heterogeneity in loading capability as a function of the covalent state of the ArCP active site.

Kinetics of Acylation versus Condensation

Through direct observation of intermediates on the 100 ms to 20 s time scale, relative rates of aminoacylation of L-cysteine versus condensation were assessed on the PCP1 active site (Figure 6). The successful detection of an on-pathway intermediate requires that the reaction rates involved in its formation be greater than the rates of its further processing. The accumulation of the cysteine-S-PCP1 intermediate to ~45% and the HPT-S-PCP1 intermediate to ~5% after 250 ms with a primed ArCP active site demonstrates that the rate of substrate acylation (denoted k_2 in Scheme 3) must be faster than the combined rate of condensation and hydrolysis (k_{-2} + k_3). While the cysteine-S-PCP1 intermediate appears in an oxidized form in Figure 6, we have observed the direct cysteine-S-PCP1 intermediate in some kinetic experiments (data not shown). Further, the steady decrease in the percent occupancy of the oxidized cysteine-S-PCP1 intermediate over time suggests that it is not an off-pathway intermediate, but is rather due to the artifactual oxidation of the thioester-bound cysteine during and after the rapid quench experiment.

General Kinetic Model

For NRP and PK synthetases in general, the rates of monomer loading versus translocation of intermediates are in competition. With the rate of cysteine acylation faster than the rate of condensation, we forward here a “crawling” model for NRPS systems, where many complex biosynthetic intermediates are stacked up at high occupancy. In the competing “fast translocation” model, wherein the rate of condensation is much faster than monomer loading, complex intermediates will not be observed at high occupancy unless substrates are withheld or enzymes truncated or mutated.

Through the linear fitting of the early time points, the initial velocity of complex intermediate accumulation at PCP1 (Sal-Cys-S-PCP1 and HPT-S-PCP1) was determined to be 19 s^{-1} (0.3 min^{-1}), for this particular set of loading conditions. This MS-based rate of intermediate accumulation is useful for semiquantitatively determining the relative rates of flux throughout the enzyme, thus identifying reaction bottlenecks.[38] Future investigations of the kinetics and timing of HMWP2 entail calculating a similar rate of accumulation of complex intermediates for PCP2 and other downstream carrier sites, monitoring the effects of substrate concentration on the initial velocity (i.e., derivation of rate constants), and observing the changes in the rate induced by full reconstitution of HMWP2 with the other enzymes and constituents required to generate yersiniabactin. Ultimately, the inclusion of more branch points within the timing of the yersiniabactin bioassembly will create the need for complex kinetic modeling of the system. This work represents the first study where by combining fast quench with large molecule MS, kinetic parameters could be determined from a NRP/PK system to visualize thiotemplate biosynthesis.

Conclusion

The structural diversity of the products NRP and PK synthetases generate is staggering. Thorough understanding of the kinetics and mechanisms of these systems will ultimately lead to improved engineering of new NRP/PK systems capable of efficiently producing novel, medicinally useful compounds. However, to study the vast number of natural NRP/PK synthetases in addition to the multitude of mutant and engineered systems, an efficient and general measurement platform is necessary. Digestion of the carrier domains at specific time points within the synthesis of the product followed by high-resolution MS analysis provides a method by which each system of interest can be reliably assessed. This direct-interrogation approach is especially well positioned to resolve questions of enzyme skipping and stuttering[36] or to resolve instances where the substrates are not readily predicted.

Acknowledgements

We thank members of the Kelleher Lab, especially Matthew Mazur, Leah Miller, and Leslie Hicks, for helpful discussions and assistance with this work. We also thank Suzanne Admiraal and Fei Liu for supplying the plasmids encoding Sfp, HMWP2, and YbtE, as well as for helpful discussions. This paper is dedicated to Dr. Christopher Walsh on the occasion of his 60th birthday. This work was supported by National Institutes of Health grant GM 067725.

