ABSTRACT: Cryptophycins (Crp) are a group of cyanobacterial depsipeptides with activity against drug-resistant tumors. Although they have been shown to be promising, further efforts are required to return these highly potent compounds to the clinic through a new generation of analogues with improved medicinal properties. Herein, we report a chemosynthetic route relying on the multifunctional enzyme CrpD-M2 that incorporates a 2-hydroxy acid moiety (unit D) into Crp analogues. CrpD-M2 is a unique non-ribosomal peptide synthetase (NRPS) module comprised of condensation—adenylation—ketoreduction—thiolation (C-A-KR-T) domains. We interrogated A-domain 2-keto and 2-hydroxy acid activation and loading, and KR domain activity in the presence of NADPH and NADH. The resulting 2-hydroxy acid was elongated with three synthetic Crp chain elongation intermediate analogues through ester bond formation catalyzed by CrpD-M2 and Crp TE. Finally, the enzyme-bound seco-Crp products were macroactonized by the Crp thioesterase. Analysis of these sequential steps was enabled through LC-FTICR-MS of enzyme-bound intermediates and products. This novel chemoenzymatic synthesis of Crp involves four sequential catalytic steps leading to the incorporation of a 2-hydroxy acid moiety in the final chain elongation intermediate. The presented work constitutes the first example where a NRPS-embedded KR domain is employed for assembly of a fully elaborated natural product, and serves as a proof-of-principle for chemoenzymatic synthesis of new Crp analogues.

Natural products have been widely applied to fight disease and offer chemical scaffolds for development of new analogues with improved/modified functions, achieved through semisynthesis, total synthesis, or chemoenzymatic synthesis.1–3 Cryptophycins are potent anticancer agents at picomolar concentrations and exert their cytotoxic effects in both vinca alkaloid- and taxol-resistant cancer cells that contribute to the proliferation of drug-resistant tumors.4 Their clinical potential and the synthetic challenges they pose have stimulated the development of alternative strategies to provide suitable amounts of material and new analogues with improved physiochemical properties for clinical studies.5 The cryptophycin gene cluster was recently elucidated and offers unique opportunities for assembly of the drug and new analogues using chemoenzymatic approaches.5 The gene cluster is comprised of two type I polyketide synthase (PKS) genes, crpA and crpB, two non-ribosomal peptide synthetase (NRPS) genes, crpC and crpD, and four tailoring enzyme genes, including a key P450 epoxidase (crpE). Previous studies from this laboratory have demonstrated the feasibility and efficiency of biocatalysts from this metabolic pathway to properly macrocyclize and regio- and stereospecifically epoxidize Crp intermediates to generate these natural products and novel analogues.6–8 CrpD is a bimodular NRPS involved in late-stage assembly of the Crp chain elongation intermediate. Bioinformatic analysis and precursor incorporation studies revealed that the substrate of its first module is methyl-β-alanine, converted from L-aspartic acid by CrpG, a β-methylaspartate-δ-carboxylase.6,9 These studies also suggested that 2-ketoisocaproic acid (2KIC, I) instead of

Received: May 23, 2011
Published: August 08, 2011
L-2-hydroxyisocaproic acid (L-2HIC, 2) was the substrate of CrpD module 2 (CrpD-M2), which was incorporated into Crp as unit D (Figure 1).6 Several natural Crp analogues contain unit D variations, including 3-methyl-2-hydroxyvalerate, 2-hydroxyvalerate, and 3-methyl-2-hydroxybutyrate (Figure S1).6 Altered bioactivity of these analogues suggests the importance of the unit D in anticancer activity, but limited efforts have been made to query analogues carrying unnatural unit D structures.10

Four distinct steps are catalyzed by CrpD-M2 in the Crp biosynthesis (Figure 1). In step I, the free-acid extender unit (e.g., 1) is activated by the CrpD-M2 adenylation (A) domain to form the corresponding acyl-AMP. This intermediate is then loaded onto the thiolation (T) domain active-site-bound phosphopantetheine arm through a transthioesterification reaction to form the acyl enzyme intermediate in step II. In the presence of reducing cofactors, the 2KIC enzyme intermediate is converted stereoselectively to L-2HIC in step III by the unique ketoreduction (2-KR) domain, based upon analysis of known products. The L-2-hydroxy group in unit D is condensed with Crp unit ABC biosynthetic intermediate, transferred from CrpD module I T domain through formation of an atypical (in an NRPS module) ester linkage by the condensation (C) domain (Figure 1, step IV).

