Biosynthetic characterization of bacillibactin in thermophilic Bacillaceae and its potential for in vitro mutational synthesis

Momona Izumi[a], Hiroya Tomita*[b,c], Kentaro Miyazaki[b], Ryo Otsuka[a], and Kohsuke Honda[b,c]

[a] Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan
[b] International Center for Biotechnology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan
E-mail: tomita.icb@osaka-u.ac.jp
[c] Industrial Biotechnology Initiative Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

Abstract:
Bacillibactin (BB) is a microbial siderophore produced by Bacillus species. BB is biosynthesized from 2,3-dihydroxybenzoic acid (2,3-DHB), Gly, and L-Thr by nonribosomal peptide synthetase (NRPS) enzymes DhbE, DhbB, and DhbF. The biosynthetic gene cluster (dhb) is also conserved in some strains of thermophilic genera, Geobacillus, Anoxybacillus and Parageobacillus. However, the production of BB from these thermophilic bacteria has not been characterized. Here, we report in vivo and in vitro characterization of BB biosynthesis in Parageobacillus sp. KH3-4 which grows at 65°C. We confirmed BB production in this thermophilic bacterium and the gene cluster active. In vitro enzymatic analysis revealed that 4'-phosphopantetheinyltransferase (PPTase) encoded in the same gene cluster is responsible for the post-translational maturation of carrier proteins. DhbE and DhbF showed substrate preference to 2,3-DHB and Gly and L-Thr, respectively, consistent with the chemical structure of BB. With the purified enzymes, we successfully reconstituted the NRPS assembly line in vitro. In addition, using chemically synthesized acyl-N-acetylcysteamine substrate analogues, BB analogues possessing methylbenzoyl groups instead of 2,3-DHB were detected. This study provides a new insight into secondary metabolism in thermophiles, and it expands the temperature limitation of NRPS enzymes.

Introduction

Iron is an essential element for organisms, playing a pivotal role in numerous enzymatic functions, e.g. heme and [Fe-S] clusters.[1,2] However, the availability of biologically accessible iron in nature is often not sufficient, because ferrous ion is easily oxidized to ferric ion, and they predominantly exist as insoluble ferric hydroxide.[3,4] To effectively acquire the requisite amounts of iron from the environment, microorganisms produce siderophores which chelate iron and facilitate its transport into the cells. Siderophores are classified depending on their chemical structures, with catecholic and

hydroxamate types being the most represented.[4,5] Both are biosynthesized by nonribosomal peptide synthetase (NRPS), which can produce a variety of peptides.[4,6] As NRPS is independent of ribosome, products of NRPS can include non-proteinogenic amino acids.[7]

Bacillibactin (BB) is a catecholic siderophore produced by various species of the genus Bacillus.[6,8,9] Its structure exhibits three-fold-symmetry, with each unit comprising 2,3-dihydroxybenzoic acid (2,3-DHB), Gly, and L-Thr (Figure 1). The biosynthetic gene cluster (dhb) possesses three NRPS genes: dhbE, dhbB, and dhbF, for compound backbone assembly, and other genes are responsible for 2,3-DHB production and accessory proteins. DhbE is an adenylation (A) domain specifically activating 2,3-DHB to its adenylated intermediate using ATP. This activated 2,3-DHB is subsequently transferred to the carrier protein (CP) domain of DhbB. These two domains organize a loading module. DhbF possesses the following two modules (module 1 and module 2), and both are composed of a condensation (C) domain, an A domain, and a CP domain. The substrates of modules 1 and 2 are Gly and L-Thr, respectively. The assembled 2,3-DHB-Gly-L-Thr unit is anchored onto a thioesterase (TE) domain within DhbF, and TE ultimately releases BB via trimeric cyclization.

Although there are various definitions, thermophiles are generally understood as microorganisms with an optimal growth temperature exceeding 45°C. To date, despite the identification of some bioactive compounds within this group,[10] little is known about secondary metabolism of thermophiles. Recent studies have highlighted thermophilic Bacillaceae, particularly the genera Geobacillus and Parageobacillus, as promising subjects for investigation.[11] Notably, it has been proposed that species within Geobacillus possess the capability to produce secondary metabolites.[12,13] Among them, geobacillin peptides have been identified, showing antibiotic activity.[14-16]

Here, we report the characterization of BB biosynthesis in the thermophilic bacterium Parageobacillus sp. KH3-4, which was isolated by cultivating the cells at 65°C.[17] In the genome database, we found the putative BB biosynthetic gene clusters in some Geobacillus, Anoxybacillus, and Parageobacillus species. The growth temperature of these bacteria is higher than that of Bacillus species which have been reported to produce BB. We confirmed BB production in Parageobacillus sp. KH3-4 under an iron starvation condition. We also carried out the enzymatic characterization of BB biosynthesis and mutational synthesis (mutasynthesis) of BB analogous compounds using chemically synthesized substrates in vitro. This study proves BB production in thermophilic bacteria for the first time.

Results and Discussion

BB biosynthetic gene cluster in Parageobacillus sp. KH3-4

We conducted an antiSMASH analysis[18] to identify secondary metabolite biosynthetic gene clusters in thermophilic bacterial strains belonging to the genera Geobacillus, Anoxybacillus, and Parageobacillus. As a result, we found that some species have a putative dhb cluster similar to that in

Bacillus subtilis. In this study, we focused on the gene cluster in Parageobacillus sp. KH3-4 (GenBank accession number AP025627),[17] which contains the genes for 2,3-DHB production from chorismic acid (dhbACB) and NRPS for BB assembly (dhbEBF) (Figure 1). Regarding BB assembly, the gene organization was identical to that of B. subtilis. The loading module is bifurcated between DhbE (A domain) and DhbB (CP domain), and the subsequent modules 1 and 2 are located on DhbF (C-A-CP-C-A-CP-TE). DhbB is a fusion protein of isochorismatase and CP domain. Additionally, this gene cluster encodes two accessory proteins for NRPS: a 4'-phosphopantetheinyltransferase PPTase for the 4'-phosphopantetheinylation of CP domains and an auxiliary protein MbtH for NRPS function.[19] Collectively, this gene organization strongly suggests the biosynthetic ability of this strain to produce BB.

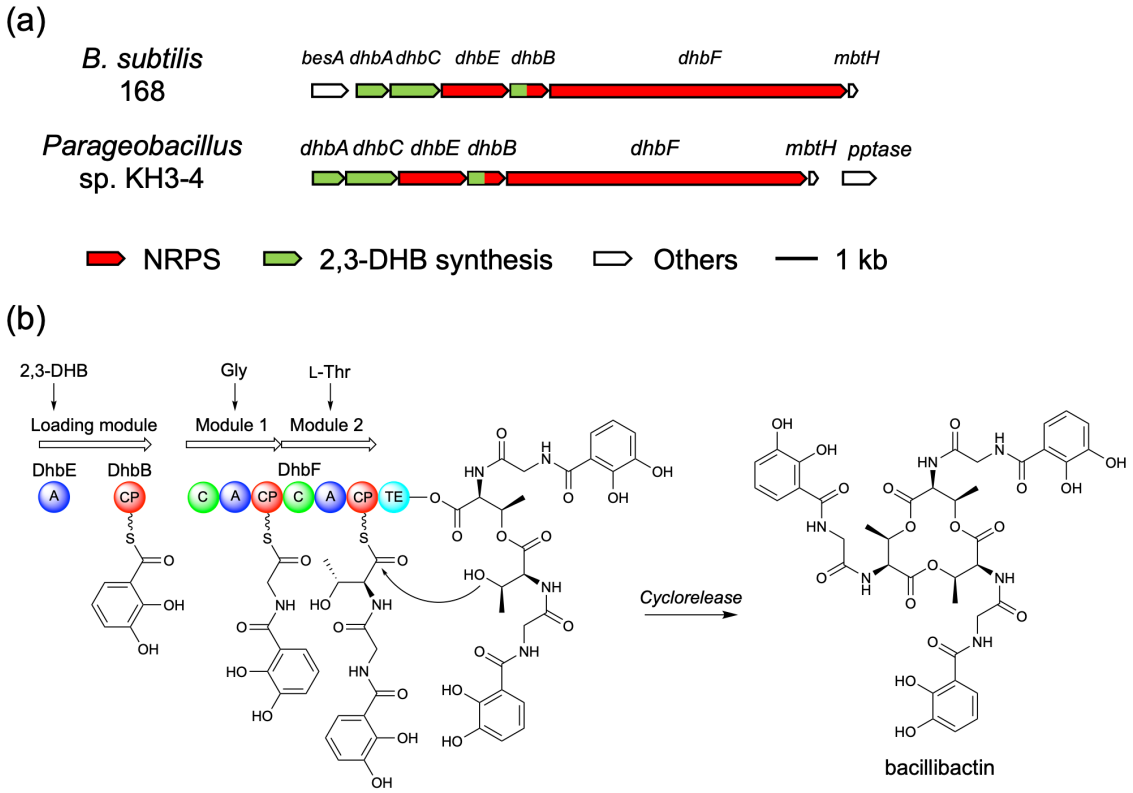


Figure 1. Gene organization of the dhb cluster. (a) Comparison of the BB biosynthetic gene clusters in B. subtilis 168 and Parageobacillus sp. KH3-4. Genes involved in peptide assembly (NRPS) and 2,3-DHB synthesis are shown in red and green, respectively. (b) Domain organization of NRPS and the biosynthetic pathway of BB. Three modules of NRPS contribute to the assembly of the 2,3-DHB-Gly-L-Thr unit, and BB is released through a trimeric cyclization process. Domain abbreviations: C, condensation domain; A, adenylation domain; CP, carrier protein; and TE, thioesterase domain.

BB production in B. subtilis and Parageobacillus sp. KH3-4

To ascertain BB production in Parageobacillus sp. KH3-4, we precultivated B. subtilis and Parageobacillus sp. KH3-4 in Luria-Bertani (LB) medium. The preculture was inoculated at a 1% volume into the BB production medium which does not contain iron. Cells exhibited robust growth, likely due to an overcarrying of residual iron from the pre-culture. After centrifugation to remove the cells, ethyl acetate was added to the supernatants to extract BB. The solvent was then evaporated, and the residue was analyzed by Liquid Chromatography-Electrospray Ionization-Mass Spectrometry (LC-ESI-MS).

In both B. subtilis and Parageobacillus sp. KH3-4, we observed an m/z value corresponding to BB with the identical retention times (Figure 2ab). These results indicate that Parageobacillus sp. KH3-4 possess the capability to produce BB and that the gene cluster is actively expressed.

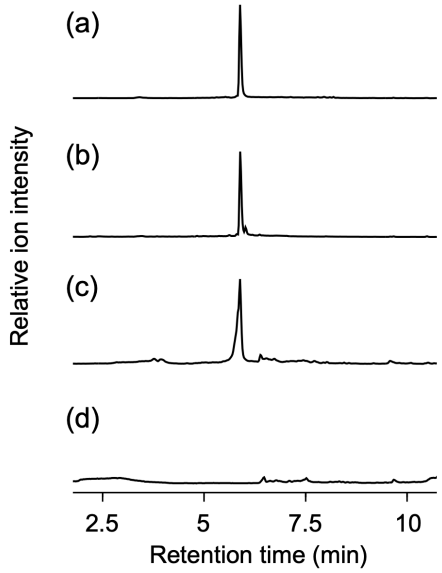


Figure 2. BB production analysis. Extracted ion chromatograms (m/z = 881.2) corresponding to [M-H]- of BB are shown. (a) B. subtilis and (b) Parageobacillus sp. KH3-4 were cultivated in BB production medium, and compounds were extracted using ethyl acetate. (c) In vitro reconstitution of BB biosynthesis with substrates and proteins (DhbE, holo-DhbB, holo-DhbF, and MbtH). After the reactions, compounds were extracted using ethyl acetate. (d) Negative control of in vitro reconstitution without holo-DhbF.