References

1. Mootz HD, Schwarzer D, Marahiel MA. ChemBioChem 2002;3:490–504. [PubMed: 12325005]
2. Marahiel MA, Stachelhaus T, Mootz HD. Chem Rev 1997;97:2651–2673. [PubMed: 11851476]
3. Schwarzer D, Finking R, Marahiel MA. Nat Prod Rep 2003;20:275–287. [PubMed: 12828367]
4. Moore BS, Hertweck C, Hopke JN, Izumikawa M, Kalaitzis JA, Nilsen G, O'Hare T, Piel J, Shipley PR, Xiang L, Austin MB, Noel JP. J Nat Prod 2002;65:1956–1962. [PubMed: 12502351]
5. Walsh CT. Science 2004;303:1805–1810. [PubMed: 15031493]
6. Cane DE, Walsh CT, Khosla C. Science 1998;282:63–68. [PubMed: 9756477]
7. Salomon CE, Magarvey NA, Sherman DH. Nat Prod Rep 2004;21:105–121. [PubMed: 15039838]
8. Rodriguez E, McDaniel R. Curr Opin Microbiol 2001;4:526–534. [PubMed: 11587928]
9. Schneider TL, Walsh CT, O'Connor SE. J Am Chem Soc 2002;124:11272–11273. [PubMed: 12236733]
10. Admiraal SJ, Khosla C, Walsh CT. J Am Chem Soc 2003;125:13664–13665. [PubMed: 14599196]
11. Admiraal SJ, Khosla C, Walsh CT. Biochemistry 2002;41:5313–5324. [PubMed: 11955082]
12. Tseng CC, McLoughlin SM, Kelleher NL, Walsh CT. Biochemistry 2004;43:970–980. [PubMed: 14744141]
13. Chen HW, Tseng CC, Hubbard BK, Walsh CT. Proc Natl Acad Sci USA 2001;98:14901–14906. [PubMed: 11752437]
14. Pfeifer V, Nicholson GJ, Ries J, Recktenwald J, Schefer AB, Shawky RM, Schröder J, Wohlleben W, Pelzer S. J Biol Chem 2001;276:38370–38377. [PubMed: 11495926]
15. Crosa JH, Walsh CT. Microbiol Mol Biol Rev 2002;66:223–249. [PubMed: 12040125]
16. Marshall CG, Burkart MD, Meray RK, Walsh CT. Biochemistry 2002;41:8429–8437. [PubMed: 12081492]
17. Quadri LEN, Sello J, Keating TA, Weinreb PH, Walsh CT. Chem Biol 1998;5:631–645. [PubMed: 9831524]
18. Miller DA, Luo L, Hillson N, Keating TA, Walsh CT. Chem Biol 2002;9:333–344. [PubMed: 11927258]
19. Pfeifer BA, Wang CCC, Walsh CT, Khosla C. Appl Environ Microbiol 2003;69:6698–6702. [PubMed: 14602630]
20. Quadri LEN, Keating TA, Patel HM, Walsh CT. Biochemistry 1999;38:14941–14954. [PubMed: 10555976]
21. Keating TA, Miller DA, Walsh CT. Biochemistry 2000;39:4729–4739. [PubMed: 10769129]
22. Suo Z, Chen H, Walsh CT. Proc Natl Acad Sci USA 2000;97:14188–14193. [PubMed: 11106385]
23. Miller DA, Walsh CT, Luo LS. J Am Chem Soc 2001;123:8434–8435. [PubMed: 11516308]
24. Shaw-Reid CA, Kelleher NL, Losey HC, Gehring AM, Berg C, Walsh CT. Chem Biol 1999;6:385–400. [PubMed: 10375542]
25. Hicks L, Weinreb P, Konz D, Marahiel MA, Walsh CT, Kelleher NL. Anal Chim Acta 2003;496:217–224.
26. Hicks LM, Connor S, Walsh CT, Kelleher NL. Chem Biol 2004;11:327–335. [PubMed: 15123262]
27. Marshall AG, Hendrickson CL, Jackson GS. Mass Spectrom Rev 1998;17:1–35. [PubMed: 9768511]
28. Kelleher NL, Senko MW, Siegel MM, McLafferty FW. J Am Soc Mass Spectrom 1997;8:380–383.
29. Keating TA, Suo ZC, Ehmann DE, Walsh CT. Biochemistry 2000;39:2297–2306. [PubMed: 10694396]
30. Gehring AM, DeMoll E, Fetherston JD, Mori I, Mayhew GF, Blattner FR, Walsh CT, Perry RD. Chem Biol 1998;5:573–586. [PubMed: 9818149]

