C-Methyltransferase and Cyclization Domain Activity at the Intraprotein PK/NRP Switch Point of Yersiniabactin Synthetase

Deborah Ann Miller,‡ Christopher T. Walsh,* and Lusong Luo‡

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

Received June 12, 2001
Revised Manuscript Received July 20, 2001

Some of the most therapeutically interesting natural products are mixed polyketide/nonribosomal peptide molecules, including bleomycin and epothilones (anticancer agents) and FK506 and rapamycin (immunosuppressive drugs).[1−3] These enzymatically assembled molecules are encoded by multimodular protein arrays that, through their chain-ordering, determine the structure of the natural product elongated and released. Among the fundamental questions in such hybrid nonribosomal peptide synthetase (NRPS)/polyketide synthase (PKS) and PKS/NRPS assembly lines are how elongating chains are switched from amide synthetase chemistry (by condensation domains, NRPS logic) to Claisen condensation chemistry (by ketosynthase domains, PKS logic) and back as the chains move through switch-point module interfaces. This study describes the first in vitro assay of one such switch point that also involves a concomitant C-methylation of the growing acyl chain.

The bacteria that are the causative agents of the plague, *Yersinia pestis*, elaborate an iron-chelating virulence factor, yersiniabactin, when in iron-deficient microenvironments. Yersiniabactin (Ybt, **1**)[4] (Figure 1) is a nonribosomal peptide in which a salicyl N-cap has been added to three cysteines that have been cyclized to thiazolines, then C-methylated and reduced to adjust the oxidation level. Between the second and third five-membered rings is a four-carbon fragment derived from two C-methylations and one malonyl unit.[5−7] Thus, Ybt is presumably derived from an NRP/PK/NRP hybrid.

This prediction has been validated by the sequencing of genes for two high-molecular weight proteins, HMWP1 and HMWP2, from a *Y. pestis* high-pathogenicity island.[8] The HMWP2/HMWP1 subunits form a 15-domain assembly line, with two NRPS modules in the 230 kDa HMWP2[9] and then a PKS module and an NRPS module in the 350 kDa HMWP1 subunit. The first NRPS/PKS module interface is at the junction of HMWP2/HMWP1 where the elongating chain moves across the interprotein boundary. The second switch point, now a PKS/NRPS intraprotein switch, is where the chain passes from the five-domain PKS module of HMWP1 to the four-domain NRPS module. The 3-(hydrophenylthiazolinyl-thiazolidinyl)-3-hydroxy-2,2-dimethyl acyl chain **2** tethered in thioester linkage on the acyl carrier protein (ACP) domain of the PKS module is the proposed donor,

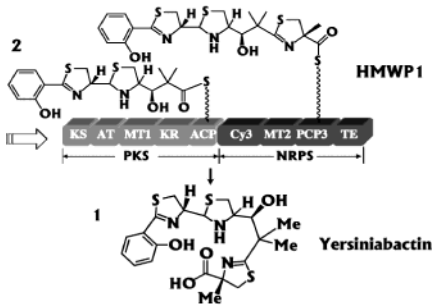

Figure 1. Structure of Ybt and HMWP1 acyl intermediates [drawn covalently tethered to the appropriate carrier protein domains through a 4′-phosphopantetheine (wavy line)]. KS = ketosynthase; AT = acyltransferase; MT = methyltransferase; KR = ketoreductase; ACP = acyl carrier protein; Cy = cyclization; PCP = peptidyl carrier protein; TE = thioesterase.

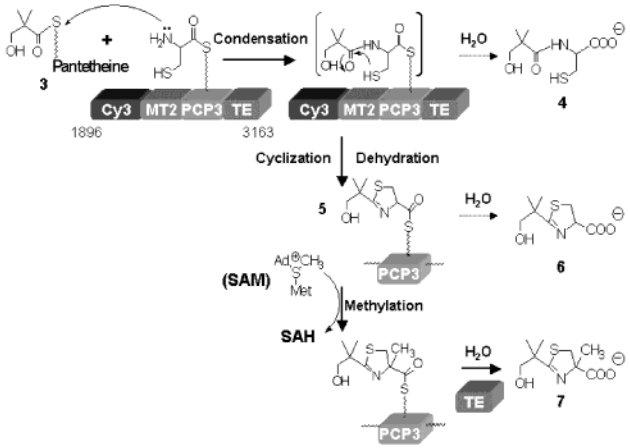

Figure 2. Condensation, cyclization, and methylation reactions catalyzed by HMWP1 NRPS module using DHPP **3** as donor substrate surrogate.