4'-Phosphopantetheinylation of DhbB by PPTase

Due to the scarcity of secondary metabolites from thermophiles, the enzymatic profiles of thermophilic NRPS proteins are of considerable interest. First, we aimed to investigate the post-translational modification of CP domains by PPTase. In NRPS, polyketide synthase (PKS), and fatty acid biosynthesis, a specific Ser residue in carrier proteins undergo 4'-phosphopantetheinylation to be a

holo-form using coenzyme A (CoA) as a substrate.[20] This reaction is catalyzed by PPTases, which are often encoded within the same biosynthetic gene cluster.

We identified a PPTase gene in the dhb cluster in Parageobacillus sp. KH3-4 and examined the ability of the enzyme to attach a 4'-phosphopantetheinyl (ppant) group to DhbB and DhbF. Recombinant DhbB and PPTase proteins were prepared using Escherichia coli as the expression host (Figure S1), and an in vitro 4'-phosphopantetheinylation reaction was performed in the presence of CoA. The resulting changes in molecular mass were analyzed by Matrix Assisted Laser Desorption/Ionization-Time Of Flight-Mass Spectrometry (MALDI-TOF-MS). In this experiment, DhbF was excluded due to its substantial molecular mass (~272 kDa), which posed challenges for effective ionization.

It was revealed that, compared to the control, the molecular mass of DhbB increased by approximately 340 Da by the PPTase reaction (Figure 3). This increase corresponds to the addition of a ppant group, indicating that DhbB was initially expressed without a ppant group (apo-DhbB) in E. coli, and was subsequently converted to a holo-form (holo-DhbB) through the action of PPTase encoded in the dhb cluster.

In E. coli, the ent cluster for the biosynthesis of another siderophore enterobactin (EB) is known.[4] E. coli BL21(DE3), the host strain used for heterologous protein expression in this study, also possesses the ent cluster. The chemical structure of EB is also a three-fold symmetric, and each unit is composed of 2,3-DHB and L-Ser, thus it is quite similar to BB. The biosynthetic mechanisms of EB and BB are also similar, and the ent cluster encodes a PPTase EntD. Therefore, we first assumed that EntD can 4'-phosphopantetheinylate DhbB and DhbF as the substrate proteins. However, in our experiments, DhbB was not 4'-phosphopantetheinylated in E. coli. As the ent cluster is reportedly not active in LB medium,[21] EntD might not be expressed in the cells. Therefore, while we could not mention about the ability of EntD to 4'-phosphopantetheinylate DhbB and DhbF, PPTase from the dhb cluster is necessary for the maturation of DhbB and DhbF in this experimental condition.

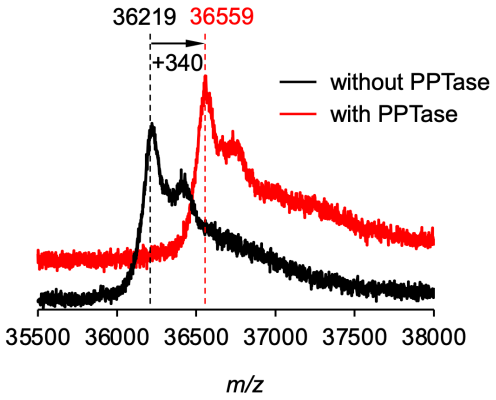


Figure 3. MALDI-TOF-MS analysis of 4'-phosphopantetheinylation of DhbB by PPTase. Black and red lines indicate the reaction products without and with PPTase, respectively. Sinapic acid was used

as the matrix. Calculated molecular mass of DhbB is 36,269.

Substrate specificity analysis of A domains in DhbE and DhbF

To evaluate the enzymatic profiles of NRPS proteins, we conducted in vitro reactions of A domains present in DhbE and DhbF. holo-DhbF was prepared by coexpression with PPTase (Figure S1). Using these proteins, we assessed enzyme activity of the A domains by examining their reaction temperature dependence and substrate specificity.

As A domains produce pyrophosphate (PPi) as a bioproduct of substrate adenylylation, we quantified PPi to evaluate enzyme activity. For this purpose, we employed a thermophilic pyrophosphatase derived from the thermophilic bacterium Thermus thermophilus HB8 to hydrolyze PPi into two equivalents of phosphate (Pi). The Malachite Green reagent binds to Pi, resulting in the formation of a green pigment, and it allowed for colorimetric quantification of the Pi concentrations.

First, the reactions were carried out at 30, 40, 50, 60, and 70°C with the natural substrates to examine temperature dependence. As a result, DhbE showed the highest activity at 60°C, demonstrating its functionality at high temperature (Figure S2). In contrast, the highest activity of DhbF was obtained at 30°C. We speculated that the diminished activity of DhbF at higher temperatures is due to its instability, as we observed protein aggregation at 60°C (data not shown). To assess this in detail, circular dichroism (CD) spectroscopy was performed on DhbF at various temperatures. The results indicate that DhbF maintained its protein secondary structures at 55°C or below but exhibited slight thermal denaturation at 60°C (Figure 4).

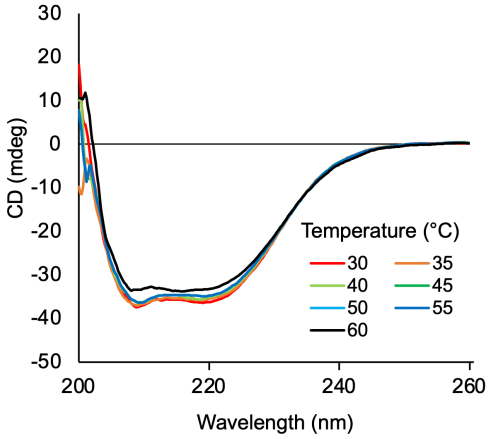


Figure 4. CD spectra in the far UV (200-260 nm) for holo-DhbF at various temperatures. holo-DhbF (1 μM) was incubated at the indicated temperatures for 10 min prior to analysis. mdeg, millidegrees.

We next investigated substrate specificity. DhbE showed higher activity with its natural substrate 2,3-DHB compared to 3,4-DHB and 2-/3-/4-hydroxybenzoic acids (HB) (Figure 5a), consistent with the chemical structure of BB. In the case of DhbF, higher activity was obtained with Gly and L-Thr

compared to the other amino acids tested (Figure 5b), because DhbF harbors two A domains corresponding to the recognition of Gly and L-Thr (Figure 1). To further analyze the functions of the A domains separately, we prepared truncated holo-proteins, DhbF1 and DhbF2, corresponding to the first and second modules of the Dhb system, respectively (Figure S1). Using these enzymes, we evaluated substrate specificity again. We found that both DhbF1 and DhbF2 precipitated at 60°C (data not shown), presumably due to structural instability. Consequently, the reactions were carried out at 30°C. As we expected, DhbF1 and DhbF2 showed the highest activity against Gly and L-Thr, respectively (Figure 5cd). Altogether, our results were in agreement with the chemical structure of BB and a previous report on BB biosynthesis in B. subtilis.6

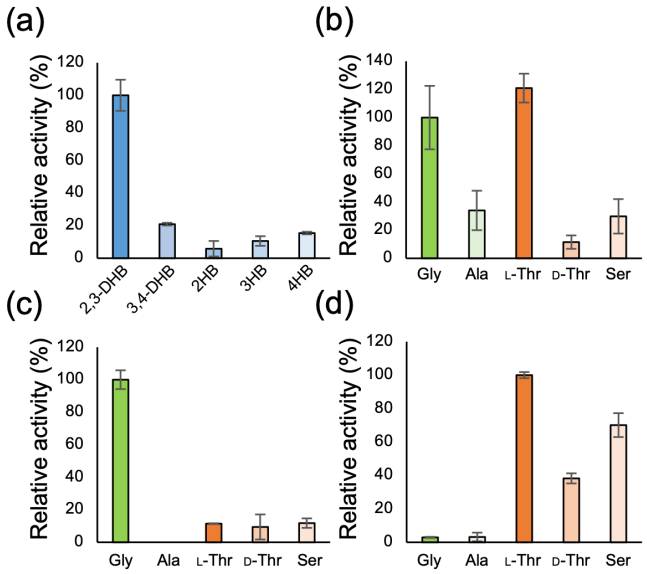


Figure 5. Substrate specificity of A domains of (a) DhbE, (b) DhbF, (c) DhbF1, and (d) DhbF2, as determined by colorimetric analysis. All reactions contained 1.5 mM substrate and the respective enzymes, and they were incubated for 1 h at 60°C (DhbE and DhbF) or 30°C (DhbF1 and DhbF2). The produced pyrophosphate was quantified using pyrophosphatase and the Malachite Green reagent. In vitro reconstitution of BB biosynthesis

Based on the aforementioned results, we conducted the in vitro reconstitution of the entire NRPS assembly line using DhbE, holo-DhbB, holo-DhbF, and MbtH (Figure S1). The enzymes were incubated with 2,3-DHB, Gly, L-Thr, and ATP at temperatures ranging from 30°C to 60°C. As a result, we observed an m/z value corresponding to BB, with a retention time identical to that of BB extracted from B. subtilis and Parageobacillus sp. KH3-4 (Figure 2). The highest BB production was observed at 45°C, whereas virtually no BB was observed at 60°C (Figure 6a). Notably, aside from BB, additional peaks were found in the UV chromatogram, which were not observed in the control (Figure 6b). Some

of these compounds exhibited the same m/z values with the monomer and dimer of the 2,3-DHB-Gly-L-Thr unit (Figure S3). Such intermediates have been reported as hydrolyzed products of BB by the intracellular enzyme YuiI in natural homeostasis.[22] However, our in vitro system lacks such enzymes involved in backward reactions, so the reason for the occurrence of the peaks remain unclear. It is likely that the intermediates anchored on the TE domain were partially hydrolyzed. While these unexpected peaks hindered the precise measurement of yield or the production efficiency, the production of BB supports our successful reconstitution of the NRPS assembly line for BB biosynthesis. Efforts will be made to optimize the reaction conditions for more efficient BB production. Based on the concentration of 2,3-DHB after the reactions, approximately 30% of the substrates was consumed.

To the best of our knowledge, NRPS reported to function at the highest temperature to date is associated with fuscachelin biosynthesis in the thermophilic bacterium Thermobifida fusca.[23] In this study, the siderophore fuscachelin was produced by cultivating the cells at 55°C, demonstrating that the Fsc NRPS enzymes are active at this temperature. However, it is noteworthy that the in vitro enzymatic evaluation was carried out at 30°C. In contrast, in our study, Parageobacillus sp. KH3-4 was cultivated and confirmed to produce BB at 65°C, and BB can be produced at 45°C in vitro. Therefore, our NRPS works at the highest temperature to date.

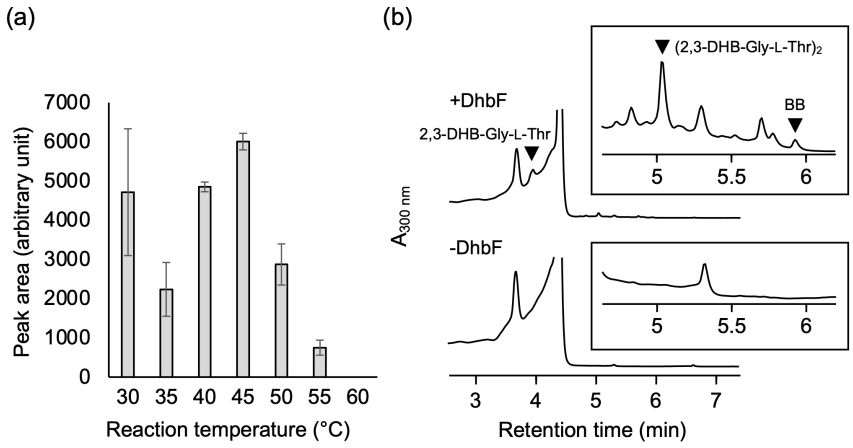


Figure 6. In vitro reconstitution of BB biosynthesis. (a) Effects of reaction temperature on BB production. (b) UV chromatogram. Arrowheads indicate the monomer and dimer of the 2,3-DHB-Gly-L-Thr unit and BB. The corresponding mass spectra are shown in Figure S3.