31. Gehring AM, Mori I, Perry RD, Walsh CT. Biochemistry 1998;37:11637–11650. [PubMed: 9709002]
32. Senko MW, Hendrickson CL, Emmett MR, Shi SDH, Marshall AG. J Am Soc Mass Spectrom 1997;8:970–976.
33. Horn DM, Zubarev RA, McLafferty FW. J Am Soc Mass Spectrom 2000;11:320–332. [PubMed: 10757168]
34. Quadri LEN, Weinreb PH, Lei M, Nakano MM, Zuber P, Walsh CT. Biochemistry 1998;37:1585–1595. [PubMed: 9484229]
35. Mazur MT, Walsh CT, Kelleher NL. Biochemistry 2003;42:13393–13400. [PubMed: 14621984]
36. Tang L, Ward S, Chung L, Carney JR, Li Y, Reid R, Katz L. J Am Chem Soc 2004;126:46–47. [PubMed: 14709052]
37. Severinova E, Severinov K, Fenyo D, Marr M, Brody EN, Roberts JW, Chait BT, Darst SA. J Mol Biol 1996;263:637–647. [PubMed: 8947564]
38. Xu G, Zhai H, Narayan M, McLafferty FW, Scheraga HA. Chem Biol 2004;11:517–524. [PubMed: 15123246]

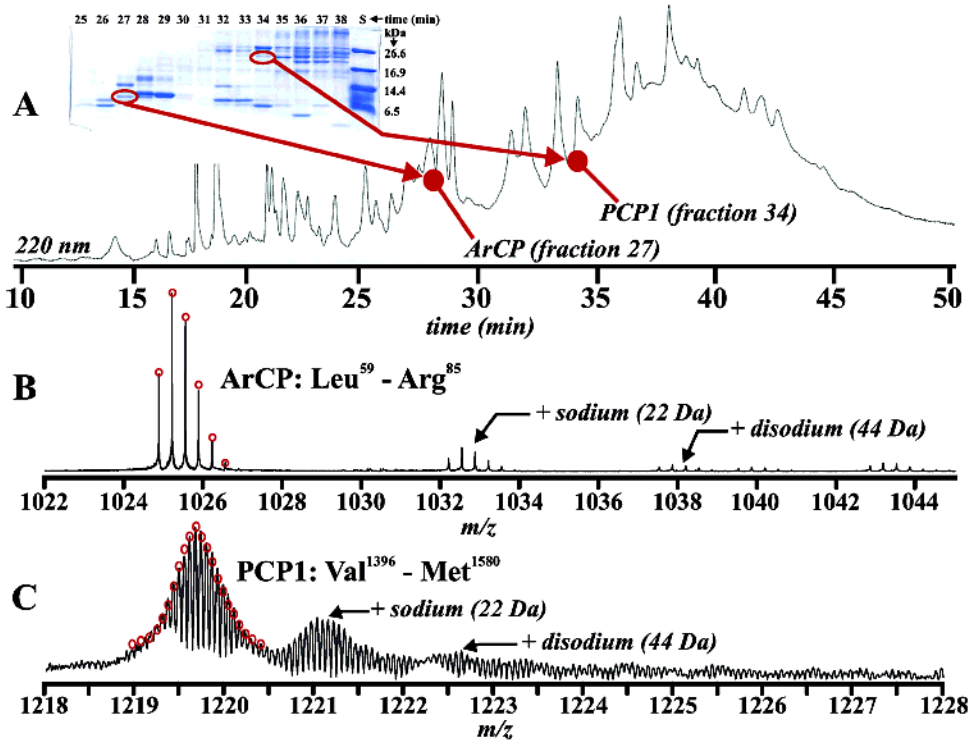


Figure 1.

(A) Reverse-phase fractionation of a cyanogen bromide digestion of apo-HMWP2 visualized by SDS-PAGE. Fractions 27 and 34 yielded identification of the ArCP active site peptide (Ile33–Met^{87}, 6186.19 Da) and the PCP1 active site peptide (Val^{1396}–Met^{1580}, 20 704.5 Da). (B) Isotopic fit of the FTMS data to the ArCP active site peptide generated by trypsin (Leu^{59}–Arg^{85}; 3071.63 Da, 25 scans). (C) Isotopic fit of the FTMS data to the CNBr-digested PCP1 active site peptide (150 scans).

