Yersiniabactin Synthetase: Probing the Recognition of Carrier Protein Domains by the Catalytic Heterocyclization Domains, Cy1 and Cy2, in the Chain-Initiating HMWP2 Subunit†

Deborah Ann Miller and Christopher T. Walsh*

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

Received December 21, 2000; Revised Manuscript Received February 21, 2001

ABSTRACT: The HMWP2 subunit of yersiniabactin (Ybt) synthetase, a 230 kDa nonribosomal peptide synthetase (NRPS) making the N-terminus of the Ybt siderophore of *Yersinia pestis*, has one cysteine-specific adenylation (A) domain, three carrier protein domains (ArCP, PCP1, PCP2), and two heterocyclization domains (Cy1, Cy2). The A domain loads the two PCP domains with cysteines that get heterocyclized by the Cy domains to yield a tricyclic hyrdoxyphenylthiazolinylthiazolinyl (HPTT) chain lodged in thioester linkage to the PCP2 domain. The interdomain recognition by the Cy1 and Cy2 domains for the three carrier proteins was tested using inactivating mutations at the conserved serine that is phosphopantetheinylated in each carrier domain (S52A, S1439A, and S1977A). These mutant forms of HMWP2 were tested for *in trans* complementation by carrier protein fragments: holo-ArCPs (S52A), holo-PCP1 and analogues (S1439A), and holo-PCP2 and analogues (S1977A). The S52A mutant tests the recognition of the Cy1 domain for donor acyl-ArCP substrates, while the S1439A mutant tests the specificity of the same Cy1 domain for downstream substrates presented by distinct PCPs. The S1439A likewise tests the recognition of Cy2 for its upstream PCP-tethered acyl donor. The S1977A mutant analogously tests the Cy2 domain for downstream Cys-PCP recognition. In all cases *in trans* complementation was successful with the carrier protein fragments, allowing kinetic probes of catalytic efficiency for PCP scaffolds and for uncoupling of the condensation and heterocyclization functions of Cy1 and Cy2. Overall, the Cy domains tested showed a definite selectivity for the upstream protein scaffold but were more relaxed toward the downstream acceptor protein. This work points to the importance of protein−protein interactions in mediating directional chain growth in NRPS and presents the first systematic exploration of how the protein scaffolds affect catalytic efficiency.

Biosynthesis of the nonribosomal peptide siderophore yersiniabactin (**1**), a virulence factor of the plague bacterium *Yersinia pestis* (*1*−*3*), is brought about by a five protein, eighteen domain enzymatic assembly line with both nonribosomal protein synthetase (NRPS)[1] and polyketide synthase (PKS) modules (*4*−*6*). YbtE is the 57 kDa salicylate activating subunit; HMWP2 is a 230 kDa six domain double NRPS module; HMWP1 is a 350 kDa protein with nine domains distributed between a PKS and then an NRPS module; YbtU is a 41 kDa putative reductase module; and YbtT is a 30 kDa putative external thioesterase (TE) domain (*7*, *8*). Genetic studies (*2*) have shown that YbtU and YbtT are required for in vivo production of yersiniabactin, but these functions have yet to be validated in vitro. Siderophore chain assembly starts on the HMWP2 scaffold with three carrier protein domains (ArCP, PCP1, PCP2), one cysteine-activating domain (A), and two peptide bond-forming condensation/heteocyclization (Cy1, Cy2) domains (Figure 1A). We have previously demonstrated that the primed, phosphopantetheinylated forms of the three carrier protein domains can be covalently loaded with salicyl, cysteinyl, and cysteinyl monomers, respectively, and that Cy1 makes the Sal-Cys peptide bond and cyclizes and dehydrates it to the hydroxyphenylthiazolinyl-S-PCP1. Then Cy2 acts analogously, moving the acyl chain to the most downstream carrier domain, PCP2, as it condenses and cyclizes the second cysteinyl moiety to produce the tandem bisthiazolinyl acyl-enzyme, HPTT-S-PCP2 (Figure 1B) (*9*, *10*).

The HMWP2 subunit, with three thiolated carrier protein domains and two condensing, heterocyclization domains, making three detectable covalent acyl-S-enzyme intermediates that can be thiolyzed or hydrolyzed to a suite of products separable by HPLC (*10*) is a good system to dissect the recognition of cyclization catalytic domains for their substrates, upstream and downstream acyl-S-carrier protein domains embedded *in cis* in the assembly line. Previous

† This work was supported in part by National Institutes of Health Grant AI 42736 to C.T.W.

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

[1] Abbreviations: A, adenylation domain; ArCP, aryl carrier protein domain; C, condensation domain; CoA, coenzyme A; Cy, cyclization domain; DHB, 2,3-dihydroxybenzoate; DHPT, 2-(2,3-dihydroxyphenyl)thiazolinyl; HMWP, high molecular weight protein; HPT, 2-(2-hydroxyphenyl)thiazolinyl; HPTT, 2-(2-hydroxyphenyl)thiazolyl-2,4-thiazolinyl; NRPS, nonribosomal peptide synthetase; PCP, peptidyl carrier protein domain; PKS, polyketide synthase; P-pant, 4′-phosphopantetheine; PPTase, phosphopantetheinyl transferase; Sal, salicylate; TCEP, tris(carboxyethyl)phosphine; TE, thioesterase domain; Ybt, yersiniabactin.