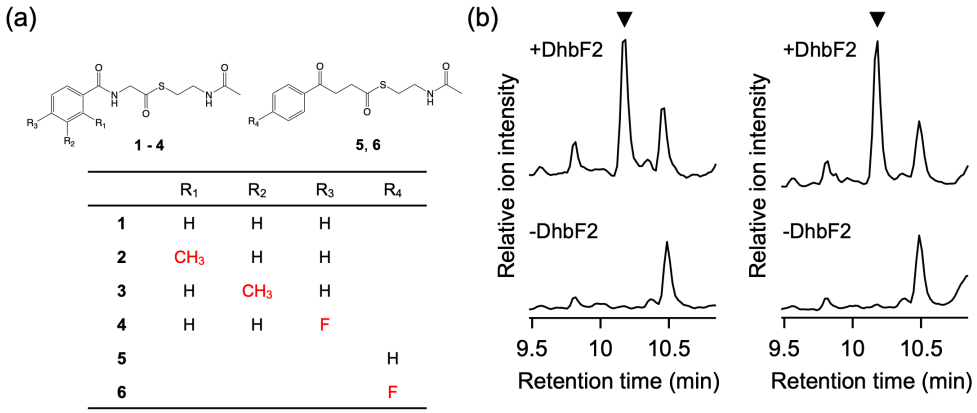
In vitro mutational synthesis of BB analogues

In vitro reactions provide the advantage of utilizing unnatural substrates independent of cell permeability and toxicity, thereby expanding the substrate and chemical libraries. It is established that acyl-N-acetylcysteamine (acyl-SNAC) thioesters can serve as the substrate analogues for C domains

of NRPS, substituting the natural substrates, acyl-CPs. This approach facilitates the evaluation of the enzymatic profiles of C domains and allows for the generation of analogous products through mutasynthesis.[24]

We chemically synthesized six acyl-SNAC thioesters: benzoyl (BA)-Gly-SNAC (1), 2/3-methyl-BA-Gly-SNAC (2 and 3), 4-fluoro-BA-Gly-SNAC (4), 3-benzoylpropionyl-SNAC (5), and 3-(4-fluorobenzoyl)propionyl-SNAC (6) (Figure 7a, Figure S5-S10). These compounds mimic the intermediate 2,3-DHB-Gly anchored on the CP domain of module 1. Consequently, DhbF2 was incubated with L-Thr, ATP, and each of these thioesters. After compound extraction with ethyl acetate, the production of the corresponding BB analogues was examined by LC-ESI-MS. As a result, with 2 and 3, we observed the m/z values of the BB analogues which harbor methylbenzoyl groups in place of 2,3-DHB (Figure 7b). MS/MS fragmentation analysis further supported these (Figure S4). This suggests that these SNAC thioesters were accepted by DhbF2 to produce the BB analogues. However, the concentrations of the products were below the detection limits of our UV detector. In contrast, the other four SNAC thioesters could not be recognized as the substrates. It is likely that specific substitutions on the benzene ring are required for substrate recognition, as 1 and 5, which contain a BA group, were not accepted. Additionally, the presence of a 4-fluorine group in 4 and 6 rendered them unacceptable.

Thermostable enzymes are physically and chemically stable and also robust against amino acid substitution and have been proposed as a suitable starting points for protein engineering.[25,26] Our thermophilic proteins would serve as templates for enzyme engineering to synthesize various BB analogues using SNAC thioesters. The native strain produces BB at 65°C, whereas our in vitro reconstitution system produced BB (and its analogues) at lower temperatures. Filling this temperature gap through protein engineering would further enhance the potential for in vitro mutational synthesis of the current BB synthesis platform.



	R_1	R_2	R_3	R_4
1	H	H	H	
2	CH_3	H	H	
3	H	CH_3	H	
4	H	H	F	
5				H
6				F

Figure 7. (a) Chemically synthesized acyl-SNAC thioesters. Compounds 1-4 contain a benzoic acid

(BA) and a Gly moieties. Compounds 5 and 6 have a 3-benzoylpropionic acid moiety, where the amino group of Gly is substituted with a methylene group. (b) In vitro mutasynthesis using compounds 2 and 3 as the substrates. Extracted ion chromatogram (m/z = 829.3) corresponding to [M+H]+ of BB analogues using 2 (left) and 3 (right) are shown. The products are indicated by arrowheads. MS/MS fragmentation analyses are presented in Figure S4.

Conclusion

In summary, we identified the dhb cluster in Parageobacillus sp. KH3-4 and confirmed it to be active. Using the recombinant NRPS enzymes, we reconstituted the NRPS assembly line in vitro. In addition, we demonstrated the potential for a new mutasynthetic approach to produce BB analogues. This NRPS system works at the highest temperature among NRPS enzymes reported to date. Altogether, this study provides new insights into the microbial survival strategy in the thermal environments, and it explores the temperature limitation and utility of NRPS enzymes.

Experimental Section

Strains, media, and culture conditions

Unless mentioned otherwise, all chemicals were purchased from Nacalai Tesque (Kyoto, Japan) or FUJIFILM Wako Pure Chemical Corporation (Osaka, Japan). E. coli JM109 and BL21(DE3) were used for plasmid construction and heterologous gene expression, respectively. JM109 was cultivated at 37°C in LB medium, and BL21(DE3) was cultivated at 37 or 30°C in LB medium or terrific broth. Terrific broth contains yeast extract (24 g L-1), tryptone (12 g L-1), K2HPO4 (9.4 g L-1), KH2PO4 (2.2 g L-1), and glycerol (4 mL L-1). Cultivation was carried out with or without ampicillin (100 mg L-1).

B. subtilis ISW1214 and Parageobacillus sp. KH3-4 were cultivated at 37°C and 65°C, respectively, in LB medium or BB production medium (iron starvation medium). BB production medium contains glucose (5 g L-1), casamino acid (3 g L-1) (Thermo Fisher Scientific, USA), (NH4)2HPO4 (1 g L-1), K2HPO4 (2.5 g L-1), KH2PO4 (2.5 g L-1), Gly (0.1 g L-1), L-Thr (0.1 g L-1), nicotinic acid (40 μM), thiamine hydrochloride (100 μM), MnSO4 (36 μM), ZnSO4 (0.3 μM), and MgSO4 (830 μM).

Production, extraction, and analysis of BB

The glasswares were washed beforehand with 6 N HCl to eliminate iron. B. subtilis and Parageobacillus sp. KH3-4 were pre-cultured in LB medium at 30°C and 65°C for 18 h, respectively. Cells were cultivated in a 500 mL Erlenmeyer flask containing 100 mL of BB production medium at 30°C (B. subtilis) and 65°C (Parageobacillus sp. KH3-4) for 48 h, respectively. The flasks were capped with air-permeable silicon rubber plugs and orbitally shaken at 180 rpm. After centrifugation (room temperature, 5,000 × g, 10 min), the supernatant was collected, and an equal volume of ethyl acetate

was added to extract compounds. The organic layer was collected, washed by water and brine, and evaporated in vacuo. The residue was dissolved in a small amount of methanol and analyzed by LC-ESI-MS. Compounds were separated using a COSMOCORE 2.6C18 column (2.1 mm I.D. × 150 mm) (Nacalai Tesque) with a linear gradient (5 to 100%) of acetonitrile containing 0.1% (v/v) formic acid as the mobile phase at a flow rate of 0.4 mL min-1.

Plasmid construction

pET-32a(+) and pETDuet-1 were purchased from Merck (Germany). All PCR reactions were carried out using KOD One (TOYOBO, Japan) following the manufacturer's instructions. For pET32a-dhbE, pET32a-dhbB, pET32a-pptase, and pET32a-mbtH, the coding regions of dhbE (GenBank accession number BDG45613), dhbB (BDG45614), pptase (BDG45617), and mbtH (BDG45616) were amplified from the genomic DNA of Parageobacillus sp. KH3-4 by PCR using the primer sets dhbE-F/R, dhbB-F/R, pptase-F/R, and mbtH-F/R, respectively (Table S1). pET32a was linearized by PCR using the primer set pET32a-F1/R1 (Table S1). The DNA fragments of pET-32a(+) and each gene were applied to Gibson assembly using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, USA) to obtain pET32a-dhbE, pET32a-dhbB, pET32a-pptase, and pET32a-mbtH.

For pET32a-ppase, the coding region of ppase (BAD71788) was amplified from the genomic DNA of Thermus thermophilus HB8 (laboratory stock) using the primer set ppase-F/R (Table S1). pET32a was linearized using the primer set pET32a-F2/R2 (Table S1), and the DNA fragments were applied to Gibson assembly to obtain pET32a-ppase.

For pETDuet-1-dhbB-pptase, pETDuet-1-dhbF-pptase, pETDuet-1-dhbF1-pptase, and pETDuet-1-dhbF2-pptase, pETDuet-1 was linearized by digestion by NcoI (New England Biolabs). The coding region of pptase was amplified from the genomic DNA of Parageobacillus sp. KH 3-4 using the primer set Duet-MCS1-pptase-F/R (Table S1). After Gibson assembly using the two DNA fragments, pETDuet-1-ppt was obtained. Likewise, the plasmid was linearized by digestion by NdeI (New England Biolabs), and the coding regions of dhbB, dhbF (BDG45615), dhbF1, dhbF2 were individually amplified using Duet-MCS2-dhbB-F/R, Duet-MCS2-dhbF-F/R, Duet-MCS2-dhbF1-F/R, and Duet-MCS2-dhbF2-F/R, respectively (Table S1). Each gene was introduced into pETDuet-1-ppt to obtain pETDuet-1-dhbB-pptase, pETDuet-1-dhbF-pptase, pETDuet-1-dhbF1-pptase, and pETDuet-1-dhbF2-pptase, respectively.

Production and purification of the recombinant proteins

For apo-DhbB, DhbF1, DhbF2, PPTase, and MbtH, the plasmids pET32a-dhbB, pETDuet-1-dhbF1-pptase, pETDuet-1-dhbF2-pptase, pET32a-pptase, and pET32a-mbtH were individually introduced into E. coli BL21(DE3). Cells were cultivated at 37°C in TB medium containing 100 mg L-1 of ampicillin until OD600 reached ~0.5. After cooling to room temperature, gene expression was induced

with isopropyl 1-thio-β-D-galactopyranoside (IPTG) at a final concentration of 0.1 mM, and cells were cultivated for a further 16 h at 30°C. The cells were harvested by centrifugation (4°C, 5,000 ×g, 10 min), resuspended in Lysis buffer (50 mM HEPES-NaOH (pH 7.5) containing 150 mM NaCl and 10% (v/v) glycerol), and disrupted by sonication. After centrifugation (4°C, 20,000 ×g, 10 min), His60 Ni Superflow Resin (Takara Bio, Japan) was added to the soluble cell extracts and mixed gently at 4°C for 30 min. After the His-tagged proteins were eluted by a stepwise gradient of imidazole (0-500 mM) in Lysis buffer, the fractions containing the pure objective proteins were combined. The buffer was exchanged to Lysis buffer using a PD-10 column (Cytiva, USA).

For holo-DhbF, pETDuet-1-dhbF-pptase was introduced into E. coli BL21(DE3). Cultivation and gene expression induction were carried out as described above. After cell disruption and centrifugation, the soluble cell extract was incubated at 60°C for 30 min. After removing thermolabile proteins by centrifugation (4°C, 20,000 ×g, 10 min), the supernatant was applied to Ni2+-affinity purification as described above. The fractions containing holo-DhbF were combined, and the buffer was exchanged to 50 mM Tris-HCl (pH 8.0) containing 150 mM NaCl. The sample was applied to size exclusion chromatography using a HiLoad 16/60 Superdex 200 (Cytiva) column with a mobile phase of 50 mM Tris-HCl (pH 8.0) containing 150 mM NaCl at a flow rate of 0.8 mL min-1. The fractions containing pure holo-DhbF were combined, the proteins were condensed, and the buffer was exchanged to Lysis buffer with a PD-10 column.