and the cysteinyl-S-peptidyl carrier protein (PCP_3) domain in the downstream NRPS module is the proposed acceptor at the PKS/NRPS switch point. This should yield the desmethyl Ybt chain that subsequently undergoes C-methylation by the methyl transferase domain (MT_2) and hydrolytic release by the thioesterase (TE) to yield Ybt.

To evaluate the catalytic capacity of the HMWP1 NRPS four-domain module to effect the PK-to-NRP chain-switching step, we have purified the 144 kDa fragment, residues 1896-3163, after expression in *Escherichia coli*, primed it with phosphopantetheine on the apo PCP_3 domain, using Sfp,[10] a member of the phosphopantetheine transferase (PPTase) family, and CoASH as substrate, to create the holo form competent to be cysteinylated on the pantetheine terminal SH by the adenylation domain of partner subunit HMWP2 (fragment 1-1382), using ATP and [^{35}S]-cysteine.[11] Since the complex natural acyl donor, an acyl-S-ACP domain embedded within the large PKS module **2**, was not available, we have prepared the simple 2,2-dimethyl-3-hydroxypropionyl-S-pantetheinyl thioester (DHPP) **3** as a soluble surrogate substrate (Figure 2), which simplifies both the acyl chain donor and the pantetheinyl-ACP scaffold.[12] When **3** was incubated with [^{35}S]-Cys-S-[1896-3163] fragment of HMWP1, the released products could be analyzed by radioactive HPLC analysis.[13] Two new peaks were detected dependent on **3** and cysteinylation of

‡ These authors contributed equally to this work.

(1) Du, L.; Sanchez, C.; Shen, B. *Metab. Eng.* **2001**, *3*, 78−95.
(2) Du, L.; Shen, B. *Curr. Opin. Drug Discovery Dev.* **2001**, *4*, 215−228.
(3) Keating, T. A.; Walsh, C. T. *Curr. Opin. Biochemistry* **1999**, *3*, 598−606.
(4) Ino, A.; Murabayashi, A. *Tetrahedron* **2001**, *57*, 1897−1902.
(5) Gehring, A. M.; DeMoll, E.; Fetherston, J. D.; Mori, I.; Mayhew, G. F.; Blattner, F. R.; Walsh, C. T.; Perry, R. D. *Chem. Biol.* **1998**, *5*, 573−586.
(6) Gehring, A. M.; Mori, I.; Perry, R. D.; Walsh, C. T. *Biochemistry* **1998**, *37*, 11637−11650.
(7) Perry, R. D.; Balbo, P. B.; Jones, H. A.; Fetherston, J. D.; DeMoll, E. *Microbiology* **1999**, *145*, 1181−1190.
(8) Pelludat, C.; Rakin, A.; Jacobl, C. A.; Schubert, S.; Heesemann, J. *J. Bacteriol.* **1998**, *180*, 538−546.
(9) Keating, T. A.; Miller, D. A.; Walsh, C. T. *Biochemistry* **2000**, *39*, 4729−4739.
(10) Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M.; Marahiel, M. A.; Reid, R.; Khosla, C.; Walsh, C. T. *Chem. Biol.* **1996**, *3*, 923−936.
(11) Suo, Z.; Walsh, C. T.; Miller, D. A. *Chem. Biol.* **1999**, *38*, 14023−14035.