10.1021/bi002905v CCC: $20.00
Published on Web 04/03/2001

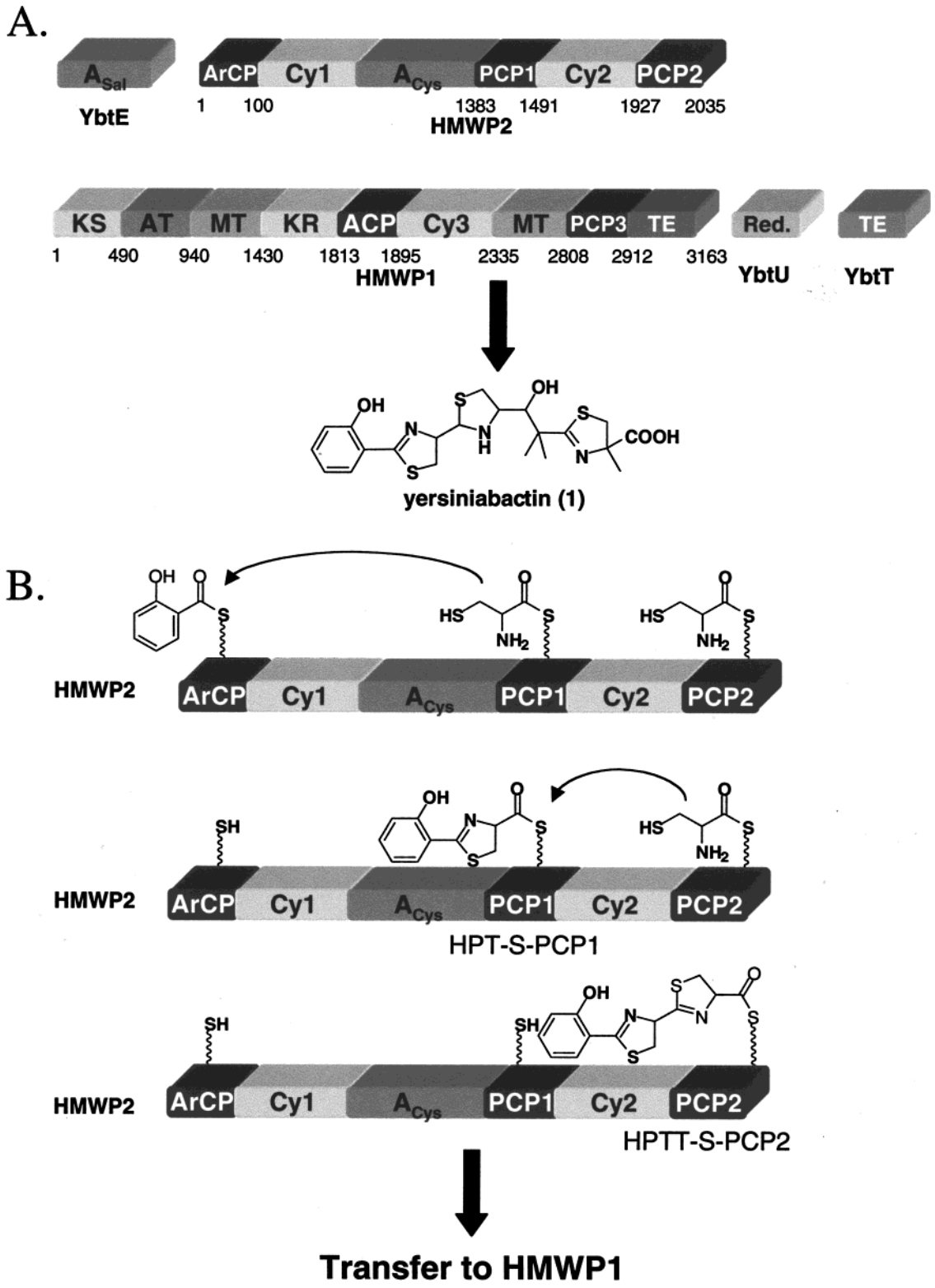

FIGURE 1: (A) Yersiniabactin (Ybt, **1**) and the five proteins that comprise the yersiniabactin synthetase. YbtE is a single domain, salicyl-AMP ligase. YbtU is a putative reductase with 27% identity to PchG, the enzyme responsible for reduction during pyochelin biosynthesis in *Pseudomonas aeruginosa*. YbtT is a putative external TE domain. HMWP2 and HMWP1 encompass the remaining 15 domains. Abbreviations: A, adenylation; ArCP, aryl carrier protein; Cy, cyclization; PCP, peptidyl carrier protein; KS, ketosynthase; AT, acyltransferase; MT, methyltransferase; KR, ketoreductase; ACP, acyl carrier protein; TE, thioesterase. The residue marking each domain boundary is indicated by the numbers below the protein. (B) Schematic of chain elongation catalyzed by HMWP2. Covalent intermediates are drawn attached to the appropriate carrier protein domains through a 4′-phosphopantetheine (wavy line).

studies have shown that ArCP and PCP domain fragments can be expressed, purified, posttranslationally phosphopantetheinylated by a member of the phosphopantetheinyl transferase family (Sfp, EntD, etc.) (*11*), and acyl/aminoacyl adenylated *in trans* by their corresponding A domain (*8*, *12*). Therefore, the stage is set for the exploration of the protein−protein interactions between the two Cy domains of HMWP2 and loaded ArCP or PCP domains. It is not clear if

recognition is composite for both the acyl/aminoacyl chain on the P-pant arm and the surface of the given carrier protein domains that are most proximal upstream and downstream in the assembly line. Previous studies have addressed the question of condensation domain specificity for the acyl/aminoacyl substrates (*13*, *14*). The three carrier protein domains of HMWP2 and a cognate PCP3 in the HMWP1 subunit are all about 10 kDa, have signature sequences [G(G/A)(D/T)**S**L] around the serine that is posttranslationally phosphopantetheinylated, but otherwise show only 16−34% homology. The three solution structures solved for carrier protein domains [an ACP from *Escherichia coli* fatty acid synthase, an ACP from actinorhodin PKS from *Streptomyces coelicolor*, and a PCP of *Bacillus brevis* tyrocidine synthetase 3 (TycC)] are structurally homologous four-helix bundles (*15*−*17*). Thus, the determinants of specificity within this protein scaffold for recognition by specific partner protein domains remain unclear.

The chain-elongating condensation domains and their heterocyclizing Cy variants of NRPS assembly lines use the upstream acyl/peptidyl-S-carrier protein domains as donor substrates and the downstream aminoacyl-S-PCP domains as monomer nucleophilic acceptor substrates in the condensations (*18*, *19*). In this work, we evaluated selectivity of the Cy1 and Cy2 domains through the use of Ser to Ala mutants in the ArCP (S52A), PCP1 (S1439A), and PCP2 (S1977A) (*10*) domains that render any one of the carrier proteins to the apo, nonfunctional form and tested the ability of various acyl-S-ArCP/PCP domains to substitute *in trans*. These *in trans* studies can generate multiple catalytic turnovers enabling a kinetic analysis of how the different protein scaffolds affect catalysis.

MATERIALS AND METHODS

ATP, coenzyme A, imidazole, and magnesium chloride were purchased from Sigma Chemical Co. Tris was purchased from J. T. Baker. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was purchased from Molecular Probes (Eugene, OR). Luria−Bertani (LB) media was prepared and used for culturing *E. coli* strains as previously described (*20*). Competent cells of *E. coli* strains DH5α were purchased from GibcoBRL; cells of BL21(DE3) were purchased from Novagen. Preparation of plasmid DNA was performed using QIAprep plasmid miniprep kits (Qiagen). Salicyl-SNAC was a gift from H. Chen (*21*). Protein concentrations were determined from the predicted molar extinction coefficient (ϵ) and absorbance at 280 nm. HPLC analyses were performed on a Beckman System Gold equipped with a VYDAC C18 reverse-phase analytical column. Detection was at 254 nm, with mobile phase A, 100 μL of formic acid/200 μL of triethylamine in 1 L of water, and mobile phase B, 4:1 acetonitrile:mobile phase A. At a flow rate of 1 mL/min, a linear gradient was maintained from 10% mobile phase B to 100% mobile phase B over 23 min. Positive ion detection electrospray mass spectrometry (ESMS) was performed by the Mass Spectrometry Facility of the Department of Chemistry and Chemical Biology, Harvard University.

Overexpression and Purification of Proteins. The HMWP2 full-length mutants [S52A, S1439A, S1977A (*10*)] and HMWP2 fragments [1−1382, 1383−2035, S1977A (1383−2035 construct) (*9*), ArCP (1−100) (*8*), PCP1, PCP2, and PCP3 (*7*, *8*)] were prepared following protocols described elsewhere. YbtE (A_{Sal} domain) was made following the protocol outlined previously (*8*). HMWP1 fragment PCP3TE (*22*), the PchE ArCP and PchD (A_{DHB} domain) (*23*), and VibB (the ArCP fragment and full length), and VibE (A_{DHB} domain) (*24*) were gifts from Z. Suo, H. M. Patel, and T. A. Keating, respectively. SrfB1 PCP (*25*), EntB ArCP, and EntE (A_{DHB} domain) (*26*) were gifts from D. E. Ehmann. Sfp, a phosphopantetheinyl transferase from *Bacillus subtilis*, was prepared previously (*27*).