For DhbE and holo-DhbB, pET32a-dhbE and pETDuet-1-dhbB-pptase were individually introduced into E. coli BL21(DE3). Cultivation, gene expression induction, and removal of thermolabile proteins were carried out in the same manner as DhbF as described above. After exchanging the buffer to 50 mM Tris-HCl (pH 8.0) using a PD-10 column, the sample was applied to anion-exchange chromatography using HiTrap Q HP 5 mL (Cytiva). Proteins were eluted with a linear gradient of 0 to 1 M NaCl in 50 mM Tris-HCl (pH 8.0). After concentrating the sample and exchanging the buffer to 50 mM Tris-HCl (pH 8.0) containing 150 mM NaCl, the objective proteins were further purified by size exclusion chromatography in the same manner as DhbF.

For pyrophosphatase, pET32a-ppase was introduced into E. coli BL21(DE3). Cultivation and gene expression induction were carried out in the same manner as DhbF. Removal of thermolabile proteins were carried out at 70°C for 30 min, and the supernatant was used as the pyrophosphatase solution.

Circular dichroism spectroscopy of DhbF

Circular dichroism (CD) spectroscopy of DhbF was carried out in the temperatures ranging of 30°C to 60°C using a Jasco J-820AC spectropolarimeter. Far-UV spectra were measured from 200 nm to 300 nm in 50 mM Tris-HCl (pH 8.0). The protein concentration was 1 μM. The pathlength was 1 mm.

4'-Phosphopantetheinylation of DhbB by PPTase

The reaction mixture contained apo-DhbB (16.5 μM), PPTase (2 μM), CoA (100 μM) (Oriental Yeast, Japan), MgCl2 (1 mM), and dithiothreitol (DTT) (5 mM) in Lysis buffer. The reaction was carried out at 60°C for 1 h. Proteins were concentrated, and the buffer was exchanged to water using Amicon Ultra-0.5 30K (Merck). The reaction solution (1 μL) was mixed with 4 μL of a saturated solution of sinapic acid in 50%(v/v) acetonitrile containing 1%(v/v) trifluoroacetic acid, and the resulted solution (1 μL) was applied to MALDI-TOF-MS using autoflex maX (Bruker, USA).

Enzymatic characterization of DhbE and DhbF

Malachite Green Phosphate Assay Kit (BioAssay Systems, USA) was brought to room temperature before use. The Malachite Green solution was prepared by following the manufacturer's instructions. For examining the reaction temperature, the DhbE reaction was carried out with DhbE (4 μM), E. coli cell extract containing T. thermophilus pyrophosphatase (0.48 mg mL-1), MgCl2 (10 mM), 2,3-DHB (1.5 mM) (Tokyo Chemical Industry, Japan), ATP (10 mM), and DTT (5 mM) in Lysis buffer at 30, 40, 50, 60, or 70°C for 5 min. The DhbF reaction was also carried out with holo-DhbF (1 μM), pyrophosphatase, hydroxylamine (200 mM) (Tokyo Chemical Industry), MgCl2 (10 mM), Gly or L-Thr (1.5 mM), ATP (10 mM) (Oriental Yeast), DTT (5 mM) in Lysis buffer at 30, 40, 50, 60, or 70°C for 5 min. For examining substrate specificity, the DhbE reaction was carried out using various compounds (2,3-DHB, 3,4-DHB, or 2-/3-/4HB) (FUJIFILM Wako Chemical or Tokyo Chemical Industry) at 60°C for 1 h. The DhbF reaction was carried out using holo-DhbF, DhbF1, or DhbF2 in the presence of various compounds (Gly, Ala, L-Thr, D-Thr, or L-Ser) at 60°C (holo-DhbF) or 30°C (DhbF1 and DhbF2) for 1 h. All reactions were initiated by the addition of ATP and the volume was 100 μL. The reactions were quenched by adding 10 μL of 6 N HCl. After centrifugation (4°C, 20,000 ×g, 10 min), a 20 μL of the aliquot was mixed with 780 μL of 50 mM Tris-HCl (pH 8.0). The solution (80 μL) was subsequently mixed with the Malachite Green solution (20 μL) and incubated at room temperature for 30 min. The absorbance at 620 nm was measured to quantify phosphate. Examinations of substrate specificity were performed in triplicate.

In vitro reconstitution of BB biosynthesis and BB analogue synthesis

For in vitro reconstitution of BB biosynthesis, the reaction mixture (100 μL) contained DhbE, holo-DhbB, holo-DhbF, MbtH (each 1 μM), ATP (10 mM), MgCl2 (10 mM), DTT (5 mM), 2,3-DHB (1.5 mM), Gly (1.5 mM), and L-Thr (1.5 mM) in 50 mM Tris-HCl (pH 8.0). The reactions were carried out between 30°C and 60°C for 16 h. The reaction mixture was added 1 N HCl (100 μL) and subsequently ethyl acetate (200 μL). After vortexed, the organic layer was separated, then ethyl acetate was removed by heating at 60°C. The residue was dissolved in 100 μL of methanol and analyzed by LC-ESI-MS. For BB analogue synthesis, the reaction mixture contained holo-DhbF2 (1 μM), ATP (10 mM), MgCl2 (10 mM), DTT (5 mM), synthetic analogues described below (1.5 mM), and L-Thr (1.5

mM) in 50 mM Tris-HCl (pH 8.0). The other conditions are identical as described above.

Chemical synthesis of acyl-N-acetylcysteamine thioesters

Free carboxy acid (hippuric acid, N-(o-toluoyl)glycine, N-(m-toluoyl)glycine, 4-fluorohippuric acid, 3-benzoylpropionic acid, or 3-(4-fluorobenzoyl)propionic acid) (Tokyo Chemical Industry or Sigma) (1 mM, 1 eq.), N-acetylcysteamine (SNAC) (Sigma) (1 eq), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide, hydrochloride (water soluble carbodiimide) (Dojindo Laboratories, Japan) (1 eq) and 4-dimethylaminopyridine (1 eq) were dissolved in dichloromethane (10 mL) and stirred overnight under nitrogen gas at room temperature. The reaction was quenched by adding an equal volume of 1 M HCl, then the organic layer was collected using a separatory funnel. The organic layer was washed with water and brine, then dried over anhydrous Na2SO4. After filtration, the solvent was removed in vacuo. The produced acyl-SNAC thioester was purified by a medium pressure liquid chromatography system Isolera Prime (Biotage, Sweden) using a silica column (Sfär C18 D-Duo 100 Å 30 μm 12 g) (Biotage) with a linear gradient of 0-2% methanol in dichloromethane.

Characterization of synthesized SNAC thioesters

The purified compounds (5-20 mg) dissolved in CDCl3 or methanol-d4 were characterized by 1H and 13C NMR (ECS400) (JEOL, Japan). Raw spectra are shown in Figure S5-S10.

Benzoyl (BA)-Gly-SNAC (1): Total yield 99%. 1H NMR (400 MHz, CDCl3): δ 7.84 (dt, 2H, J = 7.2 Hz), 7.55 (tt, 1H, J = 7.2 Hz), 7.47 (tt, 2H, J = 7.6 Hz), 6.96 (br, 1H), 5.97 (br, 1H), 4.40 (d, 2H, J = 5.6 Hz), 3.45 (q, 2H, J = 6.4 Hz), 3.09 (t, 2H, J = 6.0 Hz), 1.96 (s, 3H). 13C NMR (400 MHz, CDCl3): δ 23.21, 28.42, 39.24, 49.88, 127.34, 128.77, 132.17, 133.45, 168.17, 170.90, 197.93. LRMS m/z calcd. [M+H]+ 281.1, found 281.1.

2-Methyl-BA-Gly-SNAC (2): Total yield 46%. 1H NMR (400 MHz, CDCl3): δ 7.44 (d, 1H, J = 4.0 Hz), 7.33 (dt, 1H, J = 8.0 Hz), 7.21 (q, 2H, J = 4.0 Hz), 7.14 (t, 1H, J = 6.0 Hz), 6.48 (br, 1H, J = 6.4 Hz), 4.28 (d, 2H, J = 6.0 Hz), 3.35 (q, 2H, J = 6.4 Hz), 3.03 (t, 2H, J = 6.4 Hz), 2.44 (s, 3H), 1.89 (s, 3H). 13C NMR (400 MHz, CDCl3): δ 19.98, 23.17, 28.42, 39.22, 49.65, 125.89, 127.00, 130.51, 131.27, 135.15, 136.67, 170.54, 170.76, 197.69. LRMS m/z calcd. [M+H]+ 295.1, found 295.2.

3-Methyl-BA-Gly-SNAC (3): Total yield 99%. 1H NMR (400 MHz, CDCl3): δ 7.54-7.73 (m, 3H), 7.26-7.32 (m, 1H), 6.45-6.63 (br, 1H), 4.31 (m, 2H), 3.29-3.43 (m, 2H), 2.96-3.08 (m, 2H), 2.36 (s, 3H), 1.92 (s, 3H). 13C NMR (400 MHz, CDCl3): δ 21.43, 23.20, 28.41, 39.24, 49.87, 124.32, 128.03, 128.64, 132.90, 133.42, 138.63, 168.35, 170.89, 198.00. LRMS m/z calcd. [M+H]+ 295.1, found

295.1.

4-Fluoro-BA-Gly-SNAC (4): Total yield 82%. 1H NMR (400 MHz, methanol-d4): δ 7.93 (tt, 2H, J = 8.0 Hz), 7.23 (tt, 2H, J = 8.8 Hz), 5.49 (s, 1H), 4.25 (s, 2H), 3.29-3.32 (m, 2H), 3.04 (t, 2H, J = 6.8 Hz), 1.91 (s, 3H). 13C NMR (400 MHz, methanol-d4): δ 21.19, 27.51, 38.64, 49.23, 115.13, 115.35, 129.77, 129.86, 163.86, 166.39, 168.08, 172.15, 197.64. LRMS m/z calcd. [M+H]+ 299.1, found 299.1.

3-Benzoylpropionyl-SNAC (5): Total yield 96%. 1H NMR (400 MHz, CDCl3): δ 7.96 (d, 2H, J = 8.4 Hz), 7.58 (tt, 1H, J = 7.6 Hz), 7.47 (t, 2H, J = 8.0 Hz), 6.30 (br, 1H), 3.44 (q, 2H, J = 6.0 Hz), 3.38 (q, 2H, J = 6.0 Hz), 3.06 (t, 2H, J = 6.4 Hz), 3.03 (t, 2H, J = 6.4 Hz), 1.96 (s, 3H). 13C NMR (400 MHz, CDCl3): δ 23.22, 28.66, 33.75, 37.82, 39.62, 128.12, 128.80, 133.57, 136.27, 170.71, 197.83, 198.95. LRMS m/z calcd. [M+H]+ 280.1, found 280.1.

3-(4-Fluorobenzoyl)propionyl-SNAC (6): Total yield 65%. 1H NMR (400 MHz, CDCl3): δ 8.02 (tt, 2H, J = 7.6 Hz), 7.17 (t, 2H, J = 8.8 Hz), 6.26 (br, 1H), 3.46 (q, 2H, J = 6.0 Hz), 3.37 (t, 2H, J = 6.4 Hz), 3.09 (t, 2H, J = 6.8 Hz), 3.05 (t, 2H, J = 6.8 Hz),1.98 (s, 3H). 13C NMR (400 MHz, CDCl3): δ 23.22, 28.68, 33.59, 37.76, 39.57, 115.81, 116.03, 130.78, 132.75, 164.72, 167.27, 170.61, 196.20, 198.83. LRMS m/z calcd. [M+H]+ 298.1, found 298.1.

