C-S bond cleavage by a polyketide synthase domain

Ming Ma,[†,▽] Jeremy R. Lohman,[†,▽] Tao Liu,[‡,▽] and Ben Shen*[†,‡,§,⊥]

[†]Department of Chemistry, The Scripps Research Institute, Jupiter, FL 33458, USA; [‡]Division of Pharmaceutical Sciences, University of Wisconsin-Madison, Madison, WI 53705, USA; and [§]Department of Molecular Therapeutics and [⊥]Natural Products Library Initiative at The Scripps Research Institute, The Scripps Research Institute, Jupiter, FL 33458, USA

[▽]These authors contributed equally

*To whom correspondence should be addressed: Scripps Florida, 130 Scripps Way, #3A1, Jupiter, FL 33458; Email: shenb@scripps.edu; Tel: (561) 228-2456; Fax: (561) 228-2472

Supporting Online Materials

General procedures. *Streptomyces atroolivaceus* S-140 was provided by Kyowa Hakko Kogyo Co. Ltd, Tokyo, Japan (1), *Micromonospora aurantiaca* ATCC 27029 was purchased from the American type culture collection (ATCC), and *Catenulispora acidiphilia* DSM 44928 was purchased from German collection of microorganisms and cell cultures (DSMZ). *E. coli* DH5α or NovaBlue were used as hosts for general sub-cloning, and *E. coli* BL21(DE3) was used as the host for protein overproduction (Novagen, Madison, WI). Standard procedures were used to manipulate purified DNA (2). Genomic DNAs were purified from *S. atroolivaceus* S-140, *M. aurantiaca* ATCC 27029, and *C. acidiphilia* DSM 44928 using the salting out procedure (3). The DNA oligonucleotide primers were synthesized by the University of Wisconsin-Madison Biotechnology Center (Madison, WI). PCR was performed with a PerkinElmer GeneAmp 2400 (PerkinElmer Life Sciences, Inc.) using the high-fidelity Expand PCR system according to the manufacturer's protocol (Roche). The pRSF-TEV/LIC vector was previously described (4). Optical rotations were measured with an AUTOPOL IV automatic polarimeter (Rudolph Research Analytical). UV spectra were collected with a NanoDrop 2000C spectrophotometer (Thermo Scientific). NMR data were collected on a Bruker 700 MHz or 400 MHz Ultra Shield Magnet System. High-resolution electro-spray ionization mass spectrometry (HR-ESI-MS) data were collected on a 6230 TOF LC/MS spectrometer (Agilent Technologies).

Expression of *lnmJ-SH*, *caJ-SH*, and *maJ-SH* fragments in *E. coli* and purification of the resultant recombinant proteins. The *lnmJ-SH* (residues 6624-7070), *caJ-SH* (residues 6444-6889), and *maJ-SH* (residues 6016-6459) fragments were amplified with primers 5'-AAA ACC TCT ATT TCC AGT CGC GGC CGA AGC CCG CCG CCG C-3'/5'TAC TTA CTT AAA TGT TAG TGG CCC TCG GCC ATC GCC TC-3', 5'-AAA ACC TCT ATT TCC AGT CGC AGA GCA CAG TGA TAG CG-3'/5'-TAC TTA CTT AAA TGT TAC AGC ACG GCT TCG TTG TCG TCG G and 5'-AAA ACC TCT ATT TCC AGT CGG CCG CCG TGC CGT CGC ACA ACG C-3'/5'-TAC TTA CTT AAA TGT TAC AGC ACG CTG CCG TTG TCG TCG GCC, respectively (the sequences flanking the native domain sequence were underlined). The resultant PCR products *lnmJ-SH* (447 residues), *caJ-SH* (446 residues), and *maJ-SH* (444 residues) PCR products (Fig. S3) were introduced into a pRSF-TEV/LIC-based plasmid pBS3080 (4) using ligation independent cloning methods. Briefly, plasmid pBS3080 (4) was digested with BsmFI, purified by gel electrophoresis, treated with T4 DNA polymerase in the presence of dGTP at 20 °C for 30 min, and finally heated at 75 °C for 20 min to denature the polymerase. The PCR products were similarly treated with T4 DNA polymerase in the presence of dCTP at 20 °C for 30 min and heated at 75 °C for 20 min. The T4 DNA polymerase treated plasmid and PCR products have complementary sequences at both 5'- and 3'-ends; they were mixed at room temperature, annealed on ice for 5 min, and transformed into *E. coli* DH5α for ligation independent cloning. The plasmid pBS3109 (for *lnmJ-SH*), pBS3110 (for *caJ-SH*), or pBS3111 (for *maJ-SH*) was isolated from *E. coli* DH5α based on selection for kanamycin resistance and confirmed by DNA sequencing. The recombinant LnmJ-SH, CaJ-SH and MaJ-SH protein contains 470, 469 and 467 amino acids including the N-terminal 23 amino acids (MGSSHHHHHHSQDPGDENLYFQS) from the plasmid, respectively.

For expression, the three plasmids were introduced into *E. coli* BL21 (DE3), and the resultant strains were cultured in a lysogeny broth (LB) medium supplemented with appropriate antibiotics. All cells were grown at 18 °C and induced with 0.1 mM IPTG when OD600 reached ~0.5. They were further cultured at 18 °C for an additional 18 h. Cells were harvested by centrifugation (4 °C, 8000 rpm for 15 min) and resuspended in Buffer A (100 mM sodium phosphate, pH 7.0, containing 300 mM NaCl, 10mM imidazole) supplemented with a complete protease inhibitor tablet, EDTA-free (Roche Applied Science, Indianapolis, IN). The cells were lysed by sonication (4 × 30 s pulsed cycle), and the debris was removed by centrifugation (4 °C, 15000 rpm for 50 min). The clarified supernatant was loaded onto a pre-equilibrated Ni-NTA agarose column (Qiagen, Valencia, CA). The column was washed with 5-column volumes of Buffer B (50 mM Tris, pH 8.0, containing 300 mM NaCl, 20 mM imidazole). The His_6-tagged proteins were eluted with 6-column volumes of Buffer B containing 250 mM imidazole. The Protein purity was assessed by 12% SDS-PAGE. The determination of protein concentrations was by the Bradford assay (Bio-Rad, Hercules, CA) (The PLP content was determined by measuring the

spectrum of the enzyme in 0.1 M NaOH, assuming ε_{390} = 6,600 $M^{-1}cm^{-1}$). The proteins were dialyzed in 25 mM Tris, pH 8.0, 25 mM NaCl overnight and stored at -80 °C for future use.