Isolation and Identification of Products Synthesized by the in trans Reactions with the ArCP Domains or PCP Domains. A solution (500 μL) of 5 μM HMWP2 S52A mutant and 15 μM various ArCPs with 0.5 μM corresponding A domains (i.e., Ybt ArCP/YbtE; PchE/PchD; VibB/VibE; EntB/EntE), 0.3 μM Sfp, 75 mM Tris-HCl, pH 8.6, 10 mM $MgCl_2$, 5 mM TCEP, 1 mM CoASH, 5 mM L-cysteine, and 1 mM salicylate (or 10 mM DHB) was incubated at 30 °C for 40 min to allow phosphopantetheinylation of HMWP2 fragments. Previous experiments have shown that the A domains used are capable of loading both salicylate and DHB onto their corresponding ArCP domain (*28*). Then, 10 mM ATP was added, and the reaction mixtures were incubated for 14 h at 30 °C. The salicyl-SNAC (0.3 mM) was also tested for its ability to substitute for a salicyl-loaded ArCP in a reaction identical to that described. A similar reaction mixture was made for the *in trans* reactions with various PCP domains; the solution contained 5 μM HMWP2 S1439A or S1977A mutant and 15 μM various PCP domains, 0.3 μM Sfp, 0.5 μM YbtE, 75 mM Tris-HCl, pH 8.6, 10 mM $MgCl_2$, 5 mM TCEP, 1 mM CoASH, 5 mM L-cysteine, and 1 mM salicylate. The reactions were quenched with 100 μL of 8.5% phosphoric acid. The acidified mixture was extracted with 1 mL of ethyl acetate three times. The organic layer was combined in a glass vial and dried under reduced pressure. The dried residue was submitted for LC-MS positive ion trap mass spectrometry analysis (Chemistry Department, Harvard University).

Time Courses of Intermediate Formation at 30 °C Catalyzed by HMWP2 S52A, S1439A, and S1977A. The reactions were carried out as described above with a slight variation. After ATP was added, the reaction mixtures (500 μL) were incubated at 30 °C, and 50 μL samples were removed at various times, acidified by addition of 10 μL of 8.5% phosphoric acid, and extracted with 1 mL of ethyl acetate. The organic layer (900 μL) was dried in a SpeedVac. The dried residue was dissolved in 30% acetonitrile/water and analyzed by HPLC.

Measurement of K_M Values of S52A/ArCP, S1439A/PCP Domains and S1977A/PCP Domains at 30 °C. Reactions were performed as described above with varying concentrations of the ArCP domain or the PCP domains with 3 μM S52A, S1439A, or S1977A. Each reaction mixture was 100 μL and was incubated at 30 °C for 180 min.

Product Quantitation. Each product peak of HPLC traces was computer integrated to obtain its peak area (*S*). The integrated area (*S*) was converted into product concentration using the corresponding standard calculation curve: $(4.84 \times 10^{-6})S$ nmol for Sal-Cys, $(1.57 \times 10^{-6})S$ nmol for HPT-Cys, $(1.84 \times 10^{-6})S$ nmol for HPT-COOH or HPT, and $(4.28 \times 10^{-6})S$ nmol for HPTT-COOH or HPTT (*9*). Due to lack

of synthetic standards, the peaks of Sal-Cys-Cys, HPT-Cys-Cys, and HPTT-Cys were quantitated using the Sal-Cys, HPT-Cys, and HPTT-COOH standard curves, respectively.

RESULTS

Reconstitution of the Activity of the S52A Mutant of HMWP2 by ArCP Domains in trans. The S52A mutation eliminates the serine side chain hydroxyl in the ArCP domain (1−100) of HMWP2 that would be normally primed with phosphopantetheine, and this mutant is therefore incompetent for the aryl N-capping, chain initiation step of Ybt synthesis (*10*). We evaluated whether the Cy1 domain in the S52A mutant protein could recognize a salicyl-S-ArCP provided *in trans* by adding the pure, holo-ArCP protein fragment (1−100) (*8*) to assays containing salicylate, cysteine, YbtE, and the S52A HMWP2 mutant. If reconstitution occurs, growing Ybt chains could move first to the PCP1 domain and then to the most downstream PCP2 domain of HMWP2. Chains docked at PCP1 will have been acted upon by Cy1 to produce the hydroxyphenylthiazolinyl (HPT)-S-PCP1 acyl-enzymes while chains that have translocated down to PCP2 will have been acted upon by Cy2 for a second round of elongative condensation and heterocyclization to produce the HPTT-S-PCP2 acyl-enzyme (Figure 1B). In the normal course of events the chains would persist until transfer to the downstream HMWP1 subunit of the assembly line could occur. Absent HMWP1, we have previously shown (*9*) that the substrate cysteine also acts as a general thiolysis reagent for slow capture of the various acyl-S-enzyme intermediates, releasing the Sal-S-ArCP as Sal-Cys, the HPT-S-PCP1 intermediate as HPT-Cys, and the HPTT-S-PCP2 acyl enzyme as HPTT-Cys (Figure 2A). There is also some competing hydrolysis at PCP1 and PCP2 to yield HPT-COOH and HPTT-COOH, all of which are separable on HPLC (*9*).

Figure 2B shows the suite of products released by wild-type HMWP2 (trace a) under these assay conditons and the absence of products in the S52A mutant protein (trace b). When ArCP is added to the S52A incubations (Figure 3A), the full range of products, reflecting accumulation of all three acyl-S-enzyme intermediates, Sal-S-ArCP, HPT-S-PCP1, HPTT-S-PCP2, are again detected. We tested the concentration dependence of added ArCP and found Michaelis−Menten kinetics with a K_m of 4−5 μM for products released from PCP1 and 2.4−5 μM for the HPTT-COOH and HPTT-Cys products released from the PCP2 carrier site (Table 1). The k_{cat} (Table 2), summed over the four products of 0.07 min^{-1}, approaches the value of the wild-type HMWP2 (0.18 min^{-1}) and is limited in these assays by the rates of nonenzymatic thiolytic release of the acyl-enzymes by cysteine in solution (*10*).

The functioning *in trans* assay allowed two further assessments. First was the selectivity for the ArCP platform on which the Sal-S-pantetheinyl arm is presented to Cy1. Several other bacterial siderophore synthetases initiate chain assembly with aryl N-caps and so contain equivalent ArCP domains as chain-initiating sites for aryl acyl-S-enzyme formation, including enterobactin synthetase (ArCP in EntB) (*26*), pyochelin synthetase (ArCP in PchE) (*23*), and vibriobactin synthetase (ArCP in VibB) (*24*). The EntB ArCP, the PchE ArCP, and VibB, full length or ArCP fragment, were