Supporting Information

Additional figures and tables are available in the Supporting Information.

Acknowledgements

This work was partly supported by GteX Program Japan Grant Number JPMJGX23B4. NMR analyses were carried out at the Analytical Instrumentation Facility at the Graduate School of Engineering in Osaka University. This work used research equipment shared in the MEXT Project for promoting public utilization of advanced research infrastructure (Program for supporting construction of core facilities) under Grant Number JPMXS0441200024.

Keywords: bacillibactin, biosynthesis, mutational synthesis, nonribosomal peptide synthetase, thermophile

Conflict of Interest

The authors declare no conflict of interest.

References

[1] C. Andreini, I. Bertini, G. Cavallaro, G. L. Holliday, J. M. Thornton, J. Biol. Inorg. Chem. 2008, 13, 1205-1218.

[2] H. Beinert, R. H. Holm, E. Münck, Science 1997, 277, 653-659.

[3] M. L. Guerinot, Annu. Rev. Microbiol. 1994, 48, 743-772.

[4] M. Miethke, M. A. Marahiel, Microbiol. Mol. Biol. Rev. 2007, 71, 413-451.

[5] A. M. Timofeeva, M. R. Galyamova, S. E. Sedykh, Plants (Basel) 2022, 11, 3065.

[6] J. J. May, T. M. Wendrich, M. A. Marahiel, J. Biol. Chem. 2001, 276, 7209-7217.

[7] R. D. Süssmuth, A. Mainz, Angew. Chem. Int. Ed. Engl. 2017, 56, 3770-3821.

[8] K. Hotta, C. Y. Kim, D. T. Fox, A. T. Koppisch, Microbiology 2010, 156, 1918-1925.

[9] T. Ito, Appl. Environ. Microbiol. 1993, 59, 2343-2345.

[10] K. S. Ibrahim, M. Aishwarya, R. P. B. Kannan, Secondary metabolites from extremophiles with therapeutic benefits, Recent Advances and Future Perspectives of Microbial Metabolites, Academic Press 2023, pp. 249-267.

[11] E. B. Drejer, S. Hakvåg, M. Irla, T. Brautaset, Microorganisms 2018, 6, 42.

[12] K. Egan, D. Field, R. P. Ross, P. D. Cotter, C. Hill, Front. Microbiol. 2018, 9, 2116.

[13] J. Zebrowska, M. Witkowska, A. Struck, P. E. Laszuk, E. Raczuk, M. Ponikowska, P. M. Skowron, A. Zylicz-Stachula, Antibiotics (Basel) 2022, 11, 242.

[14] N. Garg, W. Tang, Y. Goto, S. K. Nair, W. A. van der Donk, Proc. Natl. Acad. Sci. U S A 2012, 109, 5241-5246.

[15] A. Koniuchovaitė, A. Petkevičiūtė, E. Bernotaitė, A. Gricajeva, A. Gegeckas, L. Kalėdienė, A. Kaunietis, Front. Microbiol. 2023, 14, 1207367.

[16] M. Vaičikauskaitė, M. Ger, M. Valius, A. Maneikis, E. Lastauskienė, L. Kalėdienė, A. Kaunietiset, Int. J. Biol. Macromol. 2019, 141, 333-344.

[17] K. Miyazaki, K. Hosoya, Microbiol. Resour. Announc. 2022, 11, e0020422.

[18] K. Blin, S. Shaw, H. E. Augustijn, Z. L. Reitz, F. Biermann, M. Alanjary, A. Fetter, B. R. Terlou, W. W. Metcalf, E. J. N. Helfrich, G. P. van Wezel, M. H. Medema, T. Weber, Nucleic Acids Res. 2023, 51, W46-W50.

[19] E. J. Drake, J. Cao, J. Qu, M. B. Shah, R. M. Straubinger, A. M. Gulick, J. Biol. Chem. 2007, 282, 20425-20434.

[20] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, Nat. Prod. Rep. 2014, 31, 61-108.

[21] P. Redford, P. L. Roesch, R. A. Welch, Infect. Immun. 2003, 71, 3088-3096.

[22] M. Miethke, O. Klotz, U. Linne, J. J. May, C. L. Beckering, and M. A. Marahiel, Mol. Microbiol. 2006, 61, 1413-1427.

[23] E. J. Dimise, P. F. Widboom, S. D. Bruner, Proc. Natl. Acad. Sci. U S A 2008, 105, 15311-

15316.

[24] S. Weist, R. D. Süssmuth, Appl. Microbiol. Biotechnol. 2005, 68, 141-150.

[25] J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, Proc. Natl. Acad. Sci. U S A 2006, 103, 5869-5874.

[26] W. Besenmatter, P. Kast, D. Hilvert, Proteins 2007, 66, 500-506.

Supporting Information

Table S1. Primers used in this study.

	Sequence (5'-3')
pET32a-F1	AATTCGAGCTCCGTCGACAAGCTTG
pET32a-R1	TATGTATATCTCCTTCTTAAAGTTA
pET32a-F2	ATGTGTGTCTCCTTCTTAAAGTTAAACAAAATTAT
pET32a-R2	GAATTCGAGCTCCGTCGACAAGCTTGCGGC
dhbE-F	ATGTTAAGCGGATGCCCTACATG
dhbE-R	TTATTCTGTTGTTTTTCCAGCAG
dhbB-F	AAGGAGATATACATATGGCTATCCCTGCTATTCCAGC
dhbB-R	GACGGAGCTCGAATTTTGCGAGTCTTGAGCAGATAGTA
pptase-F	AAGGAGATATACATATGATGGAATGGCAGGTTGCCAA
pptase-R	GACGGAGCTCGAATTAATATAGTAATCTTTCCAACCAA
mbtH-F	GAAGGAGATATACATATGACCAATCCTTTTG
mbtH-R	GACGGAGCTCGAATTTCTTTTTATCTTTTTCAGAC
ppase-F	GAAGGAGATATACATATGGCGAACCTGAAGAGCCTTCC
ppase-R	ACGGAGCTCGAATTCCTAGCCCTTGTAGCGGGCGATG
Duet-MCS2-dhbB-F	GACGGAGCTCGAATTTTATTCTGTTGTTTTTCC
Duet-MCS2-dhbB-R	TTGAGATCTGCCATACGGGCTTTGTTAGCAGCCGGATC
Duet-MCS2-dhbF-F	GAAGGAGATATACATATGTCCAACCGCCAAGATATAC
Duet-MCS2-dhbF-R	TTGAGATCTGCCATACGGGCTTTGTTAGCAGCCGGATC
Duet-MCS1-pptase-F	AGAAGGAGATATACCATGATGGAATGGCAGGTTGCCAA
Duet-MCS1-pptase-R	TTAAATATAGTAATCTTTCCAAC
Duet-MCS2-dhbF1-F	GAAGGAGATATACATATGTCCAACCGCCAAGATATAC
Duet-MCS2-dhbF1-R	CCAATTGAGATCTGCCATATTAGTGGTGATGATGGTGATGA GGCGGCTTGATATTTTCTGCATC
Duet-MCS2-dhbF2-F	GAAGGAGATATACATATGGTCAAAACATATGTCC
Duet-MCS2-dhbF2-R	TTGAGATCTGCCATATTAGTGGTGATGATGGTGATGTTTTGT ACCTCTCCCTCCATTTATTGA

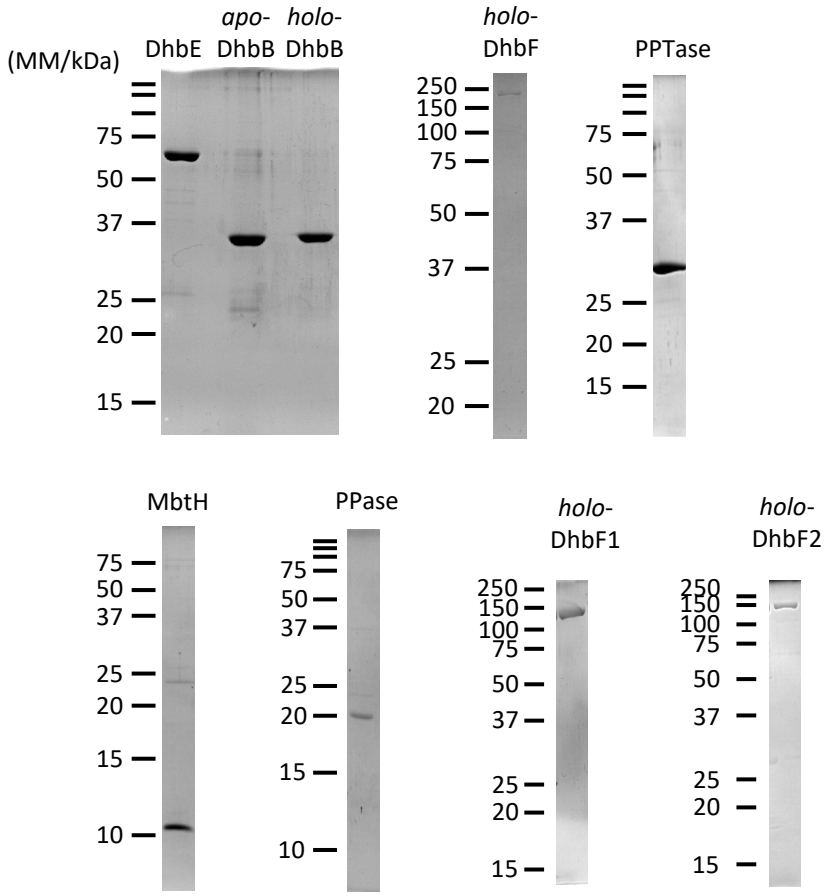


Fig. S1. SDS-PAGE analysis of the purified recombinant proteins. *apo*-Proteins were solely overexpressed in *E. coli* and are not 4'-phosphopantetheinylated, whereas *holo*-proteins were overexpressed with PPTase to be 4'-phosphopantetheinylated. DhbF1 and DhbF2 are truncated proteins which have the first and second module of DhbF, respectively.

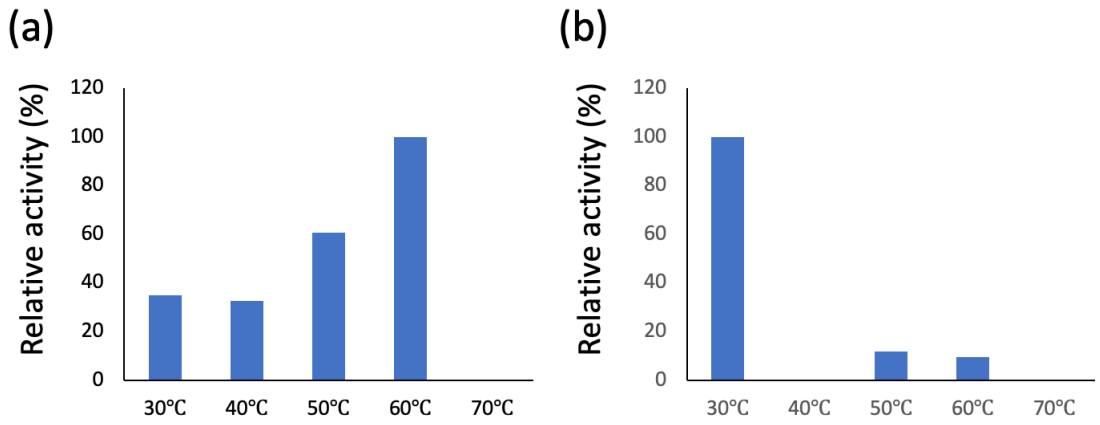


Figure S2. Temperature dependence of A domains of DhbE (a) and DhbF (b). Maximum activity of DhbE and DhbF were observed at 60°C and 30°C, respectively, and the specific activity was set at 100%.