Size exclusion chromatography of LnmJ-SH. A sample (300 µL) of 3 AU LnmJ-SH was loaded onto a Superose 6 10/300 GL column and eluted with 10 µM PLP, 100 mM KCl, 50 mM Tris pH 8.0. A calibration was performed with standards of blue dextran (2,000 KDa), thyroglobulin (660 KDa), ferritin (450 KDa), bovine serum albumin (66.0 KDa), and ribonuclease A (14.0 KDa) (4). LnmJ-SH was eluted as a homotetramer with an apparent molecular weight of 186 KDa.

Solubilization of substrates. Substrates L-cysteine (**1**), compound **9**, and compound **10** can be dissolved in water at high concentration (> 200 mM). As L-cystine (**2**) is notoriously insoluble, we used 2 equivalents of KOH with gentle heating and sonication to obtain a 0.1 M solution (5). At pH 8.5 the maximum solubility for **2** at 30 °C was around 25 mM. The other thioethers (**3** – **8**) were also solubilized using 1 equivalent of KOH but their maximum solubility cannot reach high enough to saturate the activity of enzyme; therefore only the k_{cat}/K_m values were obtained for these substrates.

High performance liquid chromatography (HPLC). HPLC analyses were carried out on a Varian HPLC system equipped with Prostar 210 pumps and Prostar 330 photodiode array detector, using an Alltech HP Altima column (C18, 4.6 × 250 mm, 5 µm). A two-phase mobile solvent system, solvent A (100% water with 0.1% acetic acid) and solvent B (90% acetonitrile in water with 0.1% acetic acid), was used. For analysis of pyruvate-derived methylquinoxalinol, the following program was used: 0-20 min, 100% solvent A to 10% solvent A and 90% solvent B; 20-22 min, 10% solvent A and 90% solvent B to 100% solvent A; 22-26 min, 100% solvent A, with a flowrate of 1mL/min and UV detection at 340 nm. The same conditions for detection of methylquinoxalinol were used for detection of thiocysteine-monobromobimane adduct from the reaction of LnmJ-SH with cystine (**2**) except for the use of UV detection at 380 nm. The same conditions for detection of methylquinoxalinol were used for detection of compound **16** and **17** in the reactions of LnmJ-SH with compound **9** and **10**, except for the use of UV detection at 235 nm. For analysis of thiophenol, the following program was used: 0-15 min, 100% solvent A to 100% solvent B; 15-21 min, 100% solvent B; 21-24 min, 100% solvent B to 100% solvent A; 24-30 min, 100% solvent A, with a flowrate of 1 mL/min and UV detection at 260 nm. The HPLC purification of compound **9**-**15** was carried out on a Agilent 1260 Infinity system using an Eclipse XDB-C18 column (21.2 x 250 mm, 7 µm), and the different programs were used (see information below).

Detection of pyruvate. The detection of pyruvate produced from the reaction of LnmJ-SH with substrates **1**-**10** is based on the formation of methylquinoxalinol, resulting from the reaction of pyruvate with o-phenylenediamine (OPD) (6, 7) (Fig. S4*A*). To 300 µL of 12 mM OPD in 3 N HCl was added 100 µL of reaction mixtures (50 mM Tris-HCl, 30 µM LnmJ-SH, 4 mM substrate, 0.2 mM PLP, 20 mM KCl, pH 8.5, 30 ºC for 1 h), and resulting mixtures vials heated to 100 ºC for 30 min. Proteins were precipitated by centrifugation and HPLC analyses conducted on 100 µL aliquots (Fig. S4*B*). The UV absorption was measured at 340 nm. Standard curves were constructed by mixing 100 µL of pyruvate acid solution instead of the reaction solution with 300 µL of 12 mM OPD in 3 N HCl with pyruvate concentrations ranging from 20 µM - 0.4 mM.

Detection of hydrogen sulfide. Production of hydrogen sulfide produced from the reaction of LnmJ-SH with L-cysteine (**1**) was verified and quantified by production of methylene blue (8) (Fig. S5). To minimize volatilization of hydrogen sulfide, 100 µL of reaction mixture (50 mM Tris-HCl, 30 µM LnmSH, 4 mM L-cysteine (**1**), 0.2 mM PLP, 20 mM KCl, pH 8.5) were put in sealed vials and incubated at 30 °C for 1 h. To each vial were injected 900 µL of water using a syringe to dilute the sample, and 100 µL of color development reagent, containing 1 mg/mL *N*,*N*-dimethyl-p-phenylenediamine sulfate and 3 mg/mL ferric chloride in 6 N HCl, were added via syringe. After 20 min the absorbance was determined at 670 nm. Standard curves were determined by replacing the reaction mixture with sodium sulfide solutions with sulfide concentrations ranging from 5 - 40 µM. The hydrogen sulfide concentration was determined as 0.39 mM in this assay.

Detection of thiocysteine. The thiocysteine produced by the reaction of LnmJ-SH with L-cystine (**2**) was derivatized with monobromobimane using standard methods (9, 10) (Fig. S6*A*). The reaction (100 µL of 50 mM Tris-HCl, 30 µM LnmJ-SH, 4 mM L-cystine (**2**), 0.2 mM PLP, 20 mM KCl, pH 8.5, 30 ºC for 1 h) was terminated by addition of 100 µL of acetonitrile. Subsequently, 5 µL of 100 mM monobromobimane in acetonitrile were added to quench reaction mixture, and incubation conducted at room temperature for 30 min in the dark. Sample was then centrifuged at 14000 rpm for 5 min, and 50 µL of the mixture were subjected to HPLC (Fig. S6*B*) and mass spectrometry analysis, respectively. The identity of the produced thiocysteine-monobimane conjugate was substantiated by MS analysis (Varian 500-MS IT Mass spectrometer) with the $[M + H]^+$ at *m/z* 344.0.

Detection of thiophenol. The detection of thiophenol produced from the reaction of LnmJ-SH with *S*-phenyl-L-cysteine (**4**) was based on HPLC analysis, using the standard thiophenol as the positive control and the reaction with boiled LnmJ-SH as the negative control (Fig. S7). The 100 µL reaction system contained 50 mM Tris-HCl (pH 8.5), 4 mM *S*-phenyl-L-cysteine (**4**), 30 µM LnmJ-SH, 0.2 mM PLP, and 20 mM KCl, and was incubated at 30 ºC for 1 h. The reaction was terminated by addition of 10 µL of 3 N HCl, centrifuged and loaded on HPLC analysis.