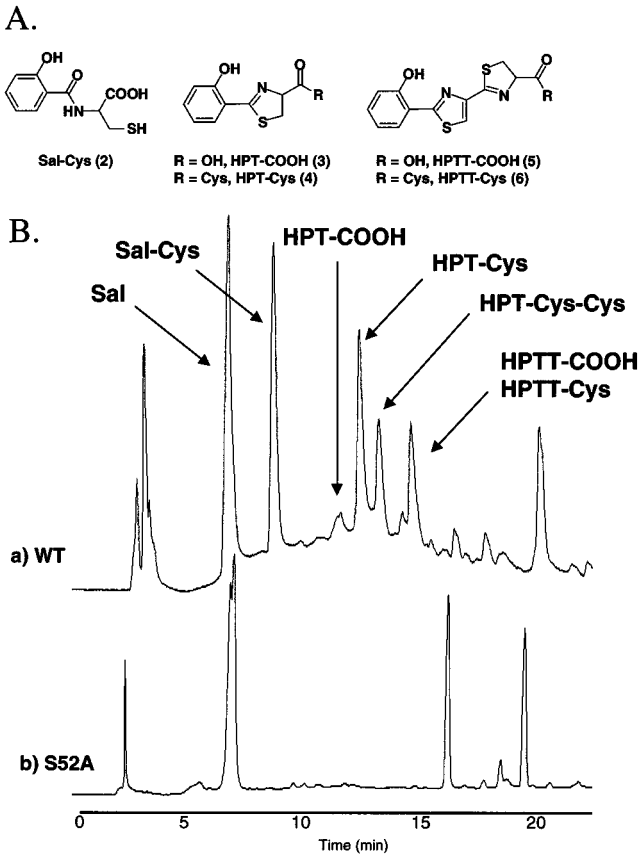

FIGURE 2: (A) Structures of the enzymatic reaction products. Abbreviations: Sal-Cys, *N*-(2-hydroxyphenyl)cysteine; HPT, 2-(2-hydroxyphenyl)thiazolinyl; HPTT, 2-(2-hydroxyphenyl)thiazolyl-2,4-thiazolinyl. Compounds **3** and **4** have spontaneously oxidized to the aromatic thiazole from the thiazoline. (B) Labeled HPLC traces of reaction products from (a) wild-type HMWP2 (3 μM) and (b) S52A (3 μM).

purified as noted in Materials and Methods, primed to their holo forms with phosphopantetheine by Sfp (*29*), salicylated using their corresponding A domain, and evaluated for utilization by the Cy1 domain in the S52A HMWP2 mutant protein. As shown in Figure 3A only the homologous Ybt ArCP produces all six products, reflecting capture of the acyl-S-HMWP2 at both the PCP1 and PCP2 sites. The EntB, PchE, and VibB proteins allow a detectable amount of Sal-Cys-Cys product, suggesting Cy1 produces a Sal-Cys-PCP1 but could not cyclize it efficiently. These results show an ArCP-Cy1 protein−protein interaction that confers specificity in the *donor* side of the first condensation/cyclization. In accord with this finding the soluble surrogate of Sal-S-ArCP, Sal-S-*N*-acetylcysteamine (Sal-S-NAC), gave no detectable product (data not shown).

The second test of Cy1 selectivity was to use 2,3-dihydroxybenzoate (DHB) in place of 2-hydroxybenzoate (salicylate) as the aryl N-capping acid substrate. We have previously determined that the YbtE enzyme will make DHB-AMP so we could now probe whether the S52A HMWP2 would accept DHB as the acyl donor in place of salicylate when presented as DHB-S-Ybt ArCP. Figure 3B shows that a new HPLC peak, identified as DHPT-Cys by mass spectrometry, the thiolysis product of the dihydroxyphenylthiazolinyl-S-PCP1 acyl enzyme, was cleanly produced, with no further elongation, and only when presented on the Ybt ArCP scaffold. Cy1 can condense and heterocy-

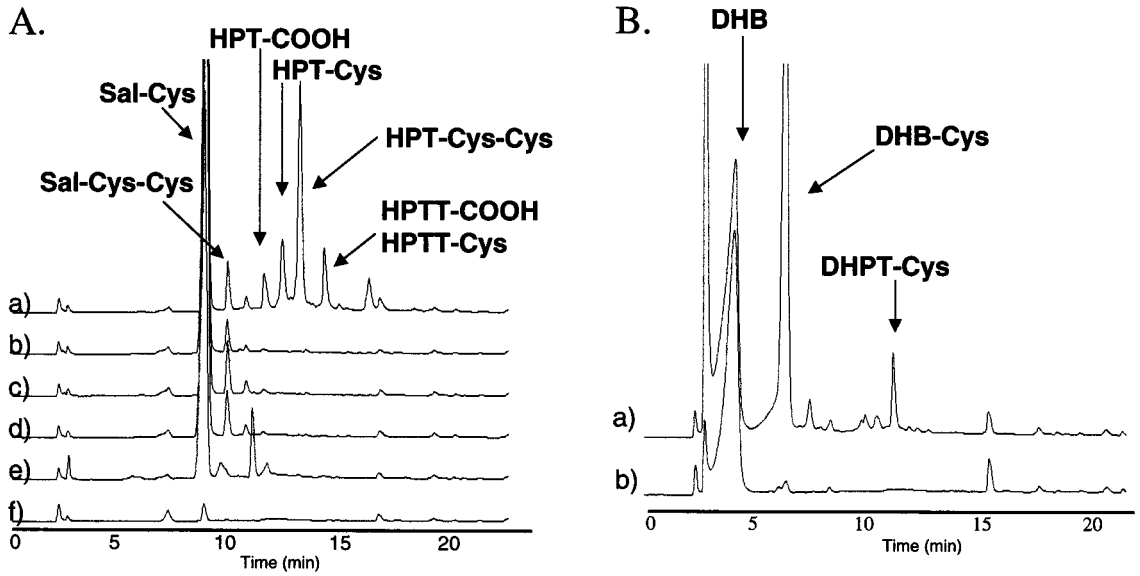

FIGURE 3: Labeled HPLC traces of reaction products from S52A (3 μM) with arylated carrier proteins. (A) Salicylate loaded onto (a) HMWP2 ArCP (1−100) (15 μM), (b) EntB (15 μM), (c) PchE (15 μM), (d) VibB full length [the bifunctional protein which consists of an isochorismate lyase (ICL) and an ArCP domain] (15 μM); (e) VibB protein fragment consisting of the ArCP domain (15 μM), and (f) control (no ArCP). (B) 2,3-Dihydroxybenzoic acid (DHB) loaded onto (a) HMWP2 ArCP (15 μM) and (b) control (no ArCP).

Table 1: K_m (μM) of HMWP2 ArCP, PCP1, and PCP2 Mutants (S52A, S1439A, and S1977A) with Carrier Proteins *in Trans*

HMWP2 species	carrier protein	HPT-COOH	HPT-Cys	HPT-Cys-Cys	HPTT-R[a]
S52A	ArCP	4 ± 3	5 ± 2	5 ± 3	2.4 ± 0.8
S1439A	PCP1		11 ± 1	>300	
S1439A	PCP3		29 ± 12	>250	
S1977A	PCP2	1.0 ± 0.5	7 ± 4	4 ± 2	8 ± 4
S1977A	PCP1	1.4 ± 0.8	2 ± 2	8 ± 9	72 ± 44
S1977A	PCP3	2.5 ± 0.7	7 ± 8	7 ± 2	66 ± 34
S1977A	SrfB1	8 ± 7	6 ± 5	29 ± 7	92 ± 65

[a] Combined production of HPTT-COOH and HPTT-Cys.

