Thioesterase from Cereulide biosynthesis is responsible for oligomerization and macrocyclization of a linear tetradepsipeptide

Graham W. Heberlig and Christopher N. Boddy

Department of Chemistry and Biomolecular Sciences
Centre for Chemical and Synthetic Biology
University of Ottawa
Ottawa, ON , K1N 6N5
cboddy@uottawa.ca

Abstract

Cereulide is a cyclic depsidodecapeptide produced in *Bacillus cereus* by two non-ribosomal peptide synthetases, CesA and CesB. While highly similar in structure and with a homologous biosynthetic gene cluster to valinomycin, recent work suggests that cereulide is produced via a different mechanism, which relys on a non-canonical coupling of two didepsipeptide-PCP bound intermediates. Ultimately this alternative mechanism generates a tetradepsipeptide-PCP bound intermediate that is prosed to differ from the tetradepsipeptide predicted from canonical activity of CesA and CesB. To test this hypothesis, we chemically synthetize both tetradepsipeptides as N-acetyl cysteamine thioesters and probed the ability of the purified recombinant terminal CesB thioesterase (CesB TE) to oligomerize and macrocyclize each substrate. Only the canonical substrate is converted cereulide, ruling out this alternative mechanism. We also show that CesB TE can use related tertradepsipeptide substrates, such as the valinomycin tetradespipetide and a hybride cereulide-valinomycin tetradespispetide in conjunction with its native substrate to generate chimeric natural products. This work clarifies the biosynthetic origins of cereulide and provides a powerful biocatalyst to access analogs of these ionophoric forming natural products.

Introduction

Cereulide, **1**, is a cyclic depsidodecapeptide produced by *Bacillus cereus* that was discovered in 1994.[1,2] It is a potassium ionophore, which rapidly depolarizes bacterial and mitochondrial membranes[3] and as such it is a potent toxin.[4,5] At lower doses it is an emetic toxin impacting the 5-HT$_3$ receptor, stimulating the vagus afferent nerve leading to vomiting.[2] Due to the potential contamination of food products with *Bacillus cereus* and the high chemical and temperature stability of cereulide, poisoning is a significant safety concern.[6–8]

The biosynthetic gene cluster responsible for cereulide biosynthesis was identified in 2004.[9–11] Cereulide is produced by two non-ribosomal peptide synthetases, CesA and CesB, encoded on the pXO1 megaplasmids.[12,13] CesA uses α-ketoisocaproate and D-Ala to generate the PCP linked D-HIC-D-Ala depsipeptide.[14–16] CesB uses α-ketoisovalerate and L-Val and in conjunction with CesA generates cereulide.[14,15] In the biosynthesis of the related ionophore valinomycin, **2**, Vlm1 generates a PCP linked D-Hiv- D-Val, which is condensed with L-lactate and then L–Val to generate a PCP linked D-Hiv-D-Val-L-lac-L-Val.[14,17] Finally the depsitetrapeptide is trimerized and macrocyclized by the C-terminal thioesterase (TE) domain of Vlm2 to generate valinomycin.[18]

Recent work suggests that despite the similarity between the structures of cereulide and valinomycin and their biosynthetic gene clusters, cereulide is produced via a different mechanism. It is proposed that a L-Hiv-L-Val-D-Hic-D-Ala PCP bound intermediate in generated (Fig 1B) rather than the expected D-Hic-D-Ala-L-Hiv-L-Val intermediate (Fig 1A).[19] Further it is suggested that some of the isocereulides[20–22] are produced by the unexpected coupling of a TE bound depsidipeptide with a PCP bound depsitetrapeptide.[19] This study sets out to test if valinomycin and cereulide have differing mechanisms of TE-mediate oligomerization and macrocyclization by characterizing the ability of the CesTE to oligomerize and macrocyclize tetradepsipeptide substrates.

Figure 1. Two proposed pathways for the biosynthesis of cereulide by CesA and CesB. A. This mechanism for cereulide formation is consistent with the expected activity of all the catalytic domains in CesA and CesB, generating a D-Hic-D-Ala-L-Hiv-L-Val-PCP bound intermediate on CesB that must be oligomerized and cyclized by the TE. B. This proposed mechanism invokes a non-canonical use of the C* domain from CesA. Unlike mechanism A, mechanism B leads to the formation of a L-Hiv-L-Val-D-Hic-D-Ala PCP bound intermediate on CesA that must be oligomerized and cyclized by the TE.[19] This work test both of these proposed mechanisms. Our data supports mechanism A and rules out mechanism B.