Chemical synthesis of compounds 9 and 10. Compound **9** was synthesized using dimethylacryl acid and N-acetylcysteamine (SNAC) as the starting material. Dimethylacryl acid (500 mg), SNAC (121 mg), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC, 955 mg) and 4-(dimethylamino)pyridine (DMAP, 60 mg) were added into 10 mL dehydrated dichloromethane. The reaction mixture was stirred at room temperature for 3 h. Then the mixture was concentrated into 5 mL, added by 30 mL methanol and loaded onto HPLC for purification, using the program: 0-24 min, 15% to 80% acetonitrile in water, at a flowrate of 17 mL/min and 266 nm for detection. Compound **11** was obtained in 93% yield. Compound **11** (176 mg) and L-cysteine (1 g) were added into 5 mL water. The mixture was stirred at room temperature for 3 h and loaded onto HPLC preparation using the same conditions for **11** purification to give compound **9**, in 22% yield (Fig. S8*A*).

Compound **10** was synthesized using pantethine as a starting material. Firstly the hydroxyl groups in pantethine were protected by the reaction with 2-methoxypropene to form acetonides. Pantethine (300 mg), 2-methoxypropene (390 mg) and camphorsulfonic acid (CSA, 3 mg) were added into 10 mL dehydrated tetrahydrofuran (THF). The reaction mixture was stirred at room temperature for 18 h, then concentrated into 5 mL, added by 30 mL methanol and loaded onto HPLC. The HPLC was carried out with the program: 0-24 min, 10% to 80% acetonitrile in water, at a flowrate of 17 mL/min and 215 nm for detection, to give compound **12** in 90% yield. Compound **12** (250 mg) was mixed with dithiothreitol (DTT, 607 mg) in 5 mL 50% methanol in water and stirred at room temperature for 1 h, to give compound **13** in 90% yield using the same HPLC purification conditions as in **12** purification. Compound **13** (192 mg), dimethylacryl acid (305 mg), EDC (585 mg) and DMAP (37 mg) were added into 5 mL dehydrated dichloromethane. The reaction mixture was stirred at room temperature for 3 h. Then the mixture was concentrated into 2 mL, added by 10 mL methanol, and loaded onto HPLC for purification. The HPLC preparation was carried out using the same conditions as in **12** purification to give compound **14**, in 86% yield. Compound **14** (209 mg) was dissolved in 2 mL acetonitrile, then 2 mL water and 2 mL trifluoroacetic acid (TFA) were added, and the reaction mixture was stirred at room temperature for 20 min, to give the deprotected compound **15** in 80% yield purified by HPLC with the same conditions as in **12** purification. Compound **15** (122 mg) and L-cysteine (412 mg) were added into 5 mL water. The mixture was stirred at room temperature for 3 h and loaded onto HPLC preparation, using the program of 15% acetonitrile in water from 0 to 24 min at a flowrate of 17 mL/min and 235 nm for detection, to give compound **10** in 10% yield (Fig. S8*B*).

*Compound **9***: colorless oil; $[\alpha]^{26}_{D}$ – 18.1 (*c* 0.16, H_2O); UV (H_2O) λ_{max} (log ε) 237 (3.63) nm; ^{1}H NMR (400 MHz, methanol-d_4) (carbon labeling shown in Fig. S11): δ 1.95 (s, H_3-1), 3.36 (t, J = 6.6 Hz, H_2-3), 3.05 (m, H_2-4), 2.92 (s, H_2-6), 1.48 (s, H_3-8 and H_3-9), 3.29 (dd, J = 13.6, 3.6 Hz, H-10a), 2.97 (dd, J =

13.6, 9.2 Hz, H-10b), 3.70 (dd, *J* = 9.2, 3.6 Hz, H-11); HR-ESI-MS for the $[M + H]^+$ ion at *m/z* 323.1094 (calculated $[M + H]^+$ ion for $C_{12}H_{22}N_2O_4S_2$ at *m/z* 323.1098) (Figs. S11, S12).

Compound ***10***: colorless oil; $[\alpha]^{26}_D$ – 8.3 (*c* 0.12, H_2O); UV (H_2O) λ_{max} (log ε) 237 (3.58) nm; ^{1}H NMR (700 MHz, methanol-d_4) (carbon labeling shown in Fig. S15): δ 3.48 (d, *J* = 11.0 Hz, H-1a), 3.40 (d, *J* = 11.0 Hz, H-1b), 3.92 (s, H-3), 3.50 (m, H_2-5), 2.44 (m, H_2-6), 3.37 (m, H_2-8), 3.05 (m, H_2-9), 2.92 (s, H_2-11), 1.48 (s, H_3-13 and H_3-14), 0.94 (s, H_3-15 and H_3-16), 3.28 (dd, *J* = 13.6, 3.7 Hz, H-17a), 2.99 (dd, *J* = 13.7, 9.1 Hz, H-17b), 3.72 (dd, *J* = 9.0, 3.6 Hz, H-18); ^{13}C NMR (175 MHz, methanol-d_4): δ 68.9 (C-1), 39.0 (C-2), 75.9 (C-3), 174.6 (C-4), 35.1 (C-5), 35.0 (C-6), 172.5 (C-7), 38.5 (C-8), 28.3 (C-9), 196.2 (C-10), 54.3 (C-11), 44.4 (C-12), 27.9 (C-13, 14), 19.9 (C-15), 19.5 (C-16), 29.3 (C-17), 54.7 (C-18), 171.1 (C-19); HR-ESI-MS for the $[M + H]^+$ ion at *m/z* 482.1994 (calculated $[M + H]^+$ ion for $C_{19}H_{35}N_3O_7S_2$ at *m/z* 482.1993) (Figs. S15-S20).