Table 2: k_{cat} (min^{-1}) of HMWP2 ArCP, PCP1, and PCP2 Mutants (S52A, S1439A, and S1977A) with Carrier Proteins *in Trans*

HMWP2 species	carrier protein	HPT-R[a]	HPT-Cys-Cys	HPTT-R[b]
WT		0.066 ± 0.004	0.080 ± 0.003	0.080 ± 0.003
S52A	ArCP	0.032 ± 0.004	0.006 ± 0.001	0.032 ± 0.003
S1439A	PCP1	0.0066 ± 0.0003	0.005 ± 0.005	
S1439A	PCP3	0.0020 ± 0.0004	0.002 ± 0.001	
S1977A	PCP2	0.018 ± 0.002	0.016 ± 0.003	0.036 ± 0.006
S1977A	PCP1	0.0057 ± 0.0008	0.0016 ± 0.0007	0.003 ± 0.004
S1977A	PCP3	0.0032 ± 0.0003	0.0027 ± 0.0003	0.007 ± 0.003
S1977A	SrfB1	0.004 ± 0.003	0.0043 ± 0.0006	0.008 ± 0.004

[a] Combined production of HPT-COOH and HPT-Cys. [b] Combined production of HPTT-COOH and HPTT-Cys.

Table 3: k_{cat}/K_m (μM^{-1} min^{-1}) of HMWP2 ArCP, PCP1, and PCP2 Mutants (S52A, S1439A, and S1977A) with Carrier Proteins *in Trans*

HMWP2 species	carrier protein	HPT-R[a]	HPT-Cys-Cys	HPTT-R[b]
1−1382	PCP1	0.00098[c]		
1−1382	1383−2035 (S1977A)	0.0085[c]		
S52A	ArCP	0.0067	0.0011	0.013
S1439A	PCP1	0.00059	0.000017	
S1439A	PCP3	0.000069	0.000008	
S1977A	PCP2	0.0045	0.004	0.0045
S1977A	PCP1	0.0038	0.0002	0.000042
S1977A	PCP3	0.000064	0.00036	0.00011
S1977A	SrfB1	0.00057	0.00015	0.000087

[a] Combined production of HPT-COOH and HPT-Cys. [b] Combined production of HPTT-COOH and HPTT-Cys. [c] Previously reported (*9*).

clize the DHB group, but the DHPT acyl chain is apparently not a substrate for the next Cy domain, Cy2, and so is not detectably elongated and translocated further downstream to the PCP2 way station.

S1439A Mutant of HMWP2: Evaluation of in trans Substitution for the Inactive PCP1 Domain as a Test for Cy1 Downstream Acyl-Enzyme Substrate and Cy2 Upstream Acyl-Enzyme Substrate. The S1439A mutant of HMWP2 disables the second carrier protein domain, PCP1, consigning it to an inactive apo form, and opens the possibility of replacing the inactive *in cis* PCP1 with holo-PCP1 and other holo-PCPs *in trans*. The holo-PCP1 domain presents the downstream nucleophilic substrate Cys-S-PCP1 in the Cy1-mediated condensation and heterocyclization but also then presents the upstream electrophilic substrate (Figure 4A), HPT-S-PCP1, in the Cy2-mediated condensation. Thus assaying S1439A reconstitution *in trans* permits interrogation of recognition and selectivity of both Cy1 and Cy2 catalytic domains.

The first experiment was reconstitution with the wild-type Ybt holo-PCP1 fragment (residues 1383−1491) with YbtE, Sal, Cys, ATP with S1439A at 3 μM, and the PCP at a 3.3-fold molar excess, 10 μM. Figure 4B shows the HPLC trace from such an incubation where the telltale products are HPT-Cys and HPT-Cys-Cys. The HPT-Cys, from thiolysis of HPT-S-enzyme, indicates that Cy1 transfers the salicyl chain from Sal-S-ArCP out to the Cys-S-PCP1 *in trans* and cyclizes it to HPT-S-PCP1. The HPT-Cys-Cys product indicates that the *in trans* HPT-S-PCP1 fragment can serve as a donor to Cy2, and HPT chain transfer to Cys-S-PCP2 in HMWP2