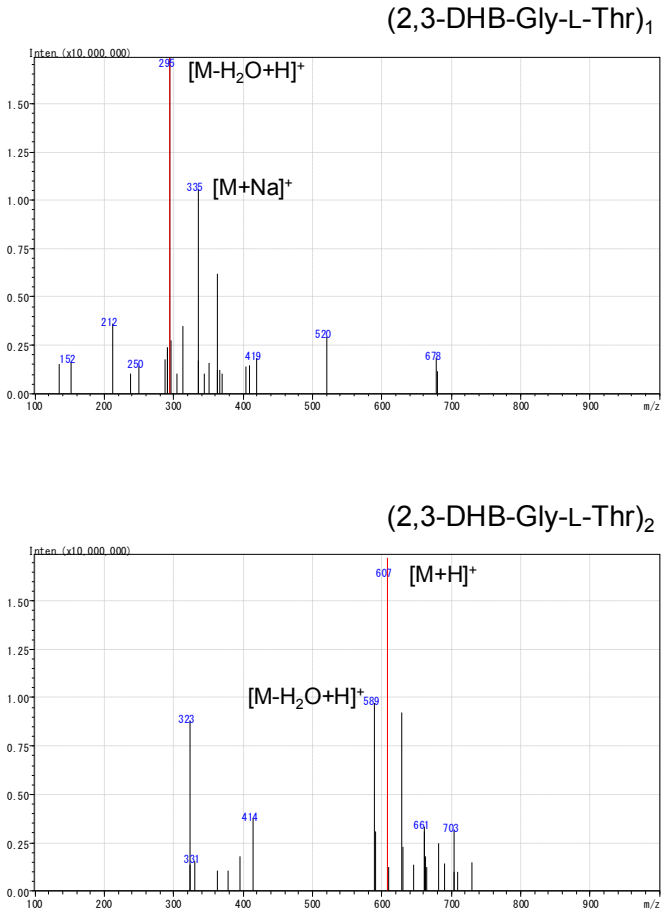


Figure S3. Mass spectra of (2,3-DHB-Gly-L-Thr)$_1$ (upper) and (2,3-DHB-Gly-L-Thr)$_2$ (lower). The corresponding chromatograms are shown in Figure 6b.

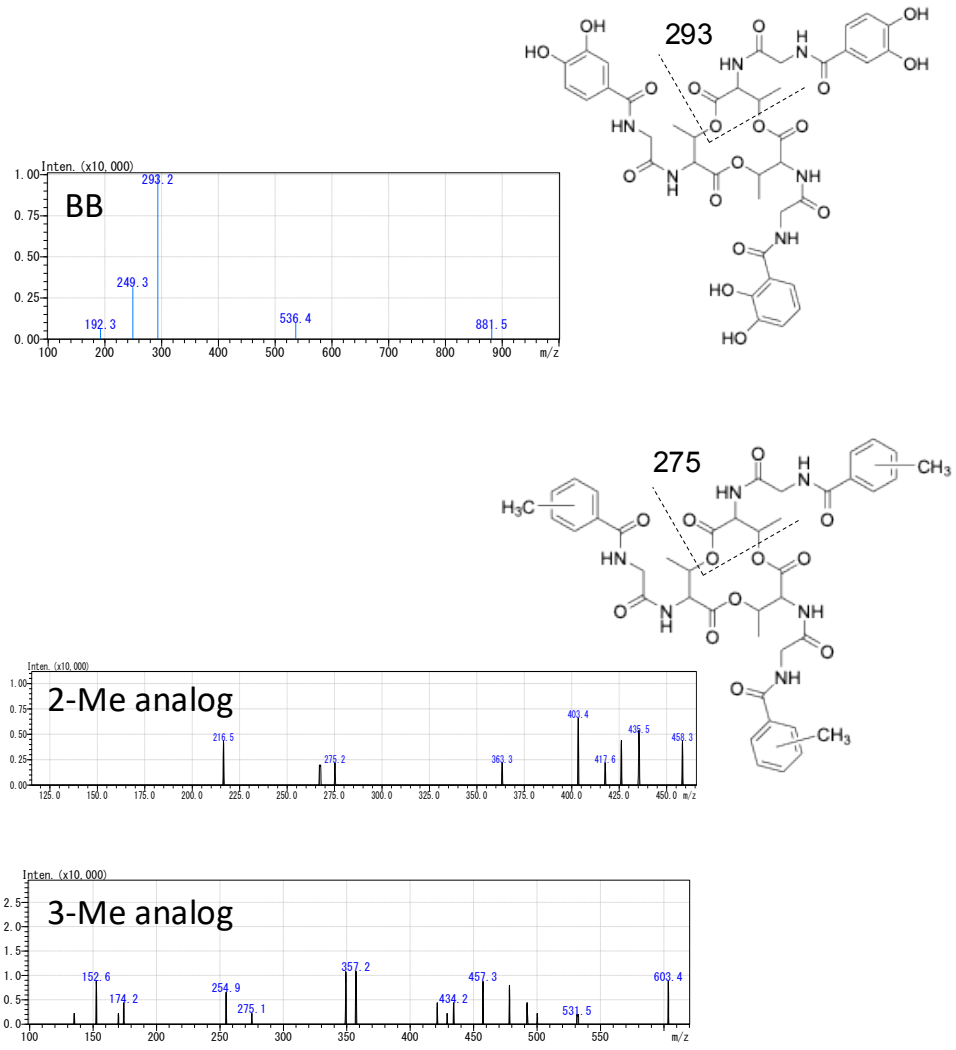
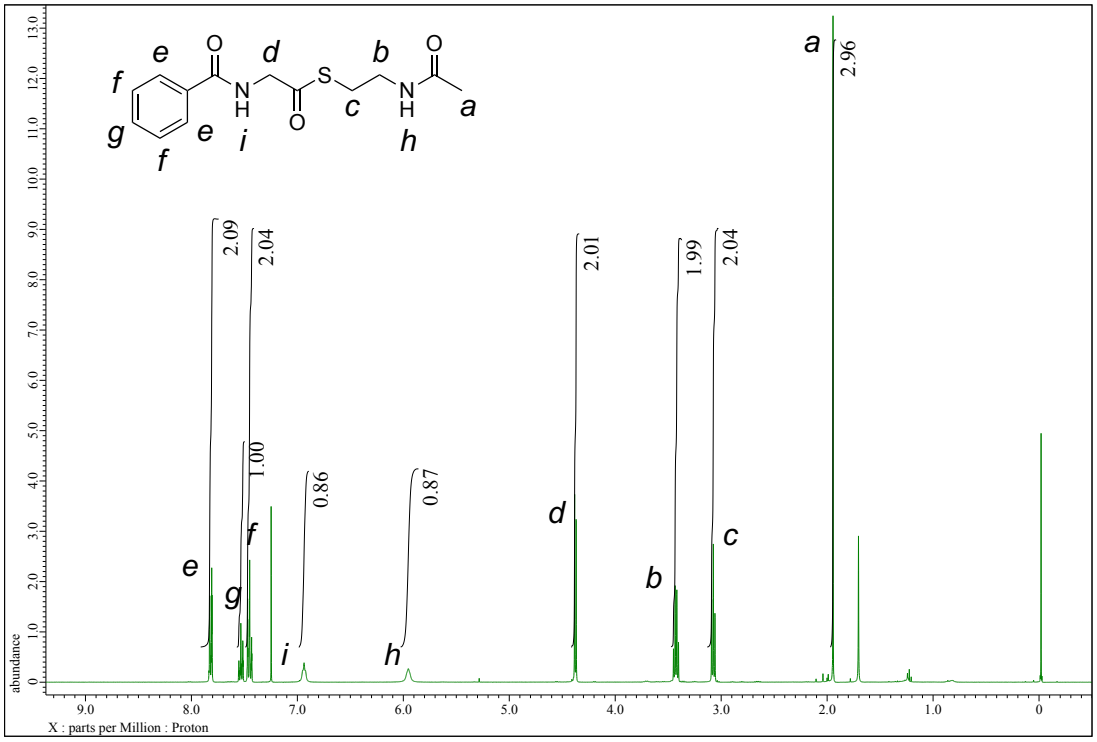


Figure S4. MS/MS spectra of BB (upper) and analogues synthetized using **2** and **3**. Key fragmentations are indicated in the chemical structures. The fragment of BB was reported by Li *et al.*[1]

(a)



(b)

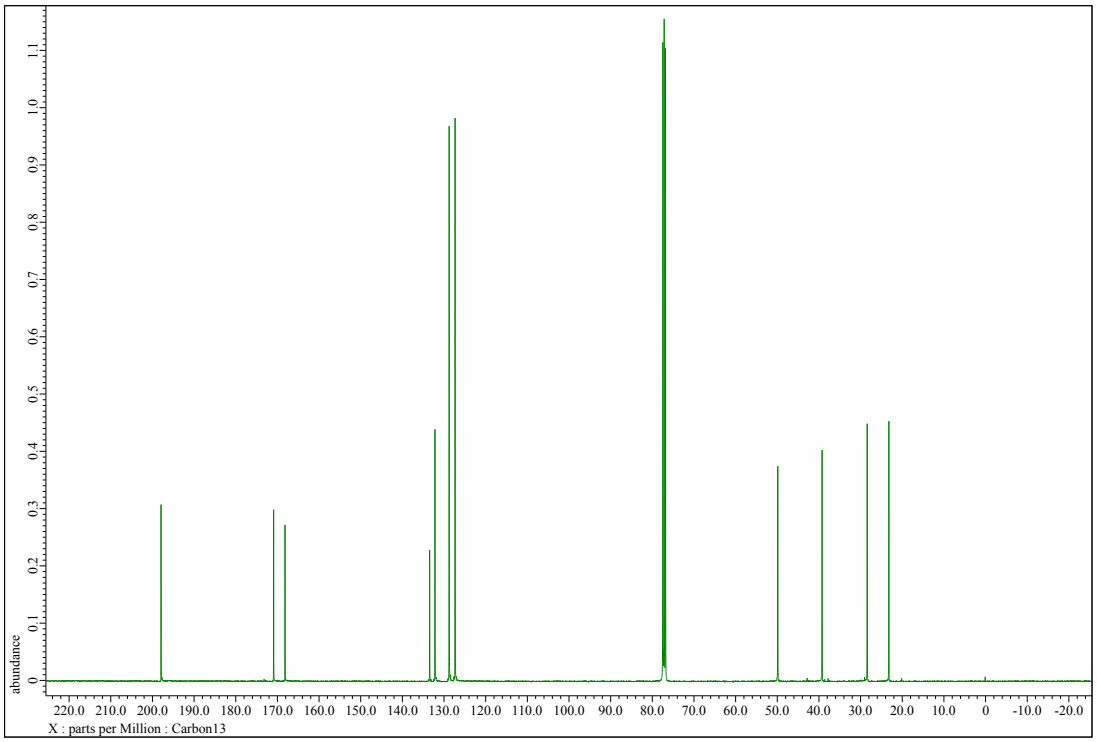
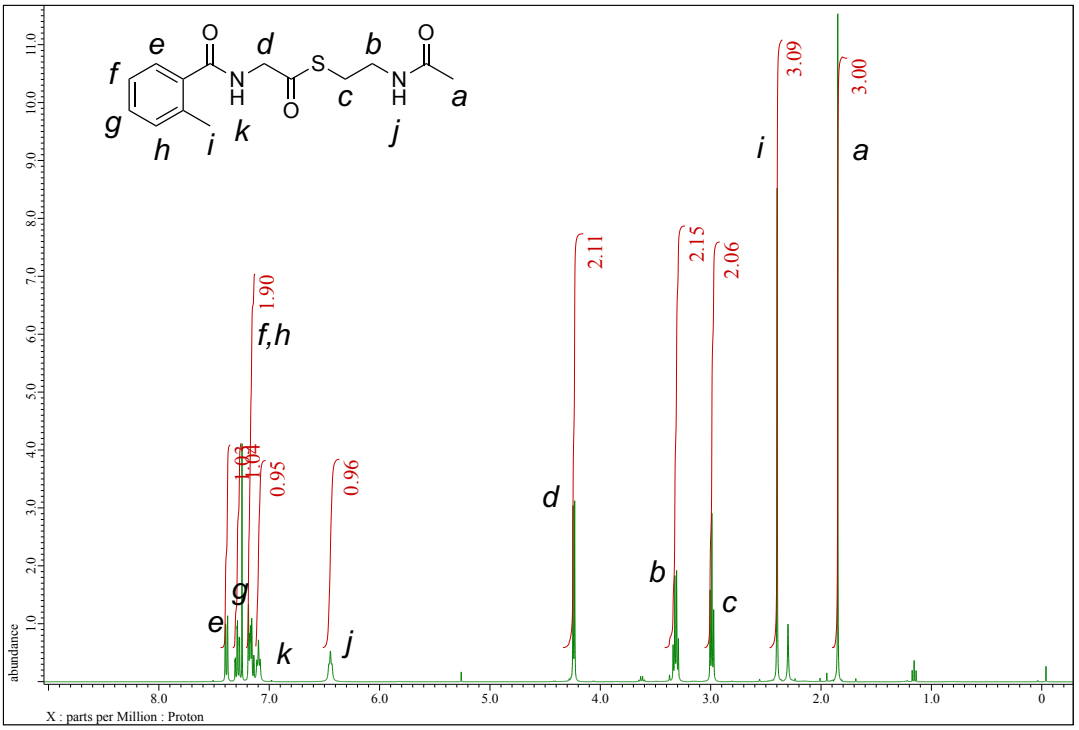


Figure S5. ^{1}H NMR (a) and ^{13}C NMR (b) spectra of benzoyl (BA)-Gly-SNAC (**1**).