Compound ***11***: colorless oil; UV (H_2O) λ_{max} (log ε) 242 (3.96), 267 (4.01) nm; ^{1}H NMR (400 MHz, methanol-d_4) (carbon labeling shown in Fig. S9): δ 1.94 (s, H_3-1), 3.34 (t, *J* = 6.7 Hz, H_2-3), 3.03 (t, *J* = 6.7 Hz, H_2-4), 6.05 (m, H-6), 1.91 (d, *J* = 1.2 Hz, H_3-8), 2.16 (d, *J* = 1.1 Hz, H_3-9); HR-ESI-MS for the $[M + H]^+$ ion at *m/z* 202.0895 (calculated $[M + H]^+$ ion for $C_9H_{15}NO_2S$ at *m/z* 202.0901), $[M + Na]^+$ ion at *m/z* 224.0713 (calculated $[M + Na]^+$ ion for $C_9H_{15}NO_2S$ at *m/z* 224.0720), $[2M + H]^+$ ion at *m/z* 403.1715 (calculated $[2M + H]^+$ ion for $C_{18}H_{30}N_2O_4S_2$ at *m/z* 403.1724), $[2M + Na]^+$ ion at *m/z* 425.1533 (calculated $[2M + Na]^+$ ion for $C_{18}H_{30}N_2O_4S_2$ at *m/z* 425.1544) (Figs. S9, S10).

Compound ***15***: colorless oil; $[\alpha]^{26}_D$ + 23.7 (*c* 0.65, MeOH); UV (H_2O) λ_{max} (log ε) 242 (3.96), 267 (4.01) nm; ^{1}H NMR (400 MHz, methanol-d_4) (carbon labeling shown in Fig. S13): δ 3.48 (d, *J* = 11.0 Hz, H-1a), 3.40 (d, *J* = 11.0 Hz, H-1b), 3.91 (s, H-3), 3.49 (m, H_2-5), 2.43 (t, *J* = 6.7 Hz, H_2-6), 3.35 (m, H_2-8), 3.03 (t, *J* = 6.7 Hz, H_2-9), 6.06 (m, H-11), 1.91 (d, *J* = 1.2 Hz, H_3-13), 2.17 (d, *J* = 1.1 Hz, H_3-14), 0.94 (s, H_3-15 and H_3-16); HR-ESI-MS for the $[M + H]^+$ ion at *m/z* 361.1789 (calculated $[M + H]^+$ ion for $C_{16}H_{28}N_2O_5S$ at *m/z* 361.1796), $[M + Na]^+$ ion at *m/z* 383.1610 (calculated $[M + Na]^+$ ion for $C_{16}H_{28}N_2O_5S$ at *m/z* 383.1615) (Figs. S13, S14).

Detection of compounds 16 and 17. The detection of compounds **16** and **17** produced by the reactions of LnmJ-SH with compound **9** and **10** is based on the HPLC analysis, using the reaction with boiling LnmJ-SH as the negative control (Fig. S21). The 100 μL reaction system contained 50 mM Tris-HCl (pH 7.2), 15 mM compound **9** or **10**, 150 μM LnmJ-SH, 0.2 mM PLP, and 20 mM KCl, and was incubated at 30 ºC for 20 min. The reaction solution was terminated by addition of 200 μL MeOH, centrifuged and loaded on HPLC analysis. The product, compound **16** or **17**, was purified from HPLC with multiple repeated reactions and their structures were identified based on NMR and HR-ESI-MS analysis (Figs. S23-S28).

Compound ***16***: colorless oil; UV (H_2O) λ_{max} (log ε) 237 (3.40) nm; ^{1}H NMR (700 MHz, methanol-d_4) (carbon labeling shown in Fig. S23): δ 1.94 (s, H_3-1), 3.36 (t, *J* = 6.7 Hz, H_2-3), 3.04 (t, *J* = 6.7 Hz, H_2-4), 2.94 (s, H_2-6), 1.49 (s, H_3-8 and H_3-9); ^{13}C NMR (175 MHz, methanol-d_4): δ 21.1 (C-1), 172.0 (C-2), 38.7 (C-3), 28.0 (C-4), 196.4 (C-5), 58.1 (C-6), 41.3 (C-7), 31.5 (C-8 and C-9); HR-ESI-MS for the $[M + H]^+$ ion at *m/z* 236.0776 (calculated $[M + H]^+$ ion for $C_9H_{17}NO_2S_2$ at *m/z* 236.0778) and $[M + Na]^+$ ion at *m/z* 258.0596 (calculated $[M + Na]^+$ ion for $C_9H_{17}NO_2S_2$ at *m/z* 258.0598) (Figs. S23-S22).

Compound ***17***: colorless oil; UV (H_2O) λ_{max} (log ε) 237 (3.42) nm; HR-ESI-MS for the $[M + H]^+$ ion at *m/z* 395.1670 (calculated $[M + H]^+$ ion for $C_{16}H_{30}N_2O_5S_2$ at *m/z* 395.1673) and $[M + Na]^+$ ion at *m/z* 417.1490 (calculated $[M + Na]^+$ ion for $C_{16}H_{30}N_2O_5S_2$ at *m/z* 417.1493) (Fig. S28).

Optimization of LnmJ-SH activity. To estimate the pH optimum for LnmJ-SH, cleavage of L-cysteine (**1**) was carried out in 100 μL of reaction mixture contained 30 μM LnmJ-SH, 4 mM L-cysteine (**1**), 0.2 mM PLP, and 50 mM KCl in 50 mM phosphate buffer or 50 mM Tris-HCl, with varying pH values from pH 6.5 to 9.0 (Fig. S29*A*). The effect of KCl on the activity of LnmJ-SH was determined by varying the

concentration of KCl from 0 - 50 mM, in 100 µL reaction mixtures containing 30 µM LnmJ-SH, 4 mM L-cysteine (**1**), 0.2 mM PLP in 50 mM Tris-HCl (pH 8.5) (Fig. S29*B*). To evaluate the effect of PLP concentration on the activity of LnmJ-SH, 100 µL reaction mixtures contained 30 µM LnmJ-SH, 4 mM L-cysteine (**1**), 20 mM KCl in 50 mM Tris-HCl (pH 8.5) with a range of PLP from 0 - 0.3 mM were carried out (Fig. S29*C*). All reactions were incubated at 30 ºC for 1 h and were carried out in duplicates. The reaction products ammonium and pyruvate were analyzed by HPLC after derivatization with OPD as noted above (Fig. S4).