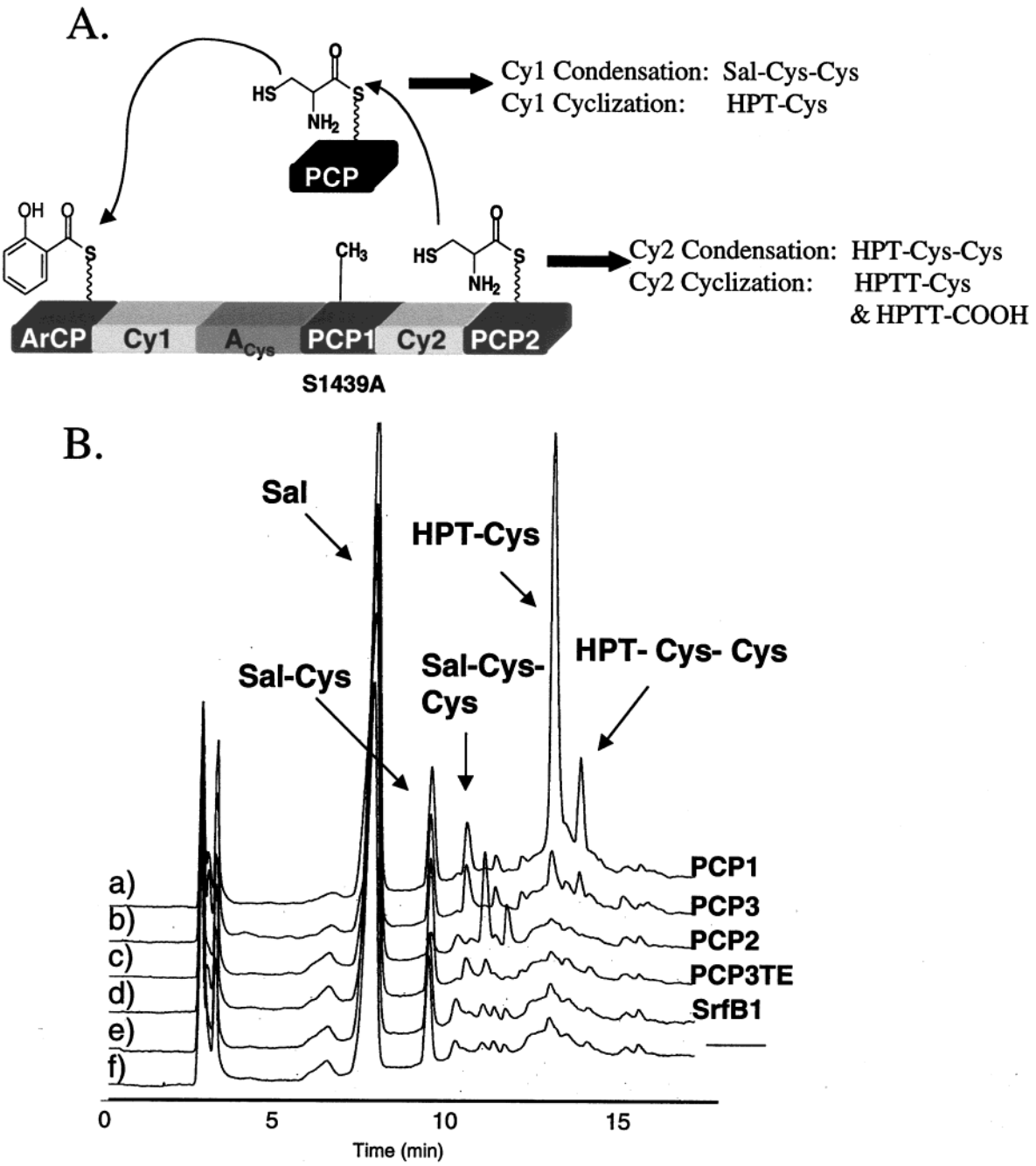

FIGURE 4: (A) Schematic of possible products released by a reaction of S1439A and an *in trans* PCP domain. S1439A allows exploration of both Cy1 and Cy2: the PCP domain *in trans* serves as the protein scaffold presenting the downstream acceptor substrate in the Cy1-mediated reactions [assayed by the release of Sal-Cys-Cys (condensation) and HPT-Cys (cyclization)] and also serves as the protein scaffold presenting the upstream donor substrate in the Cy2-mediated reactions [assayed by the release of HPT-Cys-Cys (condensation), HPTT-COOH, and HPTT-Cys (cyclization)]. (B) Labeled HPLC traces of reaction products from S1439A (3 μM) and (a) PCP1 (10 μM), (b) PCP3 (10 μM), (c) PCP2 (10 μM), (d) PCP3TE (10 μM), (e) SrfB1 (10 μM), and (f) control (no PCP).

yields the HPT-Cys-S-PCP2 acyl enzyme. This is thiolyzed faster than it is cyclized, yielding HPT-Cys-Cys as the sole product, rather than as a minor product and HPTT-Cys as the major product, as happens in wild-type HMWP2. The kinetics of product formation from the PCP1 site (HPT-Cys) and from the PCP2 site (HPT-Cys-Cys) reveal very different K_m values for the *in trans* holo-PCP1 fragment. It shows a K_m of 11 μM for making the acyl protein HPT-S-PCP1 but >300 μM for making the HPT-Cys-S-PCP2 acyl protein (Table 1). Thus Cy1 recognizes the PCP1 presented *in trans* at least 30-fold better than does Cy2 for the condensation/cyclization versus condensation chemistry. The estimated k_{cat} of 0.007 min^{-1} (for HPT-Cys and HPT-COOH) versus 0.005 min^{-1} (for HPT-Cys-Cys) indicates that about 40% of the HPT-S-PCP1 fragments, produced by Cy1, are utilized by Cy2 in competition with chain-releasing thiolysis.

Three other holo-PCP fragments were analogously provided *in trans* to S1439A HMWP2, the PCP2 from HMWP2, the PCP3 from the HMWP1 subunit (*7*, *8*), the surfactin synthetase SrfB1 PCP (*25*), and the PCP3-TE bidomain (*22*). As shown in Figure 4B, the holo-PCP3 shows a small amount of activity for HPT-Cys and HPT-Cys-Cys formation, but PCP2 does not. The PCP3-TE double domain is at the noise level as is the SrfB1 PCP. Clearly, there is selection for the

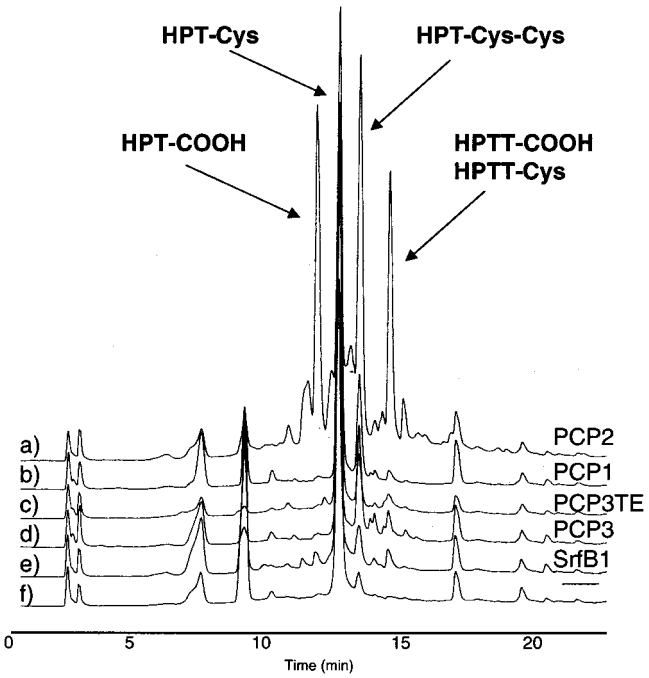

FIGURE 5: Labeled HPLC traces of reaction products from S1977A (3 μM) and (a) PCP2 (10 μM), (b) PCP1 (10 μM), (c) PCP3TE (10 μM), (d) PCP3 (10 μM), (e) SrfB1 (10 μM), and (f) control (no PCP).

cognate PCP1 platform.

S1977A Mutant of HMWP2: Evaluation of in trans Substitution of Inactive PCP2 as a Test for Cy2 Downstream Recognition of Aminoacyl-S-PCPs. The S1977A mutant of HMWP2 disables the third and last carrier protein domain, PCP2, and allows assay of Cy2 function for its downstream substrate carrier protein domain, by PCP domains supplied externally and aminoacylated with cysteine by the HMWP2 A domain. As shown in Figure 5, the exogenous addition of the cognate PCP2 leads to formation of the two bis-heterocyclic products, HPTT-Cys and HPTT-COOH, that indicate catalytic condensation and coupled heterocyclization by the Cy2 domain. Lesser, but clearly detectable, amounts of the bis-heterocyclic products are enabled also by PCP1 and PCP3 and even the SrfB1 holo-PCP, indicating that they are all cysteinylated by the A domain and then accepted by the Cy2 domain as downstream nucleophilic substrates for coupling with the HPT-S-PCP1 acyl-enzyme upstream of Cy2. The Cy2 domain seems less discriminating about the presenting PCP scaffold in the downstream versus the upstream position, comparing the behavior of the S1977A versus the S1439A mutant HMWP2s.

Kinetic analysis (Table 1) of the two sets of products, HPT-Cys and HPT-COOH derived from Cy1 versus HPT-Cys-Cys and HPTT-Cys/HPTT-COOH derived from Cy2, shows that the cognate PCP2 fragment *in trans* has about equivalent K_m values, 1−8 μM, for both Cy domains. By contrast, when PCP1 is used as the downstream acceptor for Cy2 in the *in trans* configuration, the low K_m for products generated at and released from PCP1 holds, but now the K_m for HPTT-Cys/HPTT-COOH is up to >70 μM, showing that the condensation to HPT-Cys-S-PCP1 proceeds well but the subsequent heterocyclization is disfavored. This is accentuated further with the SrfB1 holo-PCP where the K_m for Cy2 action is up to >90 μM. The noncognate PCP scaffolds are selectively disadvantaged in the Cy2 heterocyclization step to form the second of the tandem thiazoline rings. In terms of k_{cat} values, Table 2 indicates that about 75% of the chain elongation flux goes down to the *in trans* PCP when Cy2 is given the cognate PCP1 for reconstitution. This flux is about 50% to the downstream PCP when Cy2 is presented with PCP1 and the k_{cat} has dropped 10-fold. With the SrfB1 PCP, the flux is back up to 75% through Cy2 and the rate is down 4-fold.