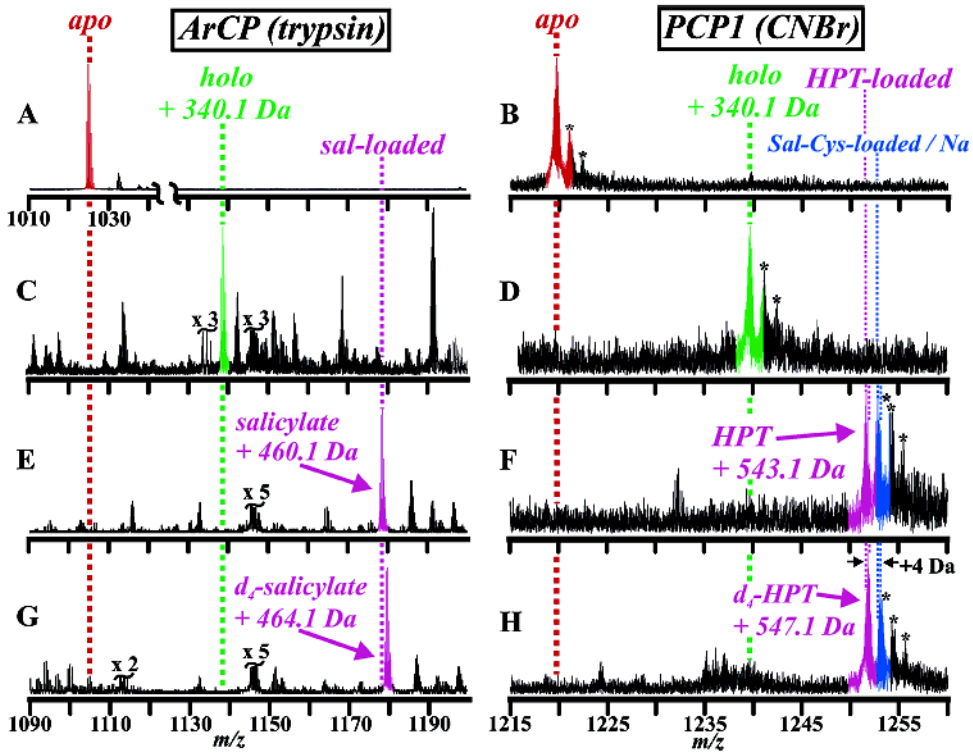


Figure 2.

Various covalent states of the trypsinized ArCP and CNBr digested PCP1 active site peptides. The asterisks denote salt adducts, and some of the unrelated peaks have been truncated for spectral clarity (i.e., ×3). (A) Apo form ArCP (3071.63 Da). (B) Apo form PCP1 (20 704.5 Da). (C) Holo-ArCP after incubation with CoA and Sfp (3411.73 Da). (D) Holo-PCP1 after a similar incubation in spectrum C (21 044.6 Da). (E) Salicylate loaded ArCP using YbtE, salicylic acid, and ATP for 30 min (3531.73 Da). (F) HPT loaded PCP1 using the same incubation parameters as (E) only with L-cysteine (21 247.6 Da). (G) The same incubation as shown in (E), but with d_6-salicylate (3535.73 Da). (H) The same incubation as shown in (F), but with d_6-salicylate (21 251.6 Da).