In this report, four sequential steps were assessed to investigate the unique incorporation of a 2-hydroxy acid subunit into Crp natural products, and to probe the intrinsic substrate flexibility and synthetic potential of CrpD-M2. Chemoenzymatic synthesis of Crps 3,11 24,12 and 5113 (Figure 1, step V) serves as a proof-of-principle for efforts to generate Crp analogues with unnatural structures of units C and D through synthetic and biochemical methods.

A CrpD-M2 expression construct was generated by amplifying a DNA fragment consisting of C, A, KR, and T domains by PCR, and cloning into the BamH1 and XhoI sites of pET28a. This construct was overexpressed in Escherichia coli BAP1 strain for production of phosphopantetheinylated proteins.14 The N-terminal His-tagged protein was purified with Ni-NTA resin to ~80% purity (Figure S2). The integrity of the purified protein was verified by peptide map fingerprinting and FTICR-MS (Figure S3A, Table S1). The CrpD-M2 T domain active site was also identified, and proper post-translational modification of the T domain active-site Ser was verified by MS/MS (Figure S3B, Table S2).15

Bioinformatic analysis can reliably predict NRPS A domain specificity on the basis of binding pocket residue motifs.16 The conserved Asp235 involved in ionic interaction with the amino group of the substrate amino acid is replaced by Val235 in CrpD-M2 A domain (Table S3). Similar to unit D of Crp, a 2-hydroxy acid moiety is also involved in at least nine other natural products: bacillaene,17 barbamide,18 bassianolide,19 beauvericin,20 cereulide,21 enniatin,22 hectochlorin,23 kutzneride,24 and valinomycin.21 A similar replacement of Asp235 is conserved across all A domains responsible for incorporation of 2-hydroxy acid into these natural products (Table S3), predicting that the CrpD-M2 A domain prefers 2-hydroxy and/or 2-keto acids.

The well-established ATP-PPi exchange assay was employed to biochemically determine the substrate specificity of the CrpD-M2 A domain with 10 acyl acid substrates (Figure 2). CrpD-M2 activated 2-KIC (1) about 20 times more efficiently than its cognate amino acid L-leucine (14), consistent with bioinformatic prediction and previous feeding experiments.6 Interestingly, L-2HIC (2), which was not incorporated into the Crp final structure during an in vivo precursor incorporation study, was among the best substrates in the assay. A similar high level of selectivity for the natural substrate 2-oxovalerate (15) was observed. CrpD-M2 specificity to two other natural unit D fragments, 3-methyl-2-oxovalerate (13) and 3-methyl-2-oxobutyrate (16), was decreased ~50% and 90%, respectively. This result, along with the observed weak activation of unnatural substrates 2-oxobutyrate (17) and phenyl pyruvate (19), suggests that the size and bulk of substrate side chains are important in CrpD-M2 A domain recognition.

In an effort to assess further the flexibility of the CrpD-M2 A domain, we tested 2-keto-γ-(methylthio)butyrate (AKGB, 12) as a substrate. Its effective activation (Figure 2) demonstrates the potential of native CrpD-M2 in producing novel Crp analogues with an altered unit D moiety. Moreover, the weak activation of 4-methylvalerate (18) by CrpD-M2 A domain reveals the importance of the α-position functional group in enzyme recognition. Therefore, CrpD-M2 A domain has relatively relaxed substrate specificity and exhibits a similar selectivity toward 2-keto and 2-hydroxy acids (Figure 1, step I). ATP-PPi exchange assays have been applied to examine substrate preference of A domains in the biosynthesis of bacillaene,17 barbamide,18 cereulide,21 enniatin,22 hectochlorin,23 kutzneride,24 and valinomycin.21 Only HctEIVA from the hectochlorin pathway displayed a similar selectivity to 2-keto- and 2-hydroxy acids compared to the CrpD-M2 A domain.23

We next monitored substrate loading directly on the T-domain of CrpD-M2 (Figure 1, step II).5,13–28 Enzyme reactions were terminated by proteolysis with trypsin, and the active-site peptides bound with extender units were separated and analyzed by LC-FTICR-MS and LC-IT-MS/MS (Tables S4 and S5, Figures S26–S35). As shown in Figure S28A,C, T domain active-site peptides bound with L-2HIC and 2-KIC showed masses of 4116.21 and 4114.15, respectively, at a charge state of 4+, matching theoretical values within ±30 ppm (Table S4). d-2HIC (20) was also loaded on the CrpD-M2 T domain active site, as shown by the observed mass of 4116.18 (Figure S28B).