Truncations of HMWP2 To Remove PCP1 and in trans Reconstitution. The above studies with inactive point mutants of the ArCP, the PCP1 and PCP2 domains, leave the inactive carrier protein domain still resident in the assembly line. An alternative approach to evaluate Cy1 recognition of the PCP domains is to remove one of the carrier protein domains by truncation and supply it back *in trans*. Truncation of the HMWP2 subunit to remove PCP1 produced the 1−1382 fragment of HMWP2 that terminates at the end of the predicted Cy1 downstream border. On addition of PCP1, PCP2, or PCP3 under conditions of apo to holo conversion and cysteinylation, HPT-Cys or Sal-Cys-Cys production could only be detected with PCP1 (data not shown). There was no signal over noise with the PCP2 or PCP3 fragments under the same conditions. The reconstitution efficiency of 1−1382 was compared upon addition of either the 100-residue PCP1 or the full second half of HMWP2 (1383−2035) in the context of an S1977A mutant to halt catalysis at Cy1. Thus, the transfer to PCP1 in the two *in trans* contexts would be comparable, by k_{cat} and K_m analysis (Table 2). An 8-fold higher catalytic efficiency, k_{cat}/K_m, was seen with the 1383−2035 protein, suggesting either recognition of the Cy1 domain for its downstream protein substrate beyond the PCP1 domain or perhaps a higher fraction of native fold in PCP1 embedded in its PCP1-Cy2-PCP2 framework. In addition, a >70-fold higher catalytic efficiency was seen with the native PCP assayed *in trans* with either S1439A or S1977A full-length HMWP2.

DISCUSSION

The HMWP2 subunit of yersinibactin synthetase, a six domain 230 kDa protein, has proven to be a tractable system for analysis of the catalytic and carrier selectivity features of a nonribosomal peptide synthetase assembly line. With its capacity to effect a tandem bis heterocyclization of two cysteine substrate molecules, this multidomain enzyme carries out intriguing chemistry as it creates the 4,2-tandem five-ring heterocycles thought to be part of the coordination site for ferric ion in the Ybt siderophore (Figure 1). We have previously reported heterologous expression of full-length HMWP2 in *E. coli* and validation of the predictions for the single Cys-activating A domain, the three carrier protein domains that tether salicylate, Cys, and Cys, and the two condensation/heterocyclization domains, Cy1 and Cy2, that function as an assembly line to condense the three substrate-S-enzyme intermediates, heterocyclize to the two thiazolines, and translocate the HPTT chain to PCP2, ready for transfer to the nine domain HMWP1 subunit (*8*, *10*, *22*).

In this work we have begun to assess selectivity of Cy1 and Cy2 for both the upstream acyl-S-ArCP/PCP and the downstream Cys-S-PCP used as cosubstrates. We have analyzed both for acyl chain selection and for recognition

of the 10 kDa carrier protein domains that act as a scaffold to present the acyl and aminoacyl substrate chains to the Cy domains. Such an assessment for composite acyl-S-PCP specificity requires that one be able to assay each of the embedded Cy1 and Cy2 domains and alter the ArCP/PCP scaffold on which the salicyl- or cysteinyl- or elongating HPT-S-pantetheinyl groups is presented.

The route we utilize here for such assays is to inactivate each of the three carrier protein domains in HMWP2, one at a time, using the S to A mutants that keep the carrier protein domains in their inactive, apo forms that cannot be phosphopantetheinylated. Acyl-S- and aminoacyl-S-ArCP and PCP protein fragments are then used to complement these mutations *in trans*. We employed several ArCP domains (Ybt 1−100, PchE, EntB, and VibB) and PCP domains (PCP1, PCP2, PCP3, PCP3TE, and SrfB1) that were available to us from prior NRPS investigations (*8*, *10*, *23*, *25*, *30*). Finally, we detected and quantified elongating chains that have arrived at ArCP, PCP1, and PCP2 carrier sites in HMWP2 via previously developed product release assays (*9*, *10*).

In these studies the S52A mutant permits us to probe the upstream acyl donor selectivity of Cy1 for the salicyl-ArCP while the S1439A PCP1 mutant allows the complementary evaluation of Cy1 for the downstream cysteinyl-S-PCP as attacking nucleophile in the condensation and subsequent competence for cyclodehydration to the thiazolinyl-S-PCP product. Since the holo-PCP1 domain is also the staging site for presentation of the HPT chain as upstream donor to Cy2, the S1439A form of HMWP2 also permits evaluation of this feature of Cy2 recognition. The third carrier protein domain mutant S1977A allows evaluation of the last feature, the recognition of Cy2 for the downstream Cys-S-PCP as the attacking nucleophile on the way to the second heterocyclization of the growing NRP chain.

The first general finding with the ArCP, PCP1, and PCP2 S to A point mutants is that the cognate holo-ArCP, PCP1, and PCP2 domains when added *in trans* can reconstitute the specific partial reactions assayed in the wild-type HMWP2 subunit. The presence of the inactive apo carrier protein domains *in cis* does not block Cy domain recognition independent of the three-dimensional architecture of the six domain HMWP2.

With the S52A HMWP2 mutant, chain initiation is restored by the homologous Ybt HMWP2 ArCP fragment (1−100) *in trans*, with a low micromolar K_m and the same rate as wild-type enzyme, where HPT-Cys formation is limited by the rate of cysteine thiolysis. All of the heterologous holo-ArCPs could be acylated with salicylate, but the salicyl group was not transferred to the cysteinyl-S-PCP1 acceptor by the Cy1 domain, showing a clear recognition of the presenting ArCP scaffold for salicyl transfer. Dihydroxybenzoyl transfer was also seen when it was presented as DHB-S-(HMWP2 ArCP) *in trans*. Analogously, the S1439A mutant could be reconstituted with cysteinyl-S-PCP1 presented *in trans* and assaying for HPT-Cys release. The cognate PCP1 was more effective than the PCP2 or PCP3 carrier protein domains for Cy1 recognition as the nucleophilic partner in the condensation with salicyl-S-ArCP, again consistent with recognition of both the aminoacyl moiety and the PCP scaffold. The homology of PCP2 and PCP3 to PCP1 is 18% and 34%, respectively.

To evaluate the comparable specificity of the Cy2 domain of the HMWP2 subunit of Ybt synthetase for the upstream HPT-S-PCP donor and the downstream cysteinyl-S-PCP acceptor, the S1439A mutant monitored the upstream and the S1977A mutant monitored the downstream substrate recognition. With the holo-PCP1 *in trans* with S1439A, HPLC analysis showed formation of HPT-Cys-Cys but no HPTT-Cys. The HPT-Cys-Cys arises from thiolysis of HPT-Cys-S−PCP2 (*9*) before cyclization of the Cys side chain to give HPTT-S-PCP2. Thus the *in trans* holo-PCP1 fragment can act as an HPT-S-PCP1 donor for Cy2 and engage in amide bond formation but not the second heterocyclization.