Determination of the kinetic parameters for LnmJ-SH, CaJ-SH, and MaJ-SH. The 100 µL reaction mixtures contained 30 µM LnmJ-SH, 0.2 mM PLP, 20 mM KCl, in 100 mM Tris-HCl (pH 8.5), with different substrates of varied concentrations. To determine kinetic parameters using L-cysteine (**1**) as the substrate, the concentration of L-cysteine (**1**) ranged from 0.4 to 100 mM in the reaction mixture. To determine kinetic parameters using L-cystine (**2**) as substrate, L-cystine (**2**) concentrations varied from 1 to 10 mM. The kinetic parameters using S-phenyl-L-cysteine (**4**) as a substrate for LnmJ-SH was determined by varying the concentration of substrate **4** from 1 to 5 mM. To determine kinetic parameters using compound **9** and **10** as the substrates, the reaction mixtures were same as above but the pH was changed into 7.2 due to the instability issues of **9** and **10** at high pH values. The concentration of **9** and **10** varied from 1 to 100 mM. All assays were carried out at 30 °C for 10 min and in duplicates on different days. The reaction products ammonium and pyruvate were subjected to analysis by methods as described above. The K_ms and k_{cat}s for L-cysteine (**1**), **9,** and **10** were generated using the Michaelis-Menten equation by nonlinear regression analysis using KaleidaGraph (Synergy Software) for LnmJ-SH, CaJ-SH and MaJ-SH. For other substrates a linear fit through 1, 3, and 6 mM concentrations produced linear results approximating k_{cat}/K_m (selected Michaelis-Menten curves determined at pH 7.2 were shown in Fig. S30).

The selected enzymes and domains (accession numbers shown in the brackets) used for the phylogenetic analysis shown in Fig. 2: (i) Members of the newly characterized PKS-SH domains A-SH (EWC63729) from *Actinokineospora* sp. EG49, Am-SH (WP_027413080) from *Aquimarina muelleri*, Bp1-SH (CAH38470) from *Burkhoderia pseudomallei* K96243, Bp2-SH (KGW99043) from *Burkhoderia pseudomallei* MSHR456, CaJ-SH (ACU71516) from *Catenulispora acidiphilia*, Fi-SH (WP_009579363) from *Fulvivirga imtechensis*, Ka-SH (EDP94334) from *Kordia algicida*, L-SH (AEH59109) from *Lysobacter* sp. ATCC 53042, LnmJ-SH (AAN85523) from *S. atroolivaceus*, M-SH (ADU09998) from *Micromonospora* sp. L5, MaJ-SH (ADL47419) from *Micromonospora aurantiaca*, Pf-SH (WP_031296991) from *Pesudoqulbenkiania ferrooxidans*, Sa-SH (WP_032717744) from *Salinispora arenicola*, Se-SH (WP_015102008) from *Saccharothrix espanaensis*, Sl-SH (WP_029381732) from *Streptomyces leeuwenhoekii*, Sm-SH (WP_028981298) from *Sporocytophaga myxococcoides;* (ii) ABL members TIL (1) (ACB04752) from *E. coli*, TIL (2) (WP_015319350) from *Natronococcus occultus*, TIL (3) (AAA25664) from *Proteus vulgaris*, TPL (1) (ELT53913) from *Aggregatibacter actinomycetemcomitans*, TPL (2) (ABI75028) from *Citrobacter freundii*, TPL (3) (WP_010696718) from *Treponema denticola*; and (iii) C-S bond cleavage enzymes alliinase (1) (AAA32639) from *Allium cepa*, alliinase (2) (Q41233) from *Allium sativaum*, alliinase (3) (BAA20358) from *Allium tuberosum*, CalE6 (AAM94792) from *Micromonospora echinospora*, CscB (1JF9_A) from *E. coli*, cysteine *S*-conjugate β-lyase (1) (WP_002979883) from *Chryseobacterium gleum*, cysteine *S*-conjugate β-lyase (2) (Q08415) from *Rattus norvegicus*, cysteine desulfhydrase (1) (WP_003060086) from *Streptococcus dysgalactiae*, cysteine desulfhydrase (2) (WP_001128188) from *Salmonella enterica*, cysteine desulfhydrase (3) (WP_004934750) from *Serratia marcescens*, GliI (AAW03309) from *Aspergillus fumigatus*, IscS (EEJ46382) from *E. coli*, NifS (1ECX_A) from *Thermotoga maritime*, SUR1 (NP_179650) from *Arabidopsis thaliana*.

References:

1. Cheng YQ, Tang GL, Shen B (2002) Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in *Streptomyces atroolivaceus* S-140. *J Bacteriol* 184(24):7013-7024.

2. Sambrook J, Russell D (2001) *Molecular cloning: a laboratory manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, United States.
3. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) *Practical Streptomyces genetics,* John Innes Foundation, Norwich, United Kingdom.
4. Lohman JR, Bingman CA, Phillips GN, Shen B (2013) Structure of the bifunctional acyltransferase/decarboxylase LnmK from the leinamycin biosynthetic pathway revealing novel activity for a double-hot-dog fold. *Biochemistry* 52(5):902-911.
5. Ramirez EC, Whitaker JR (1999) Biochemical characterization of cysteine lyase from broccoli (*Brassica oleracea var. italica*). *J Agric Food Chem* 47(6):2218-2225.
6. Stijntjes GJ, te Koppele JM, Vermeulen NP (1992) High-performance liquid chromatography fluorescence assay of pyruvic acid to determine cysteine conjugate β-lyase activity: application to S-1,2-dichlorovinyl-L-cysteine and S-2-benzothiazolyl-L-cysteine. *Anal Biochem* 206(2):334-343.
7. Koike K, Koike M (1984) Fluorescent analysis of α-keto acids in serum and urine by high-performance liquid chromatography. *Anal Biochem* 141(2):481-487.
8. Cline JD (1969) Spectrophotometric determination of hydrogen sulfide in natural waters. *Limnology and Oceanography* 14(3):454-458.
9. Ukai K, Sekiya J (1999) Rapid purification and characterization of cysteine lyase b from broccoli inflorescence. *Phytochemistry* 51(7):853-859.
10. Fahey RC, Newton GL (1987) Determination of low molecular weight thiols using monobromobimane fluorescent labeling and high performance liquid chromatography. *Methods Enzymol* 143:85-96.
11. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22(2):195-201.

Fig. S1. The ABL family of PLP-dependent enzymes that catalyze Cβ-Cγ bond cleavage as exemplified by tyrosine phenol lyases (TPLs) and tryptophan indole lyases (TILs) in comparison with the LnmJ-SH family of PKS domains that catalyze C-S bond cleavage of L-cysteine and L-cysteine *S*-conjugatesas substrates.