Results and Discussion

      To test the TE-mediated oligomerization and macrocyclization step from cereulide biosynthesis we synthesized both the proposed cereulide TE substrate[19] as a N-acetyl cysteamine (NAC) thioester (**7**, L-Hiv-L-Val-D-Hic-D-Ala-SNAC) and the expected NAC substrate based on valinomycin homology (**12**, D-Hiv-D-Ala-L-Hiv-L-Val-SNAC). For **7** (Scheme 1), a Minsunobu reaction was used to link Boc-L-Val to allyl L-α-hydroxycaproate to generate the D-Hic containing depsidipeptide. Deprotection of the allyl ester with Pd(PPh₃)₄ and coupling with D-Ala-SNAC generated the depsitripeptide. Finally Boc deprotection and coupling with L-Hiv furnished the depsitetrapeptide substrate **7**. Synthesis of **12** began with coupling Boc-D-Ala and allyl L-α-hydroxyisovalerate to form the ester linkage. From this intermediate the synthesis mirrored the route to access **7**, with allyl deprotection followed by coupling to L-Val-SNAC, Boc deprotection, and coupling with D-Hic to afford **12**.
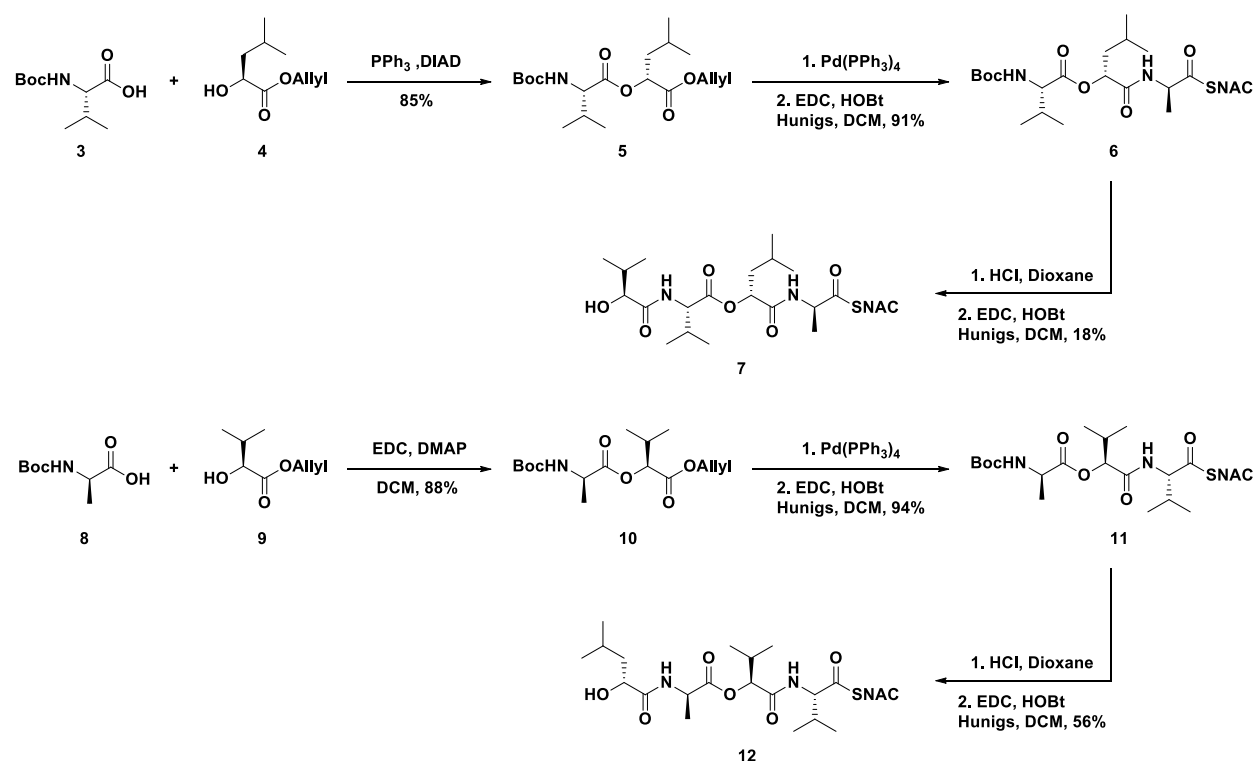
Scheme 1. Synthesis of SNAC substrates 7 and 12.

      The C-terminal TE domain was amplified from *cesB* and cloned into an expression vector under the inducible T7 promoter. Ces TE was overexpressed in *E. coli* BL21 and purified to homogeneity by metal affinity chromatography. Treatment of Ces TE with **7** in 50 mM Tris buffer pH 7.4 led to no detectable formation of cereulide by LCMS (Fig. 2). No oligomeric structures or hydrolysis of **7** were detected. This is consistent with Ces TE being unable to load **7**. In contrast incubation of **12** with Ces TE

generated cereulide, **1**, as well as dimeric, trimeric, and tetrameric SNAC products, **13**, **14**, and **15** respectively (Fig. 3). Both the no enzyme and the boiled enzyme controls showed no production of **1**, confirming its production is TE-dependent. This result clearly demonstrates that the Ces TE is capable of oligomerizing and macrocyclizing the expected depsitetrapeptide substrate derived from canonical NRPS activity, analogous to valinomycin biosynthesis.[18]
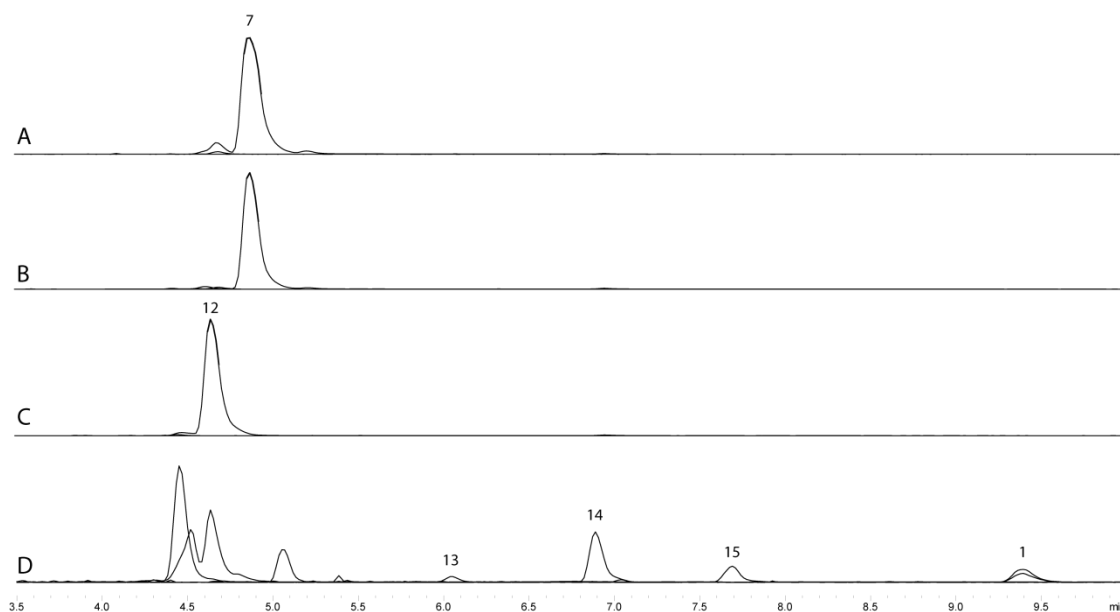