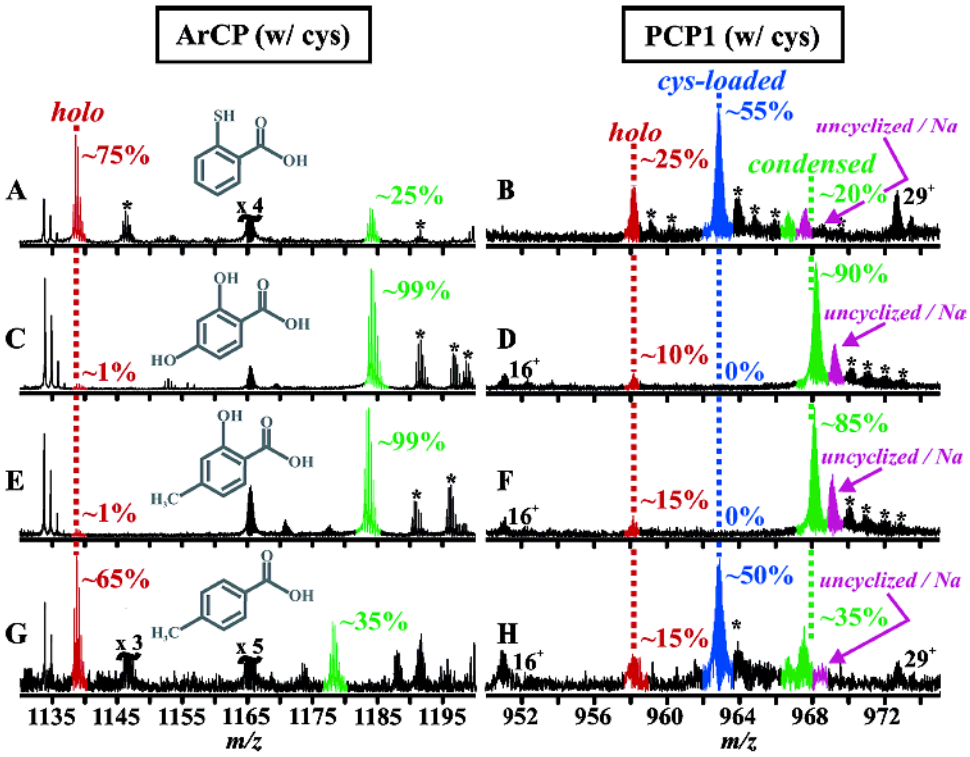


Figure 3.

Unnatural substituted benzoate incorporation into the ArCP and PCP1 active sites after 30 min. The asterisks denote salt adducts, and some of the unrelated peaks have been truncated for spectral clarity (i.e., ×3). (A) 2-Thiosalicylic acid (3547.65 Da). (B) Condensation of 2-thiosalicylic acid with L-cysteine (21 231.4 Da). (C) 2,4-Dihydroxybenzoic acid (3547.65 Da). (D) Condensation of 2,4-dihydroxybenzoic acid with L-cysteine (21 265.5 Da). (E) 4-Methylsalicylic acid (3545.73 Da). (F) Condensation of 4-methylsalicylic acid with L-cysteine (21 263.5 Da). (G) *p*-Toluic acid (3529.73 Da). (H) Condensation of *p*-toluic acid with L-cysteine (21 247.4 Da).

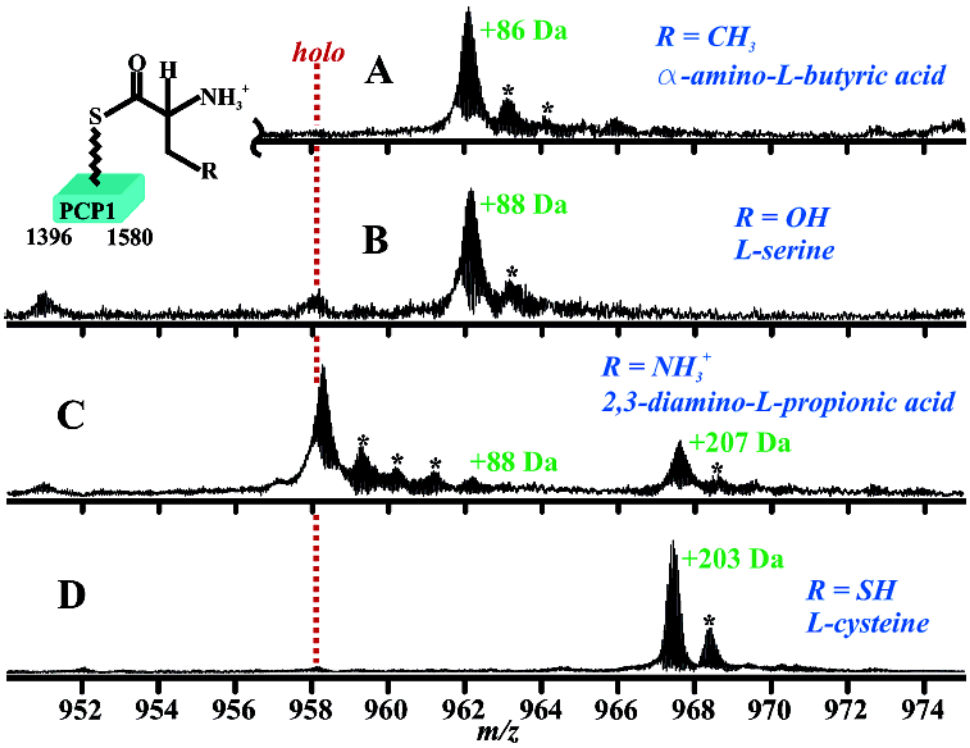


Figure 4.

Acylation and condensation efficiency of substrate analogues for the PCP1 active site including α-amino-L-butyric acid (A; 20 min, 21 130.7 Da), L-serine (B; 20 min, 21 132.7 Da), 2,3-diaminopropionic acid (C; 2 h, 21 132.7 Da (acylated) and 21 251.3 Da (condensed)), and L-cysteine (D; 20 min, 21 247.6 Da (HPT) and 21 267.6 Da (Sal-Cys)). The asterisks denote salt adducts.