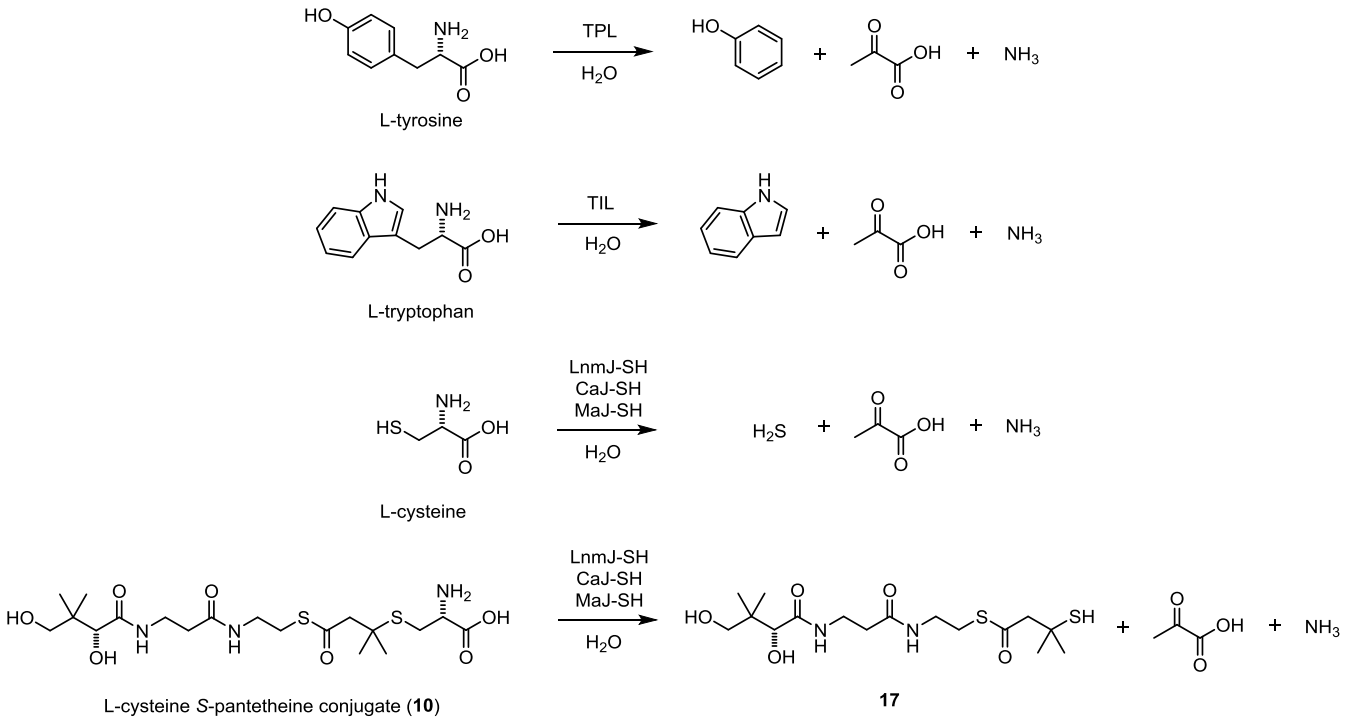


Fig. S2. Sequence alignments of selected SH domains and ABLs. Selected proteins or domains for the alignments are (organisms, accession number and percent similarity/identity to LnmJ-SH were shown in brackets): J-SH (LnmJ-SH, *Streptomyces atroolivaceus* S-140, AAN85523), M.L5 (*Micromonospora* sp. L5, ADU09998, 42/51%), Ma-SH (*Micromonospora aurantiaca*, ADL47419, 45/56%), Ca-SH (*Catenulispora acidiphilia*, ACU71516, 40/52%), K.alg (*Kordia algicida*, EDP94334.1, 29/45%), B.ps1 (*Burkholderia pseudomallei* K96243, CAH38467, 28/38%), B.ps2 (*Burkholderia pseudomallei* K96243, CAH38470, 25/36%), TPL 1 (*Citrobacter freundii*, ABI75028, 25/40%), TPL 2 (*Erwinia herbicola*, AAA66390, 24/39%), TIL 1 (*Escherichia coli*, ACB04752, 21/33%), TIL 2 (*Proteus vulgaris*, AAA25664, 23/37%). Features are annotated by: * for PLP binding residues, P for potassium binding, K for active site lysine, S for substrate binding, A for a conserved arginine that anchors the N- and C-terminal regions of the small domain.