Figure 2. LCMS traces of incubation of synthetic tetradesipeptide N-acetyl cysteamine substrates with CesB TE. A. Negative control. Boiled CesB TE (15 μM) with 1 mM 7 in 50 mM Tris pH 7.4, 14 h 22 °C. **B.** 15 μM CesB TE with 1 mM **7** in 50 mM Tris pH 7.4 for 14 h at 22 °C. No cereulide or oligomeric product is formed. **C.** Negative control. Boiled CesB TE (15 μM) with 1 mM **12** in 50 mM Tris pH 7.4, 14 h 22 °C. **D.** 15 μM CesB TE with 1 mM **12** in 50 mM Tris pH 7.4 for 14 h at 22 °C. Cereulide (**1**) as well as dimeric, trimeric and tetrameric N-acetyl cysteamine thioester products (**13**, **14**, and **15**, Fig. 3) are formed. All traces are ion extraction for m/z = 504, 888, 1272, 1658, 1170, which correspond to the [M+H]$^+$ for **7/12**, **13**, **14**, and **15**, and the [M+NH$_4$]$^+$ for **1** respectively.
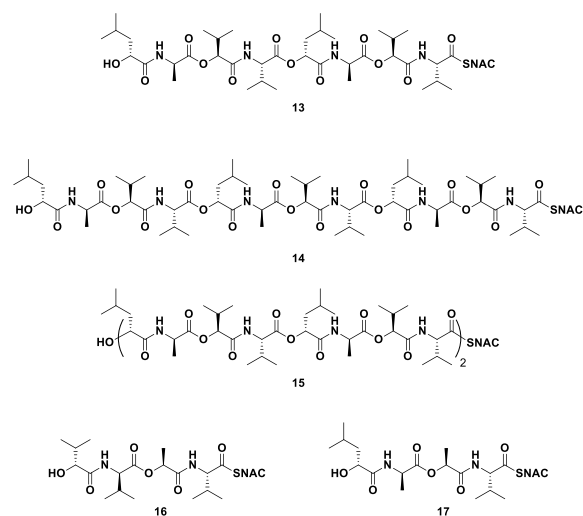
Figure 3. Structures of ceruelide oligomers and additional SNAC substrates.

Kinetic characterization of Ces TE activity with **12** as monitored by thiol release using Elman's reagent provided a $k_{cat}/K_M = 0.28 \pm 0.02$ M^{-1}s^{-1}. This is comparable, though on the low end, to specificity constants seen for other characterized TE domains,[23] such as the epothilone[24] and 6-deoxyethythronolide B[25] macrocyclizing TEs. The Ces TE could not be saturated even at 4 mM **12**, preventing k_{cat} and K_M from being independently determined. The K_M, which based on our kinetic characterization is > 4 mM, is significantly higher than other TEs responsible for oligomerization and macrocyclization of non-ribosomal peptides. For example the tyrocidine TE (TycC TE) and the gramicidin TE (GrsB TE), which both dimerize D-Phe-L-Pro-L-Val-L-Orn-L-Leu-SNAC and cyclize the product to generate gramicidin S, show K$_M$s in the range of 0.3-0.5 mM.[26,27] We speculate that in order to effectively oligomerize and macrocycle, these TEs likely possess enhanced affinity for oligomerized substrates over monomeric substrates as the oligomers have higher structural homology to the macrocyclization substrates and thus can be stabilized by more of the enzyme-substrate interactions involved in macrocyclization activity. In the case of Ces TE, we expect higher affinity for the more mature trimerized intermediate over the dimerized intermediate and monomeric substrate. This is consistent with the observation that the dimeric and trimeric SNAC intermediates do not accumulate to appreciable levels in the biochemical assays but are rather processed to products preferentially over the SNAC activated monomer. Unlike the GrsB TE, which dimerizes and macrocyclizes, Ces TE must trimerize and macrocyclize. This is also consistent with a higher K_M for the monomeric substrate of Ces TE than the monomeric substrate of GrsB TE.

Treatment of the Ces TE with the related tetradepsipeptide substrate from the valinomycin biosynthetic pathway (**16**, D-Hiv-D-Val-L-Lac-L-Val-SNAC) resulted in very little processing of the substrate (Fig S3D). While trace formation of valinomycin, **2**, and the trimeric intermediate are observable, the main product is the dimeric intermediate. Dimer formation shows that **16** is loaded onto the active site Ser of Ces TE. Attack by a second equivalent of **16**, which is required for formation of dimer and ultimately other higher order structures including **2**, occurs, however the low processivity of Ces TE with

this substrate suggests that attack by the second equivalent of substrate may be slow. Treatment of Ces TE with a mixture of **12** and **16** (Fig 4A) results in the formation of cereulide (**1**) and valinomycin (**2**) as well as hybrid cereulide-valinomycin macrocycles. Based on our data both the cereulide and valinomycin SNAC monomers can be loaded onto the active site Ser. We propose that a second equivalent of the cereulide monomer **12** is more effective than the valinomycin monomer **16** at acting as a substrate to effect off-loading for formation of dimeric and trimeric products. Following this mechanism, the hybrid cereulide-valinomycin product is generated from valinomycin monomer **16** loading onto Ces TE followed by a cereulide monomer **12** effecting off-loading via transesterification to make the activated valinomycin-cereulide dimer (D-Hiv-D-Val-L-Lac-L-Val-D-Hiv-D-Ala-L-Hiv-L-Val-SNAC). This substrate is then loaded back on the Ces TE and transesterified with the cereulide monomer **12**, generating the linear hybrid trimer with two cereulide monomers and one valinomycin monomer (Cer$_2$-Vlm-SNAC). Subsequent Ces TE loading and macrocyclization then gives the hybrid product. Because the valinomycin monomer **16** is not as effective a substrate to effect off-loading, misincorporation is limited as clearly seen in the ratio of trimeric Cer$_3$-SNAC versus Cer$_2$-Vlm-SNAC products (Fig 4A).
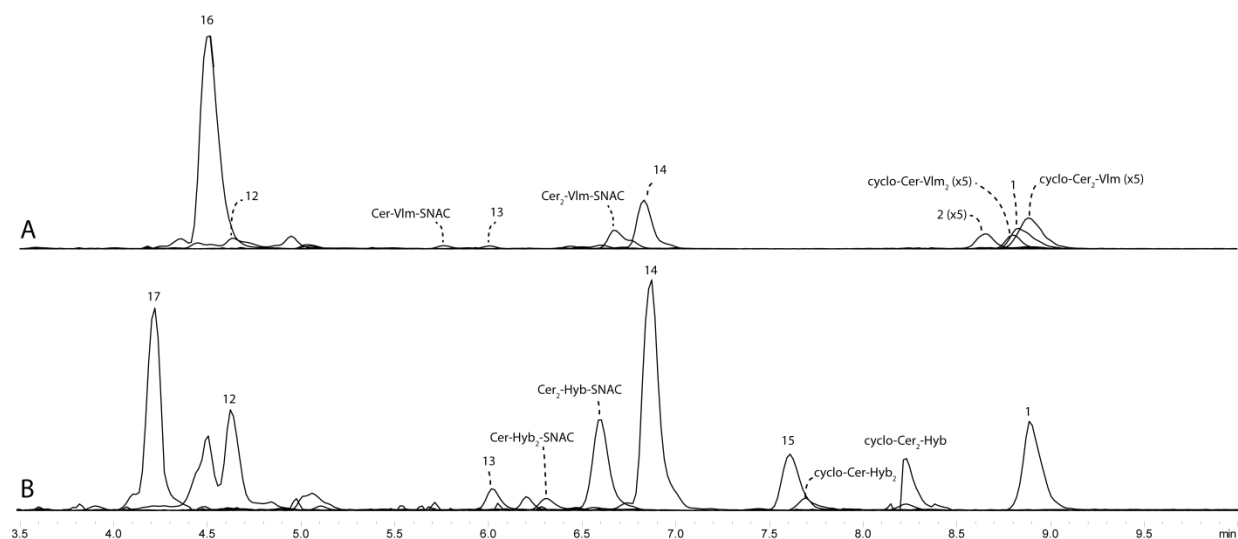