10.1021/ja016398w CCC: $20.00
Published on Web 08/03/2001

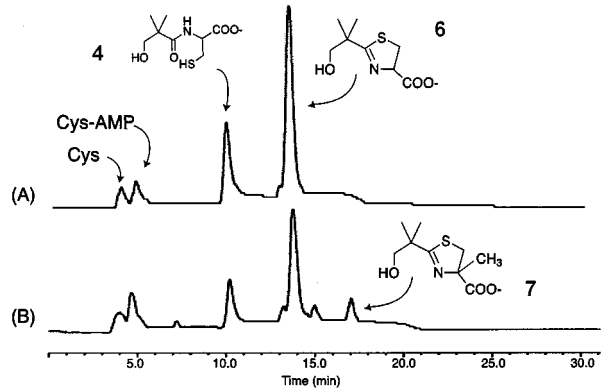

Figure 3. Radioactive HPLC traces of reaction products from [^{35}S]-cysteine and (A) HMWP1 1896-3163 + **3**; (B) HMWP1 1896-3163 + **3** + SAM.[13]

the 1896-3163 NRPS module (Figure 3A). Repetition of the incubation with unlabeled substrates followed by MS analysis showed the peak at 10 min to contain the mass (221) anticipated for the acyclic amide **4**, indicating the Cy domain of the NRPS module had functioned for amide bond condensation. The second peak (13.5 min) contained a molecular ion 18 mass units lower at 203, consistent with the cyclization/dehydration to the thiazoline-carboxylate **6**. This is the second function expected of Cy domains, and it was previously validated by us for Cy1 and Cy2 domains of the HMWP2 subunit of this siderophore synthetase assembly line.[14]

To address the prediction that the MT_2 domain in the NRPS module of HMWP1 is in fact a C-methyl transferase, the above incubations were repeated in the presence of S-adenosylmethionine (SAM) and a new radioactive peak (17 min) was detected by HPLC (Figure 3B), consistent with methyl transfer to the DHP-thiazolinyl-S-PCP_3 **5** acyl enzyme intermediate. By MS analysis in an unlabeled reaction mixture this new peak had a molecular mass of 217, as expected for the addition of 14 mass units (CH_2) to the 203 for **6**, and consistent with methylation at C2 of the DHP-thiazolinyl-S-PCP_3 to yield the C-methyl intermediate, which on thioesterase-mediated hydrolysis gives the observed product **7**, a faithful mimic of the last portion of full length Ybt.

Kinetic studies were performed to obtain a K_m for **3** (> 60 mM) and for SAM (100 μM) validating saturation behavior. Both kinetic studies[15] gave a k_{cat} of 0.01 min^{-1}. The model acyl donor, **3**, has both a minimal acyl chain and a simplified pantetheinyl-ACP scaffold. To determine if protein−protein interaction between the ACP and Cy domains would increase the reaction rate, we synthesized DHP-S-CoA,[16] which could be loaded via the PPTase Sfp onto the upstream apo-ACP[17] yielding DHP-S-ACP.[18] However, the rate of **7** production in a reaction containing DHP-S-ACP, HMWP1 NRPS fragment 1896-3163, HMWP2 1-1382, cysteine, SAM, NADPH, and ATP was the same 0.01 min^{-1}, while the K_m for the new donor substrate, DHP-S-ACP, decreased relative to that of **3** (approximately 400-fold) to ca. 150 μM. The reaction rate may be limited by TE hydrolysis, the Cy and MT domain recognition for the acyl chain, or the absence of required upstream protein−protein interactions. Nonetheless, considering the minimal structure of the surrogate substrate and the complexity of the reactions, the enzyme recognition of **3** followed by condensation, cyclization, and methylation of the downstream PCP_3 loaded-cysteine is quite remarkable. These experiments set the stage for full enzymatic reconstitution of Ybt.

These enzymatic studies utilize a model donor substrate, **3**, and the natural acceptor substrate, cysteinyl-S-NRPS module, to reconstitute the PKS/NRPS interface in the hybrid assembly line of yersiniabactin synthetase. Previous experiments have validated that the PCP_3 and the TE domain[5,17] of the four-domain NRPS module in the 1896-3163 enzyme fragment were active, and the current experiments validate that Cy3 and MT_2 are also functional for the predicted condensation/cyclization and C-methylation functions. Most noteworthy are the abilities of Cy3 to take the 2, 2-dimethyl-3-hydroxy-propionyl acyl chain presented on the pantetheinyl tether and of the MT domain to carry out the C-methylation. Presumably the MT domain removes the C2-H of the thiazolinyl-S-PCP_3 intermediate to form a delocalized carbanion that attacks the electrophilic methyl sulfonium group of SAM also bound in the MT active site. An analogous MT domain is presumed to effect the two C-methylations upstream in the PKS module as well as in the bleomycin synthetase and lovastatin synthase chain elongations.[19,20] Now that this C-MT domain can be assayed it will be of interest to determine if it is portable to other contexts for combinatorial biosynthesis.