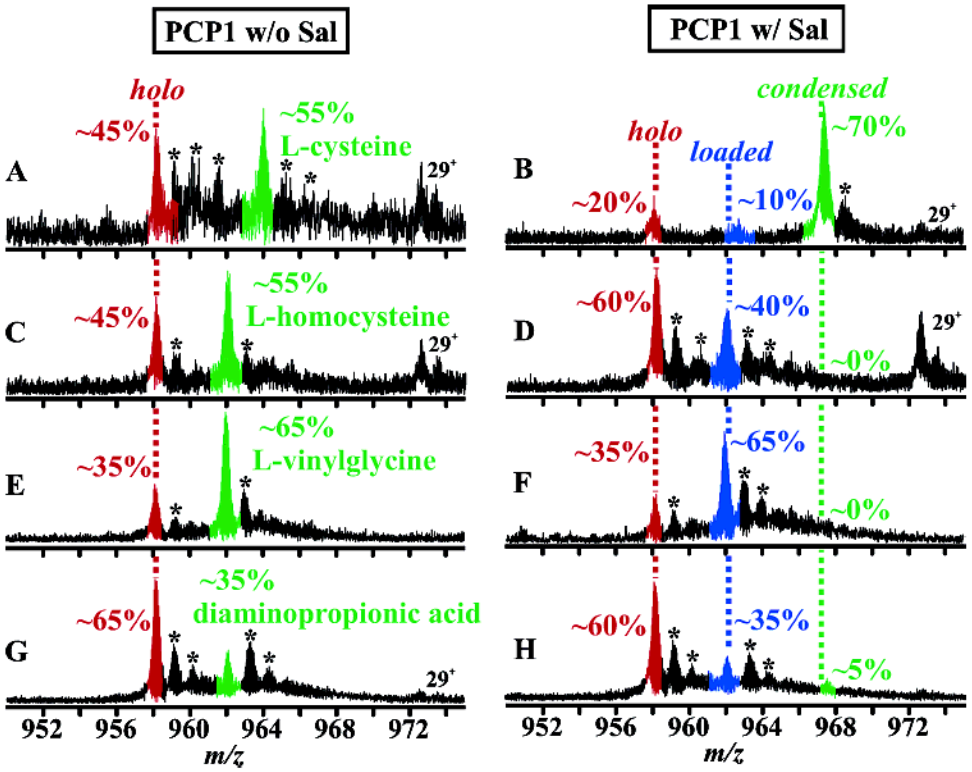


Figure 5.

Alternative substrate loading on the PCP1 active site and the effects of the covalent state of the ArCP active site on the PCP1 domain loading efficiency. (A) L-Cysteine incorporation in the absence of salicyl-S-ArCP (cysteine-S-PCP1 = 21 148.7 Da). (B) L-Cysteine incorporation and condensation with an upstream salicyl-S-ArCP (HPT-S-PCP1 = 21 247.7 Da). (C) L-Homocysteine incorporation in the absence of salicyl-S-ArCP (homocysteine-S-PCP1 = 21 128.7 Da). (D) Lack of L-homocysteine condensation with an upstream salicyl-S-ArCP. (E) 4-Vinyl-L-glycine incorporation in the absence of salicyl-S-ArCP (vinylglycine-S-PCP1 = 21 128.7 Da). (F) Lack of 4-vinyl-L-glycine condensation with an upstream salicyl-S-ArCP. (G) 2,3-Diaminopropionic acid incorporation in the absence of salicyl-S-ArCP (diaminopropionyl-S-PCP1 = 21 132.7 Da). (H) 2,3-Diaminopropionic acid incorporation with an upstream salicyl-S-ArCP (uncyclized product = 21 251.3 Da). The asterisks denote salt adducts, and irrelevant ions are denoted by their charge state.