Figure 4. . LCMS traces of incubation of two different synthetic tetradesipeptide N-acetyl cysteamine substrates with CesB TE. A. 15 µM CesB TE with 2 mM **12** and 1 mM **16** in in 50 mM Tris pH 7.4, 14 h at 22 °C. The expected natural products **1** and **2** were formed as well as two hybrid natural products comprised of two cereulide monomers and one valinomycin monomer (cyclo-Cer$_2$-Vlm) and the other of two valinomycin monomers and one cereulide monomer (cyclco-Cer-Vlm$_2$). A boiled enzyme control showed no product formation. **B**. 15 µM CesB TE with 2 mM **12** and 1 mM **17** in in 50 mM Tris pH 7.4, 14 h at 22 °C. Cereulide as well as a complex mixture of oligomeric-SNAC compounds were generated. A boiled enzyme control generated no oligomeric or cyclic products. Traces are ion extraction for [M+Na]$^+$ ions with m/z = 526 (**12**), 512 (**16**), and 498 (**17**), [M+H]$^+$ ions with m/z = 889 (**13**), 875 (Cer-Vlm-SNAC), 861 (Cer-Hyb-SNAC), 1273 (**14**), 1259 (Cer$_2$-Vlm-SNAC), 1245 (Cer-Vlm$_2$-SNAC & Cer$_2$-Hyb-SNAC), 1217 (Cer-Hyb$_2$-SNAC), and 1658 (**15**), and [M+NH$_4$]$^+$ ions with m/z = 1170 (**1**), 1128 (**2**), 1156 (cyclo-Cer$_2$-Vlm), 1132 (Cyclo-Cer-Vlm$_2$), 1142 (cyclo-Cer$_2$-Hyb), and 1114 (cyclo-Cer-Hyb$_2$).

Based on this working model, we investigated a hybrid cereulide-valinomycin hybrid monomer with the expectation that this substrate would load on the Ces TE but not be able to act as an off-loading substrate to make higher order oligomers. We synthesized the hybrid tetradepsipeptide substrate, D-Hic-D-Ala-L-Lac-L-Val-SNAC, **17** (Scheme S1). Treatment of Ces TE with **17** did not lead to macrocycle formation or oligomeric products, consistent with our working model (Fig S3B). Treatment of Ces TE with a combination of **12** and **17** generated both cereulide and the hybrid product containing two cereulide monomers and one hybrid monomer from **17**, also fully consistent with our model (Fig 4B).

Our data and working model for Ces TE oligomerizing and macrocyclizing activity are consistent with the mechanism of the closely related ionophore natural product valinomycin[18] and provide a mechanism for the biosynthesis of the isocereulides, which are produced along with cereulide in trace amounts.[20–22] The ability of Ces TE to cyclize the two non-native cereulide-valinomycin hybrids above clearly indicates that some degree of non-native side chain composition can be accommodated by the TE. Thus isocereulides A, B, C, and D are generated by the single misincorporation of α-ketoisocaproate in place of α-ketoisovalerate by A1 of CesB, α-ketoisovalerate in place of α-ketoisocaproate by A1 of CesA, D-Ser in place of D-Ala by A2 of CesA, and L-Ala in place of L-Val by A2 of CesB, respectively.

The unexpected linear tetra-, hexa-, octa-, deca-, and dodecadepsipeptide esters isolated from alcoholic extracts of cereulide producing strains were initially proposed to be due to alcoholic hydrolysis of PCP and TE linked biosynthetic intermediates.[19] To rationalize hexa- and decadepsipeptide formation, it was proposed that non-canonical C domain activity could be occurring to couple dipeptide products produced by CesA and CesB.[19] Based on our biochemical characterization of Ces TE we suggest an alternative hypothesis fully consistent with the data and the canonical activity of C domains. In biochemical treatment of Ces TE with the native SNAC activated substrate **12**, we observe formation of linear SNAC activated tetramer (hexdecadepsipeptide). This occurs when the TE-O-trimer intermediate undergoes attack by another equivalent of monomer **12** before macrocyclization can occur. In our assays the linear activated tetramer is never observed to undergo macrocyclization. Thus in the producing organism if the TE elongates the substrate beyond trimer, macrocyclization can not occur and the tetramer will either be further elongated or hydrolyzed off. Unlike cereulide, these extralong linear oligomers should be sensitive to alcoholic hydrolysis and readily generate the esters observed. Significantly, this mechanism provides a route to produce multiple equivalents of the linear esters, rather than being limited to between one and three equivalents relative to the amounts of CesA and CesB, as would be required if the esters were formed by alcoholic hydrolysis of the PCP and TE linked intermediates.