Acknowledgment. This work has been supported by the National Institutes of Health Grant GM20011 to C.T.W. We thank Rahul Kohli, Hiten Patel, and Huawei Chen for technical assistance.

JA016398W

(12) 2,2-Dimethyl-3-hydroxypropionic acid (DHP) was obtained by hydrolysis of methyl 2,2-dimethyl-3-hydroxypropionate with KOH. To a stirred solution of DHP (118 mg) in anhydrous THF (10 mL) at 25 °C, ethyl chloroformate (220 μL) and triethylamine (300 μL) were added. Following removal of the white precipitate, the resulting clear solution was then added dropwise to a solution of pantetheine (278 mg, obtained by reduction of pantethine using 2 equiv of tris-(2-carboxyethyl)phosphine) in 10 mL of 100 mM K^+HEPES buffer (pH 7.5). The reaction was stirred at 25 °C overnight. The crude product was purified using preparative HPLC to give 91 mg of DHPP **3** (24% yield): (^{1}H NMR, DMSO-d_6) 0.76 (s, 3H, $-CH_3$), 0.78 (s, 3H, $-CH_3$), 1.10 (s, 6H, CH_3), 2.21 (t, J = 6.8 Hz, 2H, C(O)CH_2CH_2NH), 3.05−3.21 (m, 6H, $-CH_2$NH, $-SCH_2CH_2$), 3.38 (s, 2H, CH_2−OH), 3.75 (s, 1H, C*H*−OH), 7.65 (t, J = 5.8 Hz, 1H, NH), 8.06 (t, J = 5.8 Hz, 1H, NH). MALDI-TOF MS (monoisotopic) $[M + H]^+$: 378.9 (378.2 calcd).

(13) Radioactive HPLC analyses were performed with a C18 small-pore reverse phase analytical column using a linear 0−75% water/acetonitrile [0.1% trifluoroacetic acid (TFA)] gradient over 30 min at a flow rate of 1 mL/min. Reactions contained 2 μM [^{35}S]-L-cysteine (2.6 Ci/mmol), 0.4 mM L-Cys, 1 mM CoASH, 0.2 μM Sfp, 100 mM Tris-HCl pH 8, 5 mM $MgCl_2$, 2 μM HMWP2 1-1382, 20 μM HMWP1 1896-3163, 1 mM NADPH, 60 mM DHPP, and 10 mM ATP plus or minus 0.75 mM SAM. Reactions were incubated 14 h at 30 °C before quenching with 10% TCA, hydrolyzing with KOH, and injecting the supernatant onto the HPLC.

(14) Miller, D. A.; Walsh, C. T. *Biochemistry* **2001**, *40*, 5313−5321.

(15) Reaction components are similar to that previously described[13] except either 50 mM DHPP and varying amounts of [^{14}C]-SAM or 87 μM [^{14}C]-SAM and varying amounts of **3** were incubated at 30 °C for 5 h. Cellulose TLC (developed in 4:1:1 butanol:acetic acid:water) was used to analyze the reaction. Phosphorimages of TLC plates were obtained and analyzed densitometrically. (TLC species confirmed by MALDI-TOF MS analysis).

(16) DHP-CoA was synthesized in a similar manner as described for **3**. MALDI-TOF MS (monoisotopic) $[M + H]^+$: 867.7 (867.2 calcd).

(17) Suo, Z.; Tseng, C. C.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 99−104.

(18) DHP-S-ACP was confirmed by MALDI-TOF MS analysis $[M + H]^+$: 10740.7 (10739.9 calcd) following HPLC purification.

(19) Du, L.; Sanchez, C.; Chen, M.; Edwards, D. J.; Shen, B. *Chem. Biol.* **2000**, *7*, 623−642.

(20) Hendrickson, L.; Davis, C. R.; Roach, C.; Nguyen, D. K.; Aldrich, T.; McAda, P. C.; Reeves, C. D. *Chem. Biol.* **1999**, *6*, 429−439.