The second general result is this uncoupling of the condensing functions of both Cy1 and Cy2 from subsequent cyclodehydrative and thiazoline formation. Perturbation of the upstream carrier protein scaffold (e.g., ArCPs with Cy1, PCPs with Cy2) leads to large decreases in the catalytic efficiency of heterocyclization versus condensation. Specifically, the isolation of the Sal-Cys-Cys product, from the action of the Cy1 domain, and the HPT-Cys-Cys product, from the action of the Cy2 domain, proves that both Cy1 and Cy2 are capable of acting as C domains. This is consistent with a mechanism where Cy domains first condense the downstream Cys-S-PCP onto the upstream acyl-S-PCP and then separately heterocyclize the acyl-Cys to thiazoline. It does not resolve the mechanistic issue of whether the amine of the Cys-S-PCP domain or the thiol of the Cys-S-PCP domain substrate is the attacking nucleophile since the net formation of the amide linkage could either be direct or after thioester formation and a rapid S to N acyl shift. Previously, we had attempted to achieve this uncoupling of condensation from cyclodehydration in Cy1 domain action by mutating the signature motif D-x-x-x-x-D-x-x-S to the core sequence characteristic of C domains, H-H-x-x-x-D-G-x-S (*10*) without success. Whether this uncoupling is due to misaligned architectural elements between the *in trans* ArCP/PCP and Cy domains and/or is a kinetic consequence of reduced efficiency/binding for subsequent cyclization by the *in trans* presentation is not yet clear.

On the whole, it appears that the point mutant S to A and *in trans* presentation of acylated/aminoacylated carrier proteins was effective, suggesting minimal disruption of interdomain chain transfer steps in HMWP2. In addition, it is interesting to note that PCP3 (a domain that is naturally recognized *in trans* by the A domain of HMWP2) was the only heterologous PCP that could be used *in trans* with S1439A, indicating that it may present a more universal structure than the PCP domains that are naturally *in cis* to their A domains. Sequence alignment does not lead to any obvious conclusions since PCP3 is less homologous to PCP1 than PCP2. However, this result lends support to the hypothesis that a universal PCP domain could be created which would allow for a range of new products to be created through combinatorial biosynthesis.

In general, this study of the S52A, S1439A, and S1977A mutants has shown that Cy domains, a variant of the more common condensation (C) domains, are more specific toward the upstream donor protein than the downstream acceptor protein. These experiments directly examined the inherent selectivity of the Cy domains for the donor or acceptor protein scaffolds since the Sal-S-P-pant- or Cys-S-P-pant-substrate portion remained the same while only the ArCPs

or PCPs were changed. Specifically, Cy1 was intolerant to variation in its upstream protein donor (ArCP) but was able to accept an alternative acyl monomer (DHB) producing DHPT-Cys. Both Cy1 and Cy2 are less selective toward their downstream (acceptor) protein scaffolds as indicated by the products observed when alternate PCP domains were used *in trans* with both S1439A (Figure 4B) and S1977A (Figure 5). However, Cy2 showed specific selectivity for its upstream protein donor; reactions with S1439A and alternative PCPs (PCP1 and PCP3) produced HPT-Cys-Cys but no HPTT-Cys, thus uncoupling the condensative and cyclodehydrative functions of the Cy domain.

Results from previous studies (*13*) in which aminoacyl-*N*-acetylcysteamine thioesters (aminoacyl-SNACs) were used to probe the specificity of C domains showed a greater specificity toward their downstream acceptor aminoacyl substrates (recognizing both the side chain size and the L- versus D-configuration) than toward their upstream donor substrates in which only the L- versus D-configurations were discriminated. These results agree with those using aminoacyl-CoAs (*14*). Taken together with our results, a prediction could be made for a biased binding pocket for C/Cy domains: upstream the P-pant binds strongly while downstream the amino acid contributes most of the binding energy.

ACKNOWLEDGMENT

We thank Huawei Chen, David E. Ehmann, Hiten M. Patel, and Zucai Suo for kindly providing salicyl-SNAC; SrfB1, EntB ArCP, and EntE; PchE ArCP and PchD; and HMWP1 fragment PCP3TE, respectively. We are grateful to Thomas A. Keating for providing VibB and VibE, for critical reading of the manuscript, and for many helpful discussions.

REFERENCES

1. Pelludat, C., Rakin, A., Jacobl, C. A., Schubert, S., and Heesemann, J. (1998) *J. Bacteriol. 180*, 538−546.
2. Geoffroy, V. A., Fetherston, J. D., and Perry, R. D. (2000) *Infect. Immun. 68*, 4452−4461.
3. Perry, R. D., Balbo, P. B., Jones, H. A., Fetherston, J. D., and DeMoll, E. (1999) *Microbiology 145*, 1181−1190.
4. Marahiel, M. A. (1997) *Chem. Biol. 4*, 561−567.
5. Konz, D., and Marahiel, M. A. (1999) *Chem. Biol. 6*, R39−48.
6. Carreras, C. W., Pieper, R., and Khosla, C. (1997) *Top. Curr. Chem. 188*, 86−126.
7. Gehring, A. M., DeMoll, E., Fetherston, J. D., Mori, I., Mayhew, G. F., Blattner, F. R., Walsh, C. T., and Perry, R. D. (1998) *Chem. Biol. 5*, 573−586.
8. Gehring, A. M., Mori, I., Perry, R. D., and Walsh, C. T. (1998) *Biochemistry 37*, 11637−11650.
9. Suo, Z., Walsh, C. T., and Miller, D. A. (1999) *Biochemistry 38*, 14023−14035.
10. Keating, T. A., Miller, D. A., and Walsh, C. T. (2000) *Biochemistry 39*, 4729−4739.
11. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) *Chem. Biol. 3*, 923−936.
12. Stachelhaus, T., Huser, A., and Marahiel, M. A. (1996) *Chem. Biol. 3*, 913−921.
13. Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000) *Chem. Biol. 7*, 765−772.
14. Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999) *Science 284*, 486−489.
15. Holak, T. A., Nilges, M., Prestegard, J. H., Gronenborn, A. M., and Clore, G. M. (1988) *Eur. J. Biochem. 175*, 9−15.
16. Crump, M. P., Crosby, J., Dempsey, C. E., Parkinson, J. A., Murray, M., Hopwood, D. A., and Simpson, T. J. (1997) *Biochemistry 36*, 6000−6008.
17. Weber, T., Baumgartner, R., Renner, C., Marahiel, M. A., and Holak, T. A. (2000) *Struct. Fold. Des. 8*, 407−418.
18. Stachelhaus, T., Mootz, H. D., Bergendahl, V., and Marahiel, M. A. (1998) *J. Biol. Chem. 273*, 22773−22781.
19. Keating, T. A., and Walsh, C. T. (1999) *Curr. Opin. Chem. Biol. 3*, 598−606.
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
21. Suo, Z., Chen, H., and Walsh, C. T. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 14188−14193.
22. Keating, T. A., Suo, Z., Ehmann, D. E., and Walsh, C. T. (2000) *Biochemistry 39*, 2297−306.
23. Quadri, L. E., Keating, T. A., Patel, H. M., and Walsh, C. T. (1999) *Biochemistry 38*, 14941−14954.
24. Keating, T. A., Marshall, C. G., and Walsh, C. T. (2000) *Biochemistry 39*, 15522−15530.
25. Weinreb, P. H., Quadri, L. E., Walsh, C. T., and Zuber, P. (1998) *Biochemistry 37*, 1575−1584.
26. Gehring, A. M., Bradley, K. A., and Walsh, C. T. (1997) *Biochemistry 36*, 8495−8503.
27. Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) *Biochemistry 37*, 1585−1595.
28. Ehmann, D. E. (2000) in *Biochemistry and Molecular Pharmacology*, Harvard University, Boston, MA.
29. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) *Chem. Biol. 3*, 923−936.
30. Keating, T. A., Marshall, C. G., and Walsh, C. T. (2000) *Biochemistry 39*, 15513−15521.

BI002905V