Thus in summary this work demonstrates that cereulide is generated highly analogously to valinomycin where the TE mediates trimerization and macorcyclization of the expected tetradepsipeptide substrate based on canonical activity of the NRPS domains present in CesA and CesB. In addition the Ces TE is selective, in our biochemical assays, for the nucleophile which hydrolyzes the depsipeptidyl-TE intermediates. This can be used to access hybrid cereulide macrocycles, with a single non-native monomer incorporated into the final macrocycle. Lastly we propose a mechanism consistent with all data to account for the formation of unexpected linear ester from cereulide producing strains, which eliminates the need to resort to non-canonical activity of the CesB C domain.

Synthesis

Boc-L-Val-D-HIC Allyl Ester (5). In a round bottom flask, 453 mg Boc-L-Val (2.08 mmol, 1.2 eq.) 172 mg L-HIC allyl ester (1.74 mmol, 1 eq.) and 915 mg triphenylphosphine (PPh$_3$, 3.49 mmol, 2 eq.) were dissolved in 18 mL dry THF. The stirring solution was cooled to 0 °C on an ice bath for 10 minutes. To this was added 685 μL diisopropyl azodicarboxylate (DIAD, 705 mg, 3.49 mmol, 2 eq.) dropwise and the reaction was stirred on ice for an additional hour then allowed to come up to ambient temperature and stirred overnight. At completion the reaction was concentrated and the title compound (550 mg, 85%) was purified by silica column chromatography (gradient 10 to 15% EtOAc in hexanes). ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.87 (ddt, J = 17.2, 10.4, 5.8 Hz, 1H), 5.30 (dd, J = 17.2, 1.4 Hz, 1H), 5.23 (dd, J = 10.4, 1.1 Hz, 1H), 5.05 (dd, J = 10.0, 3.8 Hz, 1H), 4.96 (d, J = 9.0 Hz, 1H), 4.61 (dd, J = 5.8, 0.9 Hz, 2H), 4.29 (dd, J = 9.0, 4.6 Hz, 1H), 2.18 (dt, J = 12.5, 6.4 Hz, 1H), 1.82 (tt, J = 15.5, 5.1 Hz, 1H), 1.78 − 1.67 (m, 1H), 1.68 − 1.59 (m, 1H), 1.42 (s, 9H), 0.97 (d, J = 6.8 Hz, 3H), 0.93 (d, J = 6.5 Hz, 3H), 0.92 − 0.87 (m, 6H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.83, 169.96, 155.51, 131.47, 118.84, 79.72, 71.57, 65.88, 58.64, 39.72, 31.25, 28.30, 24.53, 23.07, 21.28, 19.08, 17.38. HRMS (ESI+) Exact mass calculated for $C_{29}H_{33}NNaO_6$: 394.2200. Found 394.2192.

Boc-L-Val-D-HIC-D-Ala-SNAC (6). In a round bottom flask, 100 mg of **5** (0.27 mmol, 1 eq.) was dissolved in 2 mL of CH$_2$Cl$_2$ under a nitrogen atmosphere. To this solution 31 μL of morpholine (30 mg, 0.35 mmol, 1.3 eq.) and 16 mg of Pd(PPh$_3$)$_4$ was added in a single portion. The reaction was stirred at ambient temperature and monitored by TLC. At completion the reaction was quenched by the addition of 10 % aq. HCl, the organic layer was removed and the remaining aqueous fraction was extracted 3 x with CH$_2$Cl$_2$. The combined organic fractions were washed with brine, dried over Na$_2$SO$_4$ and concentrated in a dry round bottom flask, this intermediate was used immediately in the subsequent reaction. To the flask was added 2 mL of CH$_2$Cl$_2$ and 61 mg of D-Ala-SNAC (as HCl salt, 0.27 mmol, 1 eq.). To the resulting solution was added 140 μL of Hünig's base (104 mg, 0.81 mmol, 3 eq.), 55 mg HOBt (0.4 mmol, 1.5 eq.) and 78 mg EDC (0.4 mmol, 1.5 eq.). The reaction was stirred under argon at ambient temperature for 20 h. The reaction was quenched with NH$_4$Cl (aq), extracted 3 × with CH$_2$Cl$_2$, washed with NaHCO$_3$(aq), then with brine, dried over Na$_2$SO$_4$, and concentrated. The title compound (124 mg, 91 %) was partially purified by silica column chromatography (40 % acetone in hexanes) and taken into the subsequent reaction. ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.09 (s, 1H), 5.32 (t, J = 6.7 Hz, 1H), 5.06 (d, J = 6.2 Hz, 1H), 4.57 − 4.45 (m, 1H), 3.92 (t, J = 6.7 Hz, 1H), 3.44 − 3.26 (m, 3H), 2.99 (td, J = 6.3, 3.6 Hz, 1H), 2.94 (t, J = 6.0 Hz, 2H), 2.02 (tt, J = 13.5, 6.7 Hz, 1H), 1.92 (s, J = 2.4 Hz, 4H), 1.80 − 1.75 (m, 2H), 1.67 (dt, J = 20.2, 6.6 Hz, 1H), 1.42 (d, J = 2.7 Hz, 3H), 1.37 (s, 10H), 1.02 − 0.96 (m, 6H), 0.93 − 0.88 (m, 6H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 201.33, 172.17, 170.77, 170.45, 156.23, 80.67, 72.66, 60.17, 55.59, 40.19, 39.20, 30.02, 28.48, 28.32, 24.41, 23.29, 23.12, 21.17, 19.08, 18.66, 16.99.