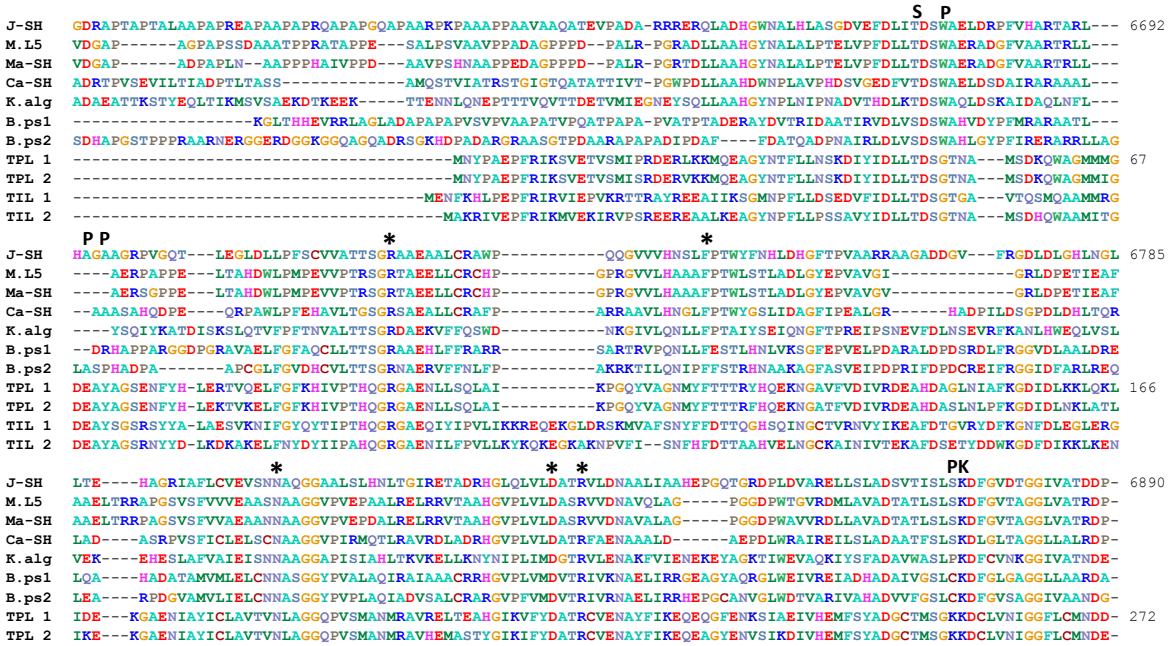
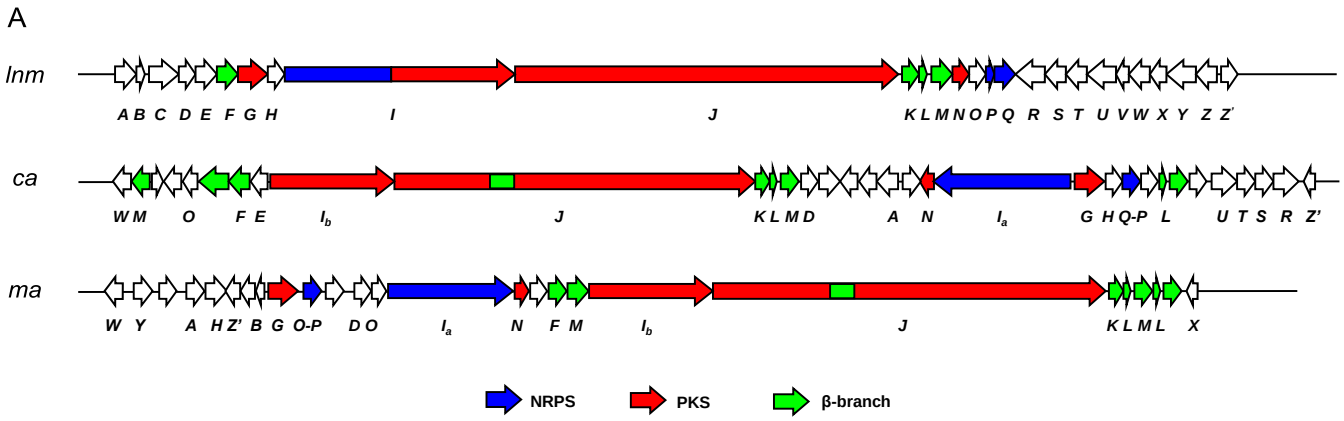
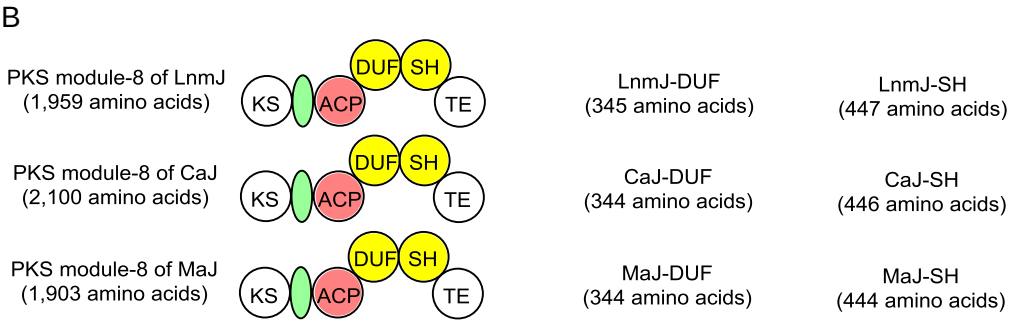


Fig. S3. Comparison of the *lnm* biosynthetic gene cluster from *S. atroolivaceus* with selected clusters that feature the new family of PKS domains that catalyze C-S bond cleavage as exemplified by LnmJ-SH, CaJ-SH from *Catenulispora acidiphilia* DSM 44928, and MaJ-SH from *Micromonospora aurantiaca*, ATCC 27092. (*A*) Comparison of the organization of the *lnm* biosynthetic gene cluster with the homologous *ca* cluster from *C. acidiphilia* DSM 44928 and *ma* cluster from *M. aurantiaca*, ATCC 27092. Genes annotated to encode the NRPS (blue) and PKS (red and green) responsible for the biosynthesis of the hybrid peptide-polyketide backbones of LNM and its homologues are color-coded to highlight the sequence similarity among the three clusters. (*B*) Architectural conservation of the PKS domains within PKS module-8 of LnmJ and its homologues CaJ and MaJ and amino acid sequence comparison among the three selected PKS module of LnmJ, CaJ, and MaJ (% amino acid identity/% amino acid similarity).





	LnmJ module-8	LnmJ-DUF	LnmJ-SH	CaJ module-8	CaJ-DUF	CaJ-SH	MaJ module-8	MaJ-DUF	MaJ-SH
LnmJ module-8	-	-	-	45/65	-	-	50/71	-	-
LnmJ-DUF	-	-	-	-	57/72	-	-	55/69	-
LnmJ-SH	-	-	-	-	-	40/52	-	-	45/56
CaJ module-8	45/65	-	-	-	-	-	48/67	-	-
CaJ-DUF	-	57/72	-	-	-	-	-	60/72	-
CaJ-SH	-	-	40/52	-	-	-	-	-	45/56
MaJ module-8	50/71	-	-	48/67	-	-	-	-	-
MaJ-DUF	-	55/69	-	-	60/72	-	-	-	-
MaJ-SH	-	-	45/56	-	-	45/56	-	-	-

Fig. S4. Detection of pyruvate produced from LnmJ-SH catalyzed β-elimination reactions in vitro. *(A)* Pyruvate reaction with o-phenylenediamine (OPD) to form methylquinoxalinol. *(B)* HPLC analysis of LnmJ-SH reaction products, derivatized by reaction with OPD: (I) methylquinoxalinol standard (●) from the reaction of pyruvate with OPD; (II) negative control, L-cysteine (**1**) as the substrate with boiled enzyme; (III) L-cysteine (**1**) as the substrate; (IV) L-cystine (**2**) as substrate; (V) *S*-phenyl-L-cysteine (**4**) as the substrate; (VI) L-tyrosine as the substrate; (VII) L-tryptophan as the substrate; (VIII) D-cysteine as the substrate.