(a)



(b)

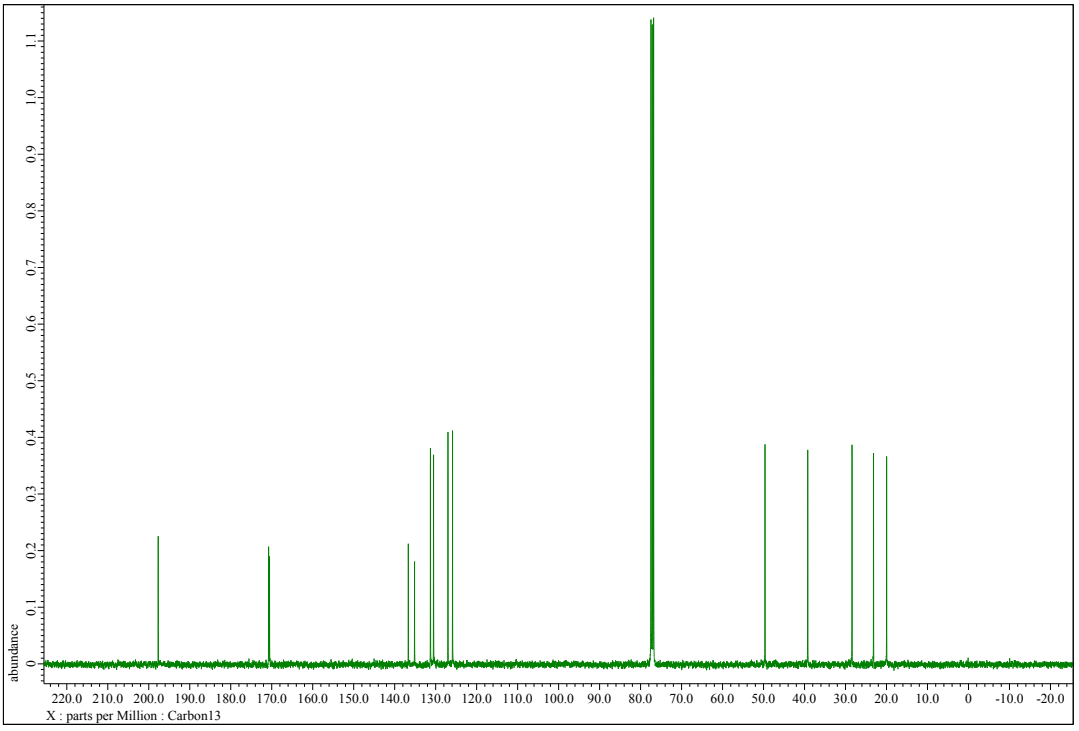


Figure S6. ^{1}H NMR (a) and ^{13}C NMR (b) spectra of 2-methyl-BA-Gly-SNAC (**2**).

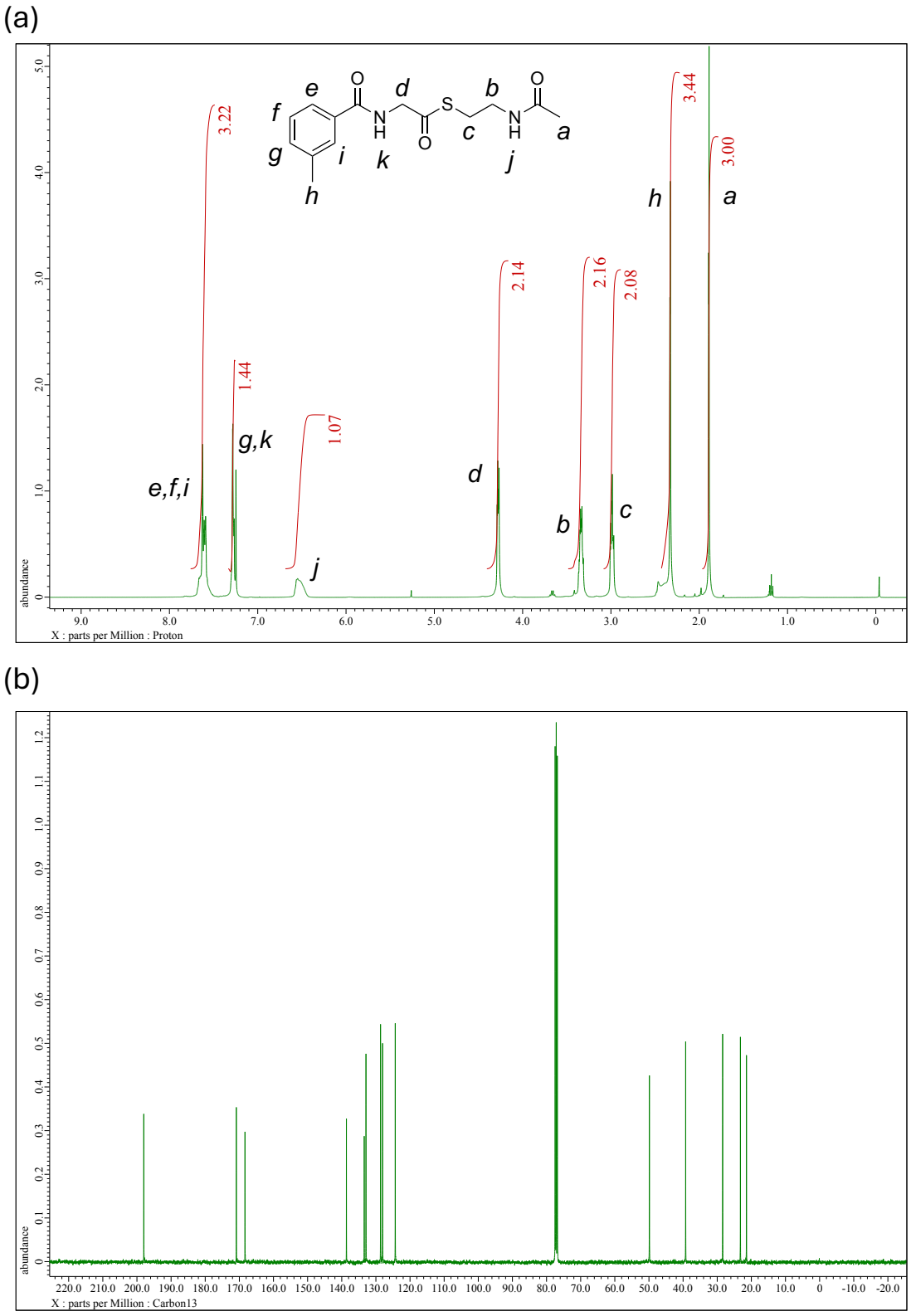
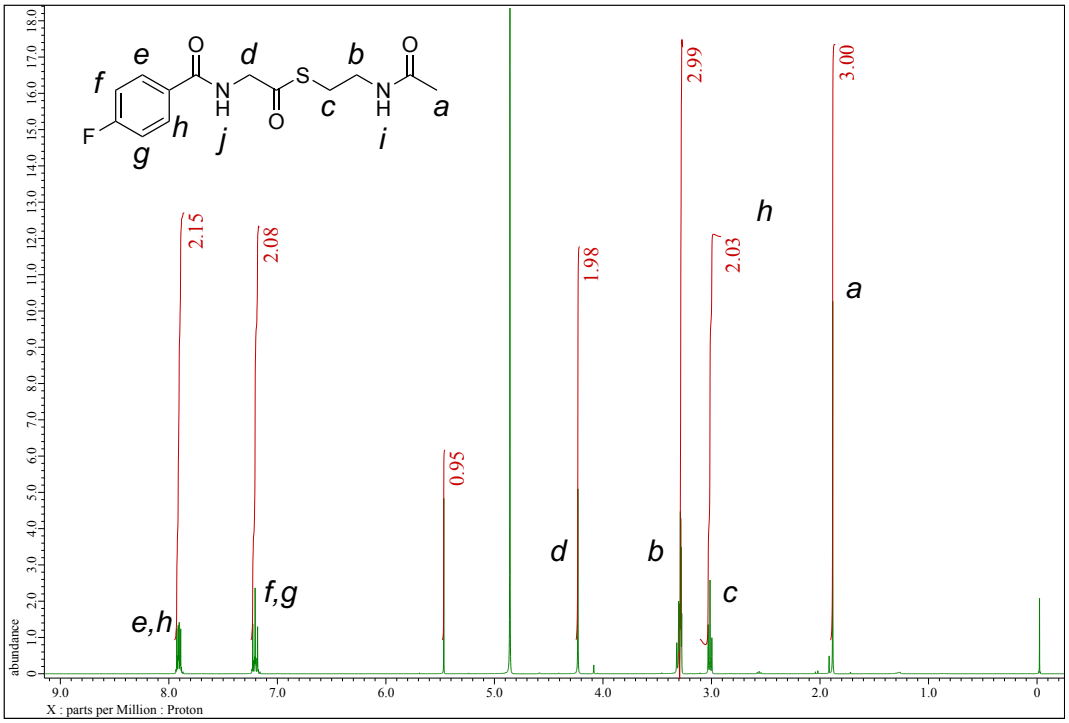


Figure S7. ^{1}H NMR (a) and ^{13}C NMR (b) spectra of 3-methyl-BA-Gly-SNAC (**3**).