HO-L-HIV-L-Val-D-HIC-D-Ala-SNAC (7). To a round bottom flask was added 100 mg **6** (0.2 mmol, 1 eq.) in a minimal amount of EtOAc and cooled to 0˚C. To this was added 2 mL of 4M HCl in dioxane (Sigma), and the reaction allowed to warm to ambient temperature. The reaction was monitored by TLC and at completion all solvent was removed by rotary evaporation. The unpurified intermediate was used immediately and was dissolved in 1 mL of CH$_2$Cl$_2$ and to this was added 105 μL of Hünig's base (78 mg, 0.6 mmol, 3 eq.), 26 mg of L-α-hydroxyisovaleric acid (0.21 mmol, 1.1 eq.), 40 mg HOBt (0.3 mmol, 1.5 eq.), and 60 mg EDC (0.3 mmol, 1.5 eq.). The reaction was stirred at ambient temperature overnight and

at completion was quenched with NH₄Cl(aq), extracted 5 × with CH₂Cl₂, washed with NaHCO₃(aq), then with brine, dried over Na₂SO₄, and concentrated. The title compound (18 mg, 18%) was purified by silica column chromatography (50 % acetone in hexanes). ^{1}H NMR (600 MHz, CDCl₃) δ 7.77 (d, J = 8.1 Hz, 1H), 7.35 (d, J = 6.8 Hz, 1H), 6.16 (s, 1H), 5.28 (dd, J = 10.6, 3.0 Hz, 1H), 4.62 (p, J = 7.4 Hz, 1H), 4.14 (dd, J = 9.0, 6.8 Hz, 1H), 3.92 (d, J = 3.3 Hz, 1H), 3.52 (td, J = 13.8, 6.6 Hz, 1H), 3.39 – 3.32 (m, 1H), 3.17 (dt, J = 13.4, 6.1 Hz, 1H), 2.16 – 2.07 (m, 2H), 1.88 – 1.70 (m, 3H), 1.48 (d, J = 7.3 Hz, 3H), 1.08 (d, J = 6.7 Hz, 3H), 1.04 (d, J = 3.1 Hz, 3H), 1.03 (d, J = 2.8 Hz, 3H), 0.97 (d, J = 6.3 Hz, 3H), 0.94 (d, J = 6.2 Hz, 3H), 0.90 (d, J = 6.9 Hz, 3H). ^{13}C NMR (150 MHz, CDCl₃) δ 201.79, 175.88, 172.10, 171.54, 171.13, 76.11, 72.94, 59.13, 55.02, 40.98, 40.29, 38.90, 31.82, 29.57, 28.58, 24.48, 23.33, 23.04, 20.86, 19.06, 19.03, 18.95, 17.00, 16.09. HRMS (ESI+) Exact mass calculated for C₂₃H₄₁N₃NaO₇S: 526.2563. Found 526.2551.

Boc-D-Ala-L-HIV Allyl Ester (10). In a round bottom flask, 200 mg L-2-hydroxyisovalerate allyl ester (1.26 mmol, 1 eq.) and 264 mg Boc-D-Ala (1.39 mmol, 1.1 eq.) were dissolved in 6.5 mL of CH₂Cl₂. To this solution was added 155 mg 4-dimethylaminopyridine (DMAP, 1.26 mmol, 1 eq.) and 365 mg 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC, 1.9 mmol, 1.5 eq.) were added at ambient temperature. The resulting solution was stirred for 14 h at ambient temperature. The reaction was quenched with NH₄Cl (aq), extracted 3 × with CH₂Cl₂, washed with NaHCO₃ (aq), washed with brine, dried over Na₂SO₄, and concentrated. The title compound (368 mg, 88 %) was purified by silica column chromatography (20 % EtOAc in hexanes). ^{1}H NMR (400 MHz, CDCl₃) δ 5.90 (ddt, J = 17.2, 10.4, 5.8 Hz, 1H), 5.33 (dq, J = 17.2, 1.5 Hz, 1H), 5.25 (ddd, J = 10.4, 2.4, 1.2 Hz, 1H), 5.05 (d, J = 7.6 Hz, 1H), 4.68 – 4.59 (m, 2H), 4.49 – 4.36 (m, 1H), 2.36 – 2.21 (m, 1H), 1.44 (d, J = 6.0 Hz, 12H), 1.01 (d, J = 6.9 Hz, 3H), 0.98 (d, J = 6.9 Hz, 3H). ^{13}C NMR (100 MHz, CDCl₃) δ 172.79, 168.91, 131.50, 118.97, 84.84, 65.80, 56.04, 30.15, 28.33, 27.92, 18.74, 18.66, 17.10. HRMS (ESI+) Exact mass calculated for C₁₆H₂₇NNaO₆: 352.1736. Found 352.1736.

Boc-D-Ala-L-HIV-L-Val-SNAC (11). In a round bottom flask, 124 mg of **10** (0.37 mmol, 1 eq.) was dissolved in 2 mL of CH₂Cl₂ under a nitrogen atmosphere. To this solution 43 µL of morpholine (43 mg, 0.49 mmol, 1.3 eq.) and 30 mg of Pd(PPh₃)₄ was added in a single portion. The reaction was stirred at ambient temperature and monitored by TLC. At completion the reaction was quenched by the addition of 10 % aq. HCl, the organic layer was removed and the remaining aqueous fraction was extracted 3 x with CH₂Cl₂. The combined organic fractions were washed with brine, dried over Na₂SO₄ and concentrated in a dry round bottom flask, this intermediate was used immediately in the subsequent reaction. To the flask was added 4 mL of CH₂Cl₂ and 120 mg of L-Val-SNAC (as HCl salt, 0.37 mmol, 1 eq.). To the resulting solution was added 197 µL of Hünig's base (146 mg, 1.12 mmol, 3 eq.), 77 mg HOBt (0.56 mmol, 1.5 eq.) and 109 mg EDC (0.56 mmol, 1.5 eq.). The reaction was stirred under argon at ambient temperature for 20 h. The reaction was quenched with NH₄Cl (aq), extracted 3 × with CH₂Cl₂, washed with NaHCO₃(aq), then with brine, dried over Na₂SO₄, and concentrated. The title compound (192 mg, 94 %) was partially purified by silica column chromatography (40 % acetone in hexanes) and taken into the subsequent reaction. ^{1}H NMR (600 MHz, CDCl₃) δ 7.15 (d, J = 8.6 Hz, 1H), 6.50 (d, J = 29.8 Hz, 1H), 5.47 (dd, J = 41.1, 5.9 Hz, 1H), 5.13 (t, J = 12.6 Hz, 1H), 4.42 (dd, J = 17.2, 9.8 Hz, 1H), 4.30 – 4.16 (m, 1H), 3.36 – 3.27 (m, 2H), 3.00 – 2.88 (m, 3H), 2.27 (td, J = 13.3, 6.6 Hz, 1H), 1.90 (d, J = 1.6 Hz, 4H), 1.40 (d, J = 7.2 Hz, 3H), 1.35 (s, 8H), 0.95 – 0.88 (m, 14H). ^{13}C NMR (150 MHz, CDCl₃) δ 199.98, 172.91, 170.67, 170.01, 155.58, 80.13, 78.28, 64.55, 49.77, 39.05, 30.14, 30.03, 28.41, 28.24, 23.00, 19.38, 18.79, 17.84, 17.43, 16.54.