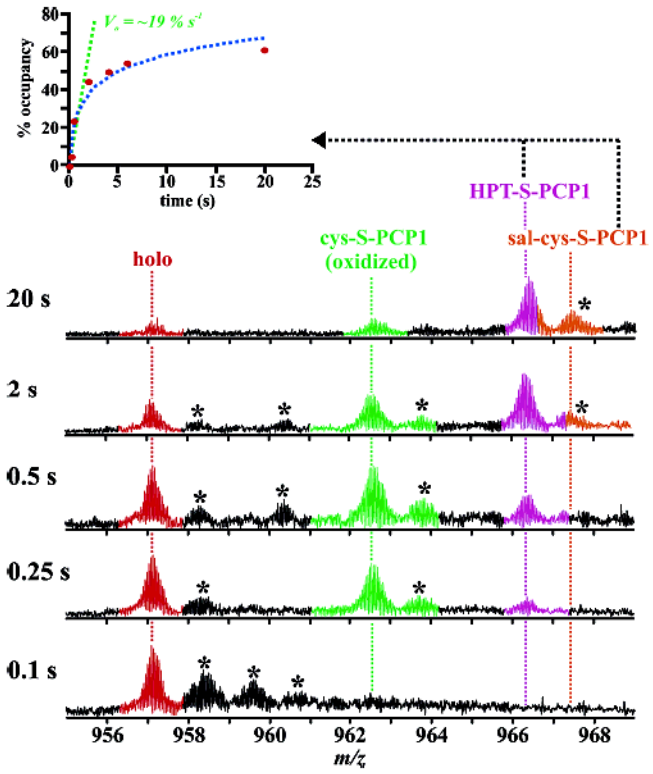
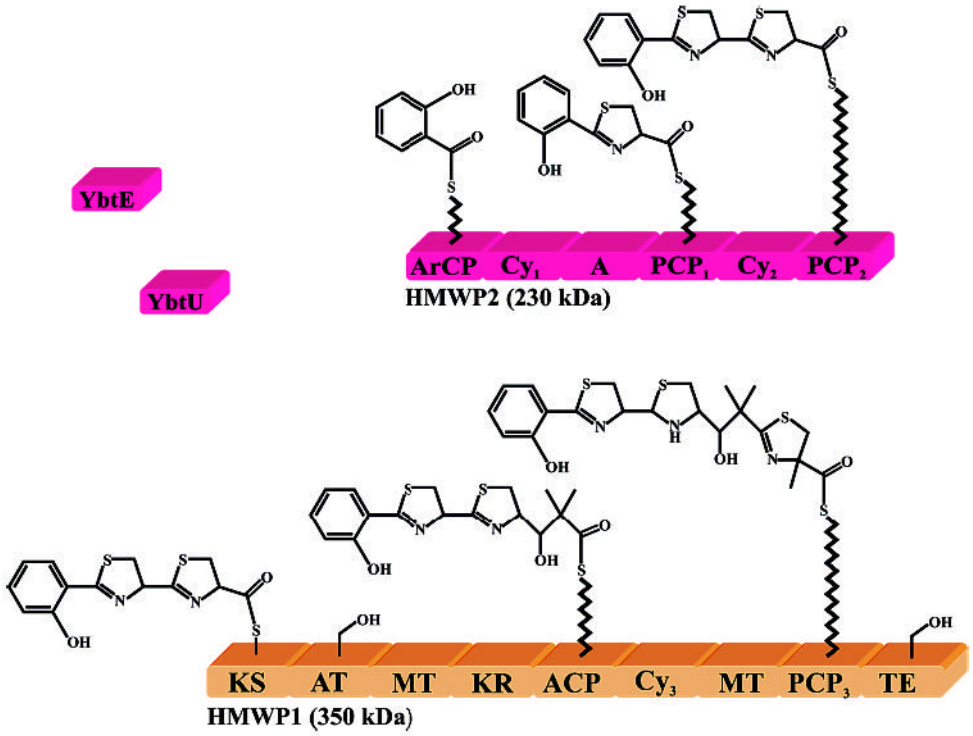
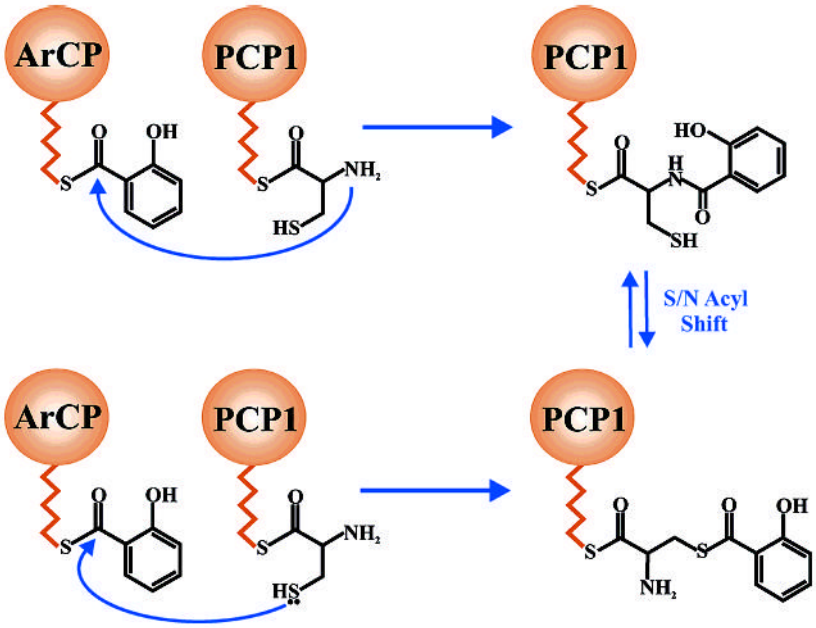