(a)



(b)

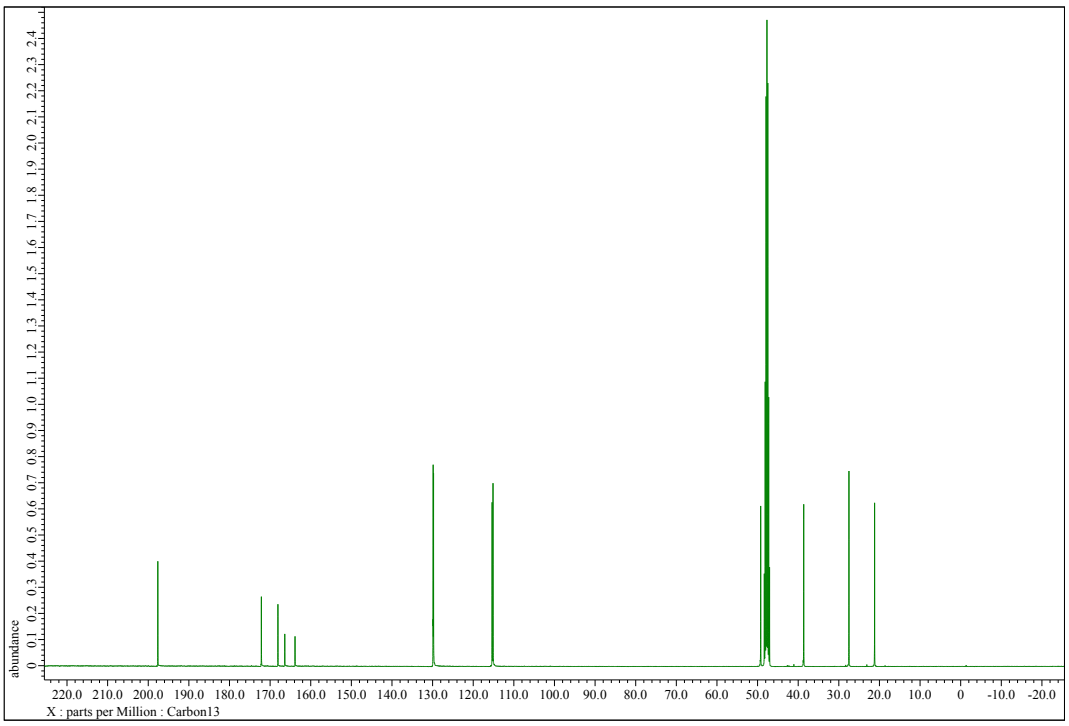
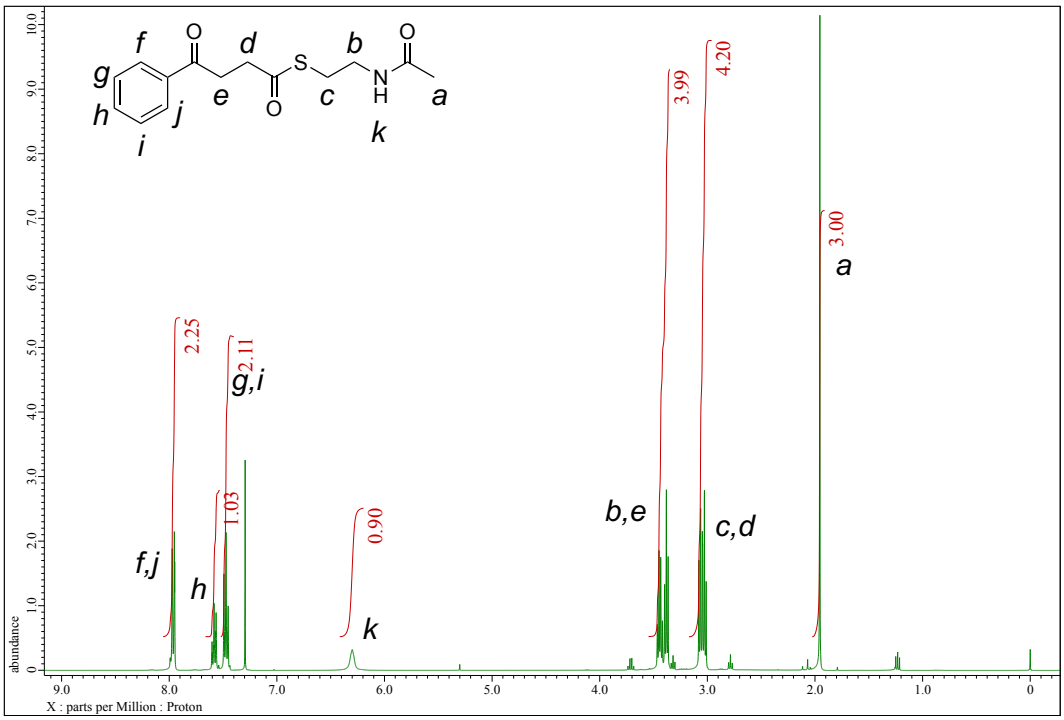


Figure S8. ^{1}H NMR (a) and ^{13}C NMR (b) spectra of 4-fluoro-BA-Gly-SNAC (**4**).

(a)



(b)

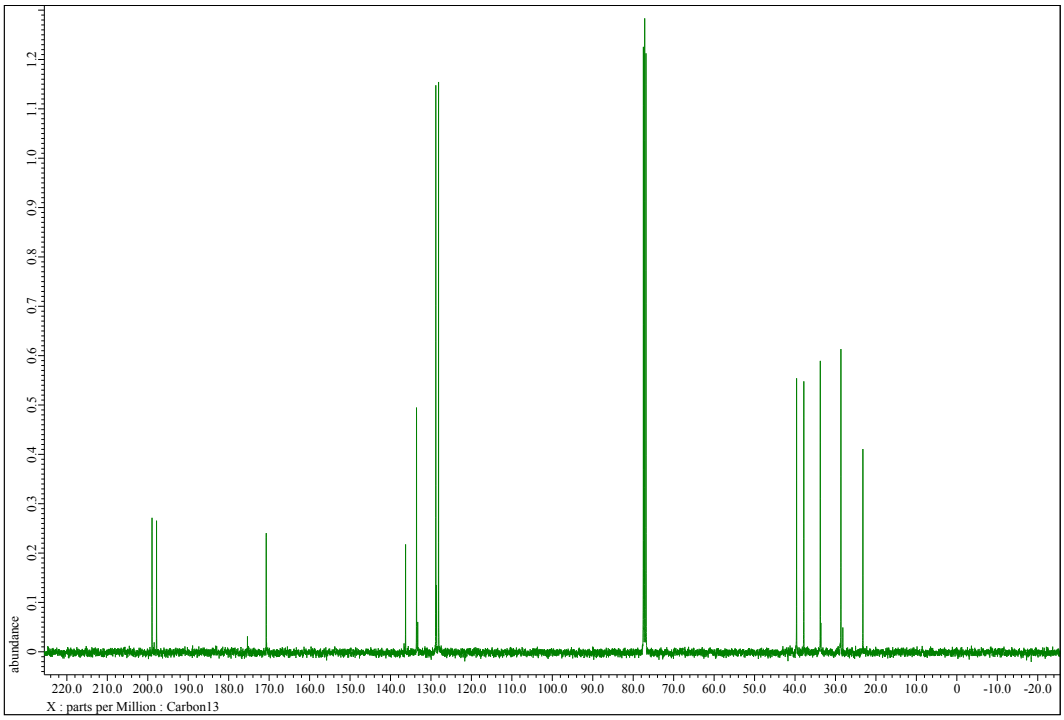
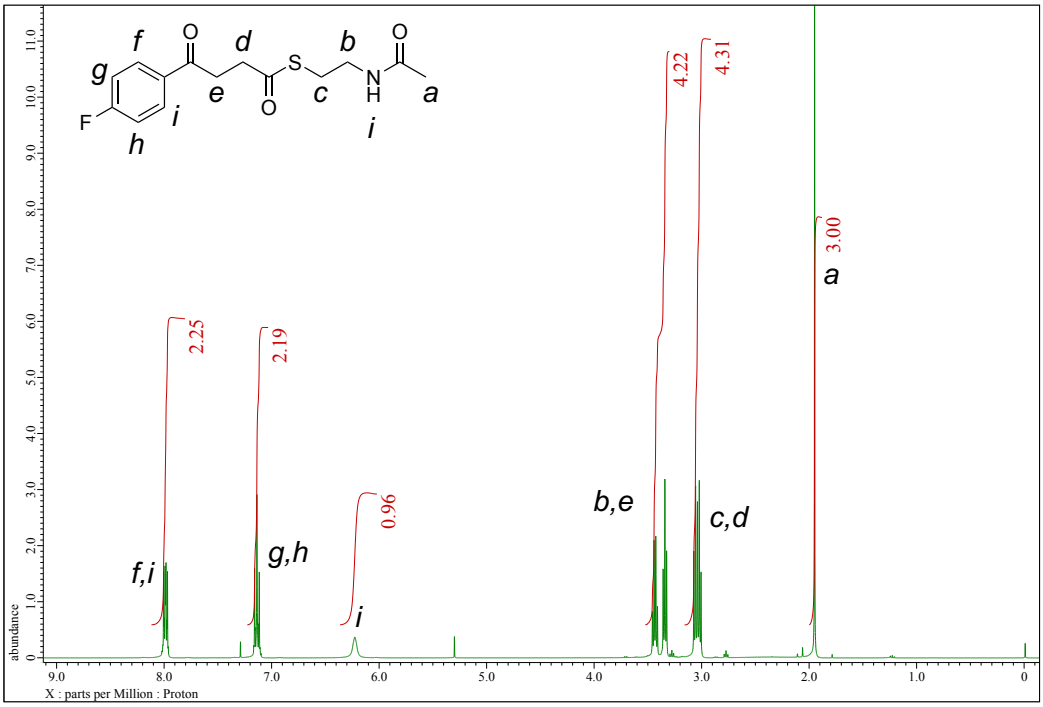


Figure S9. ^{1}H NMR (a) and ^{13}C NMR (b) spectra of 3-benzoylpropionyl- SNAC (**5**).

(a)



(b)

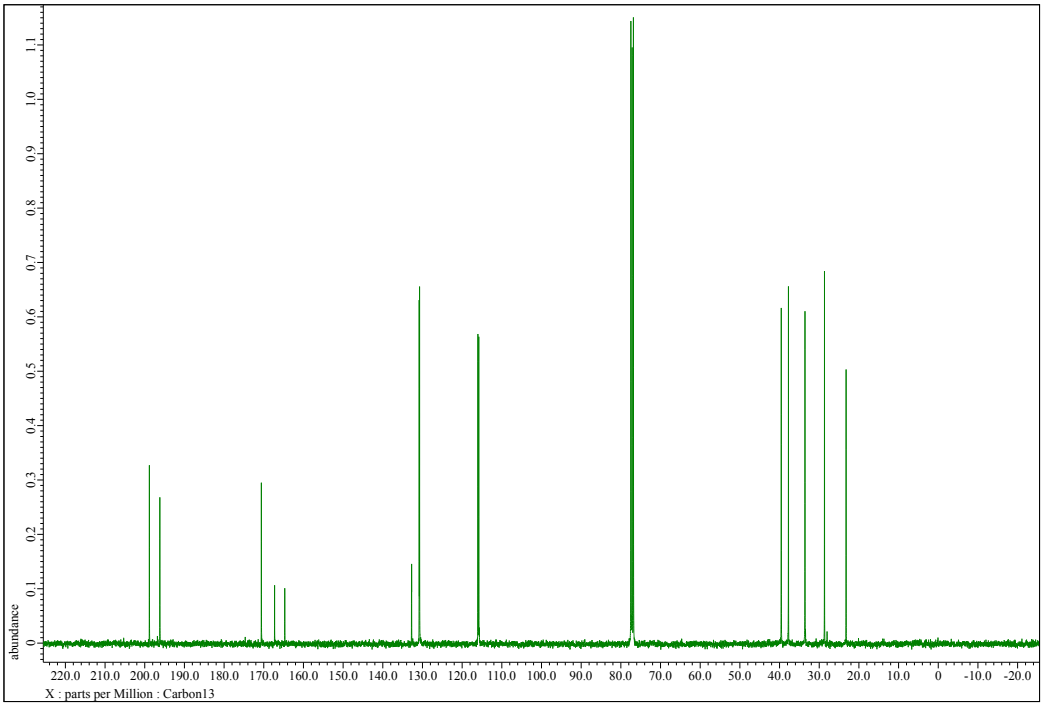


Figure S10. ^{1}H NMR (a) and ^{13}C NMR (b) spectra of 3-(4-fluorobenzoyl)propionyl-SNAC (**6**).

References

[1] J. Li, S. Liu, Z. Jiang, and C. Sun, *Tetrahedron* **2017**, *35*, 5245-5252.