HO-D-HIC-D-Ala-L-HIV-L-Val-SNAC (12). To a round bottom flask was added 55 mg **11** (0.11 mmol, 1 eq.) in a minimal amount of EtOAc and cooled to 0˚C. To this was added 1 mL of 4M HCl in dioxane (Sigma), and the reaction allowed to warm to ambient temperature. The reaction was monitored by TLC and at

completion all solvent was removed by rotary evaporation. The unpurified intermediate was used immediately and was dissolved in 1 mL of CH_2Cl_2 and to this was added 60 μL of Hünig's base (44 mg, 0.33 mmol, 3 eq.), 17 mg of D-α-hydroxyisocaproic acid (0.12 mmol, 1.1 eq.), 19 mg HOBt (0.16 mmol, 1.5 eq.), and 33 mg EDC (0.16 mmol, 1.5 eq.). The reaction was stirred at ambient temperature overnight and at completion was quenched with NH_4Cl(aq), extracted 5 × with CH_2Cl_2, washed with $NaHCO_3$(aq), then with brine, dried over Na_2SO_4, and concentrated. The title compound (32 mg, 56%) was purified by silica column chromatography (50 % acetone in hexanes). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.28 – 7.25 (m, 1H), 7.22 (d, J = 9.1 Hz, 1H), 6.12 (s, 1H), 5.11 – 5.04 (m, 1H), 4.65 (s, 1H), 4.55 (dd, J = 13.9, 7.0 Hz, 1H), 4.51 – 4.45 (m, 1H), 4.14 (d, J = 5.9 Hz, 1H), 3.49 – 3.35 (m, 2H), 3.35 – 3.20 (m, 1H), 3.14 – 3.04 (m, 1H), 2.84 (dt, J = 7.1, 6.5 Hz, 1H), 2.41 – 2.23 (m, 3H), 1.96 (s, 3H), 1.89 – 1.79 (m, 1H), 1.60 – 1.46 (m, 3H), 1.44 (t, J = 6.6 Hz, 3H), 1.03 – 0.86 (m, 26H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 199.88, 176.11, 172.67, 171.29, 169.99, 78.99, 70.45, 64.46, 48.18, 43.55, 38.72, 30.26, 29.94, 28.58, 24.50, 23.37, 23.09, 21.45, 19.48, 18.84, 18.05, 17.19, 16.74. HRMS (ESI+) Exact mass calculated for $C_{23}H_{41}N_3NaO_7S$: 526.2563. Found 526.2558.

Protein expression

The CesB TE gene was generated synthetically and cloned into pET21 (pGWH43). Expression conditions are modified from Alonzo et al.,[15] briefly, the CesB TE was expressed in *E. coli* BL21 Rosetta grown in LB media supplemented with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. Cells were induced at an OD_{600} = 0.45 with 0.1 mM IPTG and further incubated at 16 °C for 16 hours. Cells were harvested by centrifugation and resuspended in buffer A (50 mM Tris pH 8.0, 300 mM NaCl, 10% v/v glycerol) and lysed by sonication. The lysate was centrifuged at 10000 rpm for 45 minutes and the clarified supernatant was incubated with 0.75 mL Ni-NTA superflow resin (QIAGEN, Valenica, CA) at 4 °C for 1 hour. The resin was loaded into a column and first washed with buffer A supplemented with 20 mM imidazole then CesB TE was eluted with buffer A supplemented with 150 mM imidazole. The purified protein was exchanged into dialysis buffer (25 mM HEPES, 100 mM NaCl, pH 8) and concentrated by centrifugation (Amicon 10000 MWCO). The concentrated protein was flash frozen and stored at –78 °C. Protein concentration was determined by the Bradford assay (Bio-rad). Approximately 8 mg of purified protein was obtained per L of cell culture.

Enzyme assays

Enzymatic assays were carried out in 50 mM Tris (pH7.4) at 22 °C, with 15 μM Ces TE. Total substrate concentration was 2 mM (50mM stock solutions in DMSO) for single substrate assays and 3 mM total substrate for mixed assays, and DMSO up to 10% solution volume if necessary. At completion (14 h) assays were quenched with an equal volume of 0.5% formic acid in acetonitrile before analysis by LCMS. Substrate mix and match assays were carried out as above but with 2 mM **12** and 1 mM of the appropriate non-native substrate **7** or **16**. LCMS analysis was carried out on a Shimadzu LC20A HPLC coupled with a Shimadzu LCMS-2020 using a Thermo Scientific Hypersil C18 column (100 x 2.1 mm). The HPLC conditions was 0.4 mL/min with a linear gradient from 5% B to 95% B over 5 minutes, holding at 95% B for 6 minutes, then returning to 5% B and reequilibrating for 2 minutes. Column oven was set to 40 °C and solvent composition was A: H_2O + 0.05% formic acid; B: MeCN + 0.05% formic acid.

Supporting information: Supporting information is available containing additional synthetic procedures, ^{1}H and ^{13}C NMR spectra, detailed protein purification protocols, and kinetic characterization.

References

1. Agata, N. *et al.* A novel dodecadepsipeptide, cereulide, isolated from Bacillus cereus causes vacuole formation in HEp-2 cells . *FEMS Microbiol. Lett.* **121**, 31–34 (1994).