Figure 6.
Kinetic time course of PCP1 intermediate formation. The inset is a combined plot of percent occupancy for the Sal-Cys-S-PCP1 and HPT-S-PCP1 intermediates versus time.

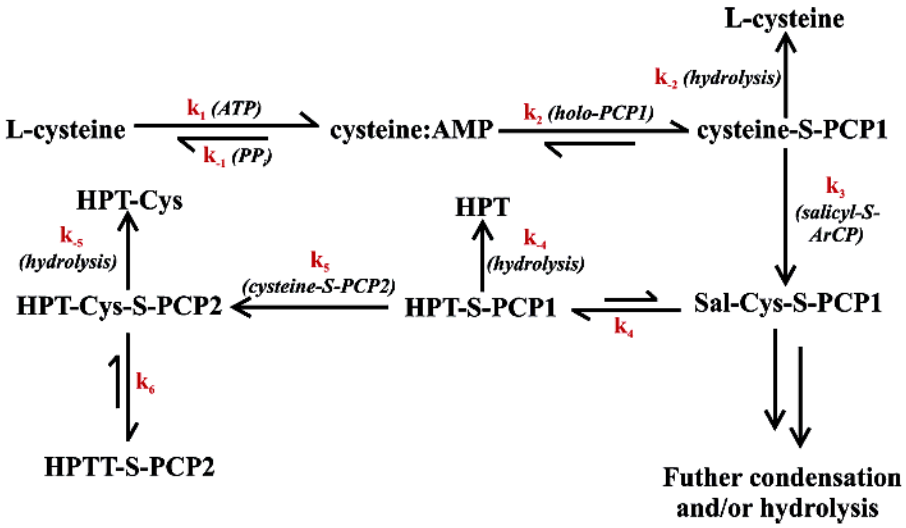


Scheme 1.
Thiotemplate Bioassembly of Yersiniabactin by Yersiniabactin Synthetase YbtE, YbtU, HMWP2, and HMWP1



Scheme 2.

Alternative Mechanisms of the Condensation Event between the ArCP and PCP1 Active Sites of Yersiniabactin Synthetase



Scheme 3.
Complex Intermediate Accumulation at the PCP1 Active Site

Table 1

Alternative Substrate Incorporation into the ArCP Active Site[a]

Substrate	% holo	% loaded
salicylic acid	1%	99%
benzoic acid	20%	80%
2,3-dihydroxybenzoic acid	5%	95%
2,4-dihydroxybenzoic acid	1%	99%
2,6-dihydroxybenzoic acid	15%	85%
4-methylsalicylic acid	1%	99%
4-hydroxybenzoic acid	80%	20%
p-toluic acid	65%	35%
2-thiosalicylic acid	75%	25%
anthranilic acid	100%	0%
3-hydroxypicolinic acid	50%	50%
2-hydroxynicotinic acid	85%	15%
L-phenylalanine[b]	100%	0%
L-tyrosine[b]	100%	0%

[a] Each reaction contained holo-HMWP2 incubated in 2 mM of both the designated substrate and L-cysteine for 30 min prior to trypsinization.

[b] Further incubation for 60 min still resulted in no loading of the substrate.

Table 2

Selectivity of Acylation and Condensation in the PCP1 Domain[a]

Substrate	% holo	% loaded	% condensed
L-cysteine	5%	0%	95%
L-homocysteine	60%	40%	0%
L-serine	5%	95%	0%
4-vinyl-L-glycine	40%	60%	0%
2,3-diamino-L-propionic acid	50%	45%	5%[b]
α-amino-L-butyric acid	1%	99%	0%

[a] Each reaction contained holo-HMWP2 incubated in 2 mM of both the noted substrate and salicylic acid for 20 min prior to CNBr digestion.

[b] Percent occupancy increases to ~30% in a 2 h reaction (see Figure 4C).