Since only L-2HIC-containing Crp analogues have been isolated and characterized from the cyanobacterium Nostoc sp. ATCC S3789, we suspect that factors other than A domain selectivity, such as substrate availability and/or downstream processing, determine the final outcome. It is well-known that 2-keto acids are indispensable intermediates in amino acid biosynthesis, such as 1, 13, and 16 in the biosynthesis of leucine, isoleucine, and valine, respectively. The availability of free 2-hydroxy acid may be ascribed to a pathway-specific enzyme. For example, A domains for the biosynthesis of bassianolide,17 beauvericin,20 and enniatin22 are specific to d-2-hydroxy isovalerate (d-2HIV), and a pathway-specific NADPH-dependent reductase is essential.

**Figure 2.** Examination of CrpD-M2 A domain substrate specificity using the radio-PPi exchange assay. (A) Relative activity of the A domain normalized to 2. (B) Extender units investigated in this assay.
for stereospecific reduction of 2-keto isovalerate. Since a corresponding reducing enzyme is not encoded in the Crp gene cluster, it is possible that 2KIC is the native substrate of CrpD-M2 A domain.

The loaded 2KIC is proposed to be reduced to l-2HIC by the α-KR domain of CrpD-M2 (Figure 1, step III). This type of KR domain is also embedded in NRPS modules of cereulide,21 hectochlorin,23 kutzneride,24 and valinomycin,21 and its α-keto reduction activity and stereospecificity were biochemically confirmed in the cereulide system.21 Bioinformatic analysis indicates that CrpD-M2 KR domain is grouped with these NRPS KR domains and is phylogenetically distinct from the more thoroughly studied PKS β-KR domains (Figure S4A). The stereochemical outcome of PKS β-KR domains can be predicted on the basis of conserved amino acid sequence motifs.30–32 A similar analysis was performed to predict 2-hydroxy chirality introduced by KR domains from CrpD-M2, CesA, and CesB (Figure S4B). The corresponding KR domain in CesA and CesB produces l-2HIC and d-2HIV, respectively. However, none of these enzymes contain the conserved motifs expected for either type A or type B β-KR polypeptides. Given the phylogenetic distance and positional difference of the keto group in their respective substrates between PKS β-KR domains and NRPS α-KR domains, this sequence divergence was not unexpected. The first KR domain in the hybrid PKS/NRPS Pks/J involved in bactenial biosynthesis catalyzes both β- and α-ketone reduction, with 10-fold preference toward the former reaction. Based on structural analysis, the product of β-ketone reduction was consistent with the KR A-type sequence motif. However, as with NRPS α-KR domains, the Pks/J KR cannot be grouped as type A or B (Figure S4B). Therefore, it is not possible to classify NRPS α-KR domains using current PKS β-KR domain bioinformatic tools,30–32 and product chirality must be determined experimentally (Figure S5). In the CrpD-M2 KR reaction, the S-configuration (e.g., L-2HIC) is expected since all natural Crp contain this stereocchemistry.

After loading of 2KIC (Figure S28C), addition of reducing cofactors (e.g., NAD(P)H) to the CrpD-M2 reaction resulted in a mass shift observed by FTICR-MS (Figure S28D,E). The increase of 0.5 m/z unit (a 2-Da shift in the deconvoluted mass) is consistent with 2KIC conversion to 2HIC as the product of the α-KR reaction. Both NADH and NADPH operated within a similar (1–2-fold) efficiency as hydride donors, based on peak abundance (Figure S28D,E). Future structural analysis of the CrpD-M2 α-KR domain may contribute to understanding stereochemical control in this enzyme subclass.