2. Agata, N., Ohta, M., Mori, M. & Isobe, M. A novel dodecadepsipeptide, cereulide, is an emetic toxin of Bacillus cereus. *FEMS Microbiol. Lett.* **129**, 17–20 (1995).

3. Mikkola, R., Saris, N. E. L., Grigoriev, P. A., Andersson, M. A. & Salkinoja-Salonen, M. S. Ionophoretic properties and mitochondrial effects of cereulide. The emetic toxin of B. cereus. *Eur. J. Biochem.* **263**, 112–117 (1999).

4. Shinagawa, K., Konuma, H., Sekita, H. & Sugii, S. Emesis of rhesus monkeys induced by intragastric administration with the HEp-2 vacuolation factor (cereulide) produced by Bacillus cereus. *FEMS Microbiol. Lett.* **130**, 87–90 (1995).

5. Shinagawa, K., Ueno, Y., Hu, D., Ueda, S. & Sugii, S. Mouse lethal activity of a hep-2 vacuolation factor, cereulide, produced by Bacillus cereus isolated from vomiting-type food poisoning. *J. Vet. Med. Sci.* **58**, 1027–1029 (1996).

6. Schoeni, J. L. & Lee Wong, A. C. Bacillus cereus food poisoning and its toxins. *Journal of Food Protection* **68**, 636–648 (2005).

7. Ehling-Schulz, M., Fricker, M. & Scherer, S. Bacillus cereus, the causative agent of an emetic type of food-borne illness. *Molecular Nutrition and Food Research* **48**, 479–487 (2004).

8. Agata, N., Ohta, M. & Yokoyama, K. Production of Bacillus cereus emetic toxin (cereulide) in various foods. *Int. J. Food Microbiol.* **73**, 23–27 (2002).

9. Toh, M. *et al.* Cereulide, the emetic toxin of Bacillus cereus, is putatively a product of nonribosomal peptide synthesis. *J. Appl. Microbiol.* **97**, 992–1000 (2004).

10. Horwood, P. F., Burgess, G. W. & Jane Oakey, H. Evidence for non-ribosomal peptide synthetase production of cereulide (the emetic toxin) in Bacillus cereus. *FEMS Microbiol. Lett.* **236**, 319–324 (2004).

11. Ehling-Schulz, M. *et al.* Identification and partial characterization of the nonribosomal peptide synthetase gene responsible for cereulide production in emetic Bacillus cereus. *Appl. Environ. Microbiol.* **71**, 105–113 (2005).

12. Ehling-Schulz, M. *et al.* Cereulide synthetase gene cluster from emetic Bacillus cereus: Structure and location on a mega virulence plasmid related to Bacillus anthracis toxin plasmid pXO1. *BMC Microbiol.* **6**, (2006).

13. Rasko, D. A. *et al.* Complete sequence analysis of novel plasmids from emetic and periodontal Bacillus cereus isolates reveals a common evolutionary history among the B. cereus-group

plasmids, including Bacillus anthracis pXO1. *J. Bacteriol.* **189**, 52–64 (2007).

14. Magarvey, N. A., Ehling-Schulz, M. & Walsh, C. T. Characterization of the cereulide NRPS alpha-hydroxy acid specifying modules: activation of alpha-keto acids and chiral reduction on the assembly line. *J. Am. Chem. Soc.* **128**, 10698–9 (2006).

15. Alonzo, D. A., Magarvey, N. A. & Schmeing, T. M. Characterization of cereulide synthetase, a toxin-producing macromolecular machine. *PLoS One* **10**, (2015).

16. Alonzo, D. A., Chiche-lapierre, C., Tarry, M. J., Wang, J. & Schmeing, T. M. Structural basis of keto acid utilization in nonribosomal depsipeptide synthesis. *Nat. Chem. Biol.* (2020). doi:10.1038/s41589-020-0481-5

17. Jaitzig, J., Li, J., Süssmuth, R. D. & Neubauer, P. Reconstituted biosynthesis of the nonribosomal macrolactone antibiotic valinomycin in Escherichia coli. *ACS Synth. Biol.* **3**, 432–8 (2014).

18. Huguenin-Dezot, N. *et al.* Trapping biosynthetic acyl-enzyme intermediates with encoded 2,3-diaminopropionic acid. *Nature* **565**, 112–117 (2019).

19. Marxen, S. *et al.* Depsipeptide intermediates interrogate proposed biosynthesis of cereulide, the emetic toxin of Bacillus cereus. *Sci. Rep.* **5**, (2015).

20. Pitchayawasin, S. *et al.* Molecular diversity of cereulide detected by means of nano-HPLC-ESI-Q-TOF- MS. *Int. J. Mass Spectrom.* **235**, 123–129 (2004).

21. Stark, T. *et al.* Mass spectrometric profiling of Bacillus cereus strains and quantitation of the emetic toxin cereulide by means of stable isotope dilution analysis and HEp-2 bioassay. *Anal. Bioanal. Chem.* **405**, 191–201 (2013).

22. Marxen, S. *et al.* Chemodiversity of cereulide, the emetic toxin of Bacillus cereus. *Anal. Bioanal. Chem.* **407**, 2439–2453 (2015).

23. Horsman, M. E., Hari, T. P. A. & Boddy, C. N. Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? *Nat. Prod. Rep.* **33**, 183–202 (2016).

24. Boddy, C. N., Schneider, T. L., Hotta, K., Walsh, C. T. & Khosla, C. Epothilone C macrolactonization and hydrolysis are catalyzed by the isolated thioesterase domain of epothilone polyketide synthase. *J. Am. Chem. Soc.* **125**, 3428–3429 (2003).

25. Horsman, M. E., Hari, T. P. A. & Boddy, C. N. Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: Logic gate or a victim of fate? *Natural Product Reports* **33**, 183–202 (2016).

26. Trauger, J. W., Kohli, R. M., Mootz, H. D., Marahiel, M. A. & Walsh, C. T. Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **407**, 215–218 (2000).

27. Hoyer, K. M., Mahlert, C. & Marahiel, M. A. The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization. *Chem. Biol.* **14**, 13–22 (2007).