Next, the ability of CrpD-M2 to form the seco-Crp intermediate was investigated using synthetic SNAC-ABC chain elongation intermediates 3–5 as the starting point (Figure 1B). The synthetic scheme followed our previously established route33 and the SNAC-ABC intermediates were confirmed using NMR and high-resolution mass spectrometry (Supporting Information). The intermediate with the monomethylated unit C (3-amino-2(R)-methylpropionyl, 3) was then combined with CrpD-M2 and 2. The C domain of CrpD-M2 is proposed to catalyze formation of an ester bond with the 2-hydroxy group of unit D as the nucleophile (Figure 1, step IV). Formation of the ester bond was confirmed by detecting the reaction products released from CrpD-M2 T domain following addition of the excised Crp TE (Figure 1, step V; Figure 3A,B). Both cyclic 6 (Figure 3A) and linear 9 (Figure 3B) were observed in the extracted ion chromatograms (EICs). Previously, a didomain NRPS (T-C), Fum14p, was shown to form a C–O bond in the biosynthesis of the fungal mycotoxin fumonisins. The only other C domain shown to catalyze C–O bond formation is a discrete enzyme, SgcC5, in C-to-T biosynthesis. In both cases, donor substrates are tethered to T domains, while the nucleophile (–OH) is from a small-molecule extender unit. This study represents the first example of a C domain in a complete NRPS module that catalyzes the incorporation of non-amino acid moieties as an ester synthase.

Assuming both cyclic and linear products share similar ionization efficiency in the positive ion mode, we conclude that cyclic 6 was formed as the predominant species over linear 9 (Figure 3A,B). This result demonstrates chemoenzymatic synthesis of 6 through five catalytic steps (Figure 1, steps I–V). Formation of 6 was confirmed by MS/MS (Figure S6A) and LC co-elution with an authentic standard (Figure S7). Observed isotope patterns of cyclic product and MS3 fragments were also consistent with the +2 37Cl shift from unit B.

Similarly, the synthetic chain elongation intermediate with desmethyl (3-amino-propionyl, 4) or gem-dimethyl (3-amino-dimethylpropionyl, 5) unit C and l-2HIC (2) were used as substrates in the CrpD-M2 reaction, along with Crp TE for off-loading/cyclization. Both Crp 24 (7) and Crp S1 (8) were generated by this chemoenzymatic route, indicating the versatility of the CrpD-M2 C domain and Crp TE (Figure S8A,E). The corresponding hydrolyzed linear products (10, 11) were also observed (Figure S8B,F). Assuming similar ionization efficiency between linear and cyclic products, linear 10 was the predominant species compared to cyclic 7 (Figure S8A,B), and cyclic 8 was present in similar abundance compared to its linear counterpart, 11 (Figure S8E,F). Using SNAC-ABCD analogues as native substrate mimics, Beck et al. investigated the impact of unit C methylation on Crp TE-mediated macrocyclization.5 The analogue bearing the 3-amino-dimethylpropionyl group was found to produce more cyclic product (6:1) than the one with a 3-amino-propionyl moiety (5:1) but less than the one with a 3-amino-2(R)-methylpropionyl group (10:1). We found a similar order of unit C reactivity when the native T domain-bound substrates generated by CrpD-M2 were supplied to CrpD TE.
In an effort to further probe the flexibility of CrpD-M2 to substrate stereochirality, d-2HIC (20) was substituted for 2 in the chemoenzymatic reaction. Following loading of chain elongation intermediate 3, no cyclic depsipeptide product was detected (Figure 3C), but a small amount of linear product was formed (Figure 3D), suggesting that the C domain of CrpD-M2 was able to recognize the stereoisomer of its acceptor substrate and catalyze ester bond formation with its donor ABC chain.

In summary, non-amino acid extender subunits selected and processed by NRPS enzymes have been found in a handful of natural products isolated from bacteria and fungi. However, a systematic approach to further explore unit D selection, loading, reduction, elongation through an ester bond, and final product formation. FTICR-MS was employed to assess these five sequential biochemical reactions that occurred through a complete NRPS module including A, C, KR, T, and TE domains. This is also the first study in which a NRPS bearing an embedded KR domain was used to directly generate bioactive compounds from fully elaborated natural and unnatural chain elongation intermediates to provide cyclic cryptophycins 3, 24, and 51. Thus, CrpD-M2 as a chemoenzymatic reagent for unit D stereochirality selection are the CrpD-M2 KR domain and Crp TE rather than its C or A domain.

**ASSOCIATED CONTENT**

Supporting Information. Methods, figures, and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

Corresponding Author
davidhs@umich.edu

Author Contributions
*These authors contributed equally to this work.

**ACKNOWLEDGMENT**

Y.D. was supported by a Rackham Predoctoral Fellowship. This work was supported by NIH grant CA108874, GM076477 and the Hans W. Vahlteich Professorship (to D.H.S.). Work in K.H.’s laboratory is supported by an NSF Career Award (CHE-05-47699).

**REFERENCES**