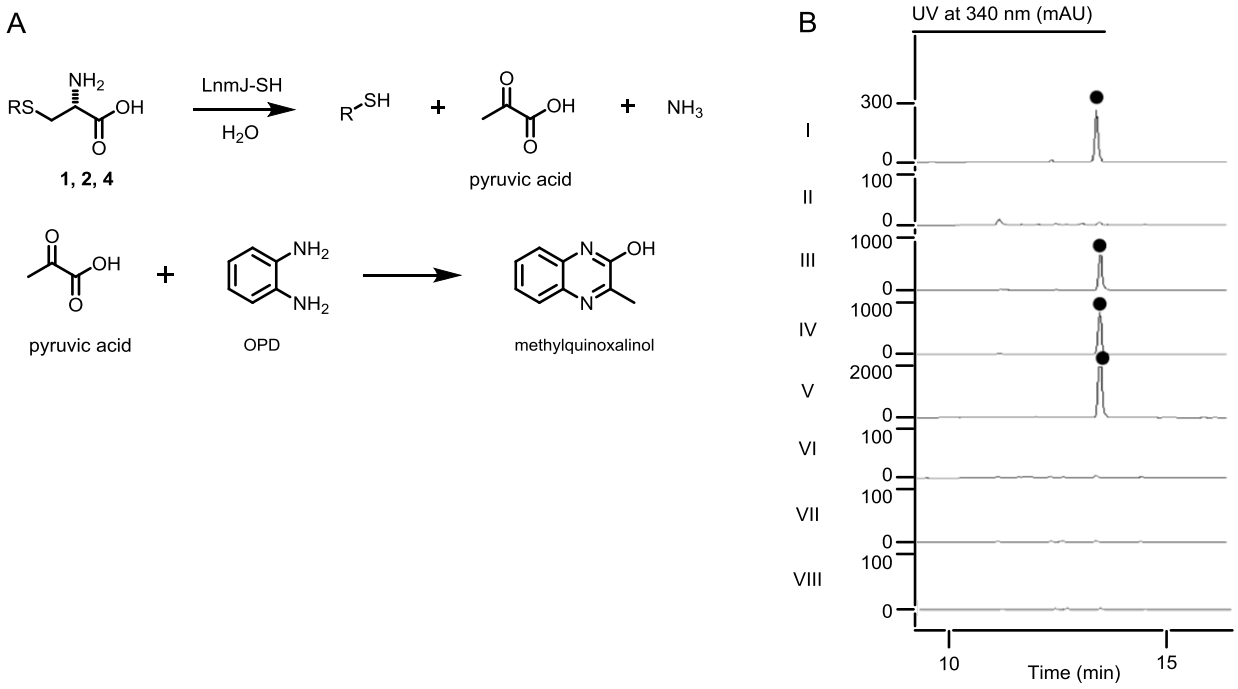


Fig. S5. Detection of the hydrogen sulfide produced from the reaction of LnmJ-SH with L-cysteine (**1**) by methylene blue colorimetric method.

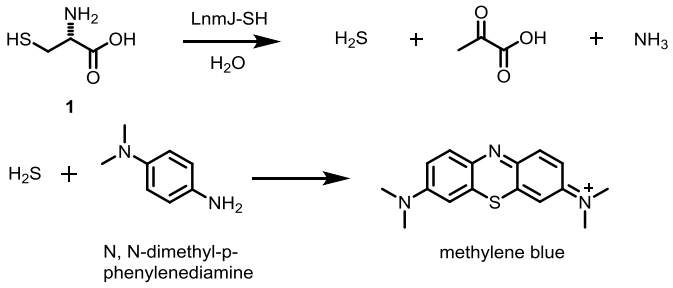


Fig. S6. Detection of thiocysteine produced from the reaction of LnmJ-SH with L-cystine (**2**). *(A)* LnmJ-SH catalyzed the β-elimination of L-cystine (**2**) to produce thiocysteine, pyruvate, and ammonia, and thiocysteine was derivatized with monobromobimane (MBB) to form a thiocysteine-monobimane conjugate. *(B)* HPLC analysis of the thiocysteine-monobimane conjugate (●). (I) Reaction of LnmJ-SH with L-cystine (**2**) (▼); (II) Reaction of boiled LnmJ-SH with L-cystine (**2**) as the negative control.

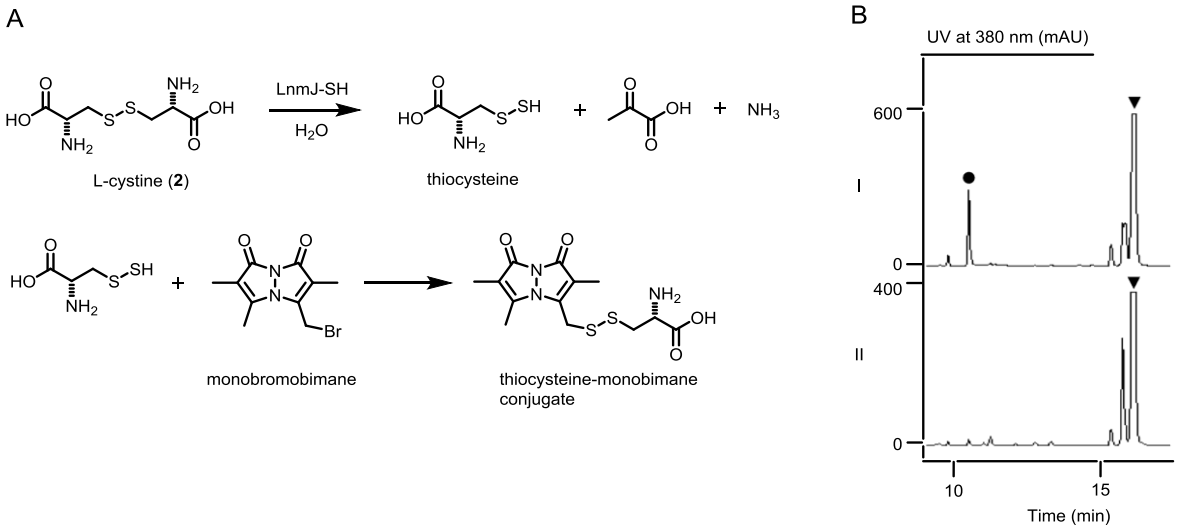


Fig. S7. Detection of thiophenol produced from the reaction of LnmJ-SH with *S*-phenyl-L-cysteine (**4**). *(A)* LnmJ-SH catalyzed *S*-phenyl-L-cysteine (**4**) to produce thiophenol, pyruvate, and ammonia. *(B)* HPLC analysis of thiophenol. (I) thiophenol standard (♦); (II) reaction of boiled LnmJ-SH with *S*-phenyl-L-cysteine (**4**) (◊) as the negative control; (III) reaction of LnmJ-SH with *S*-phenyl-L-cysteine (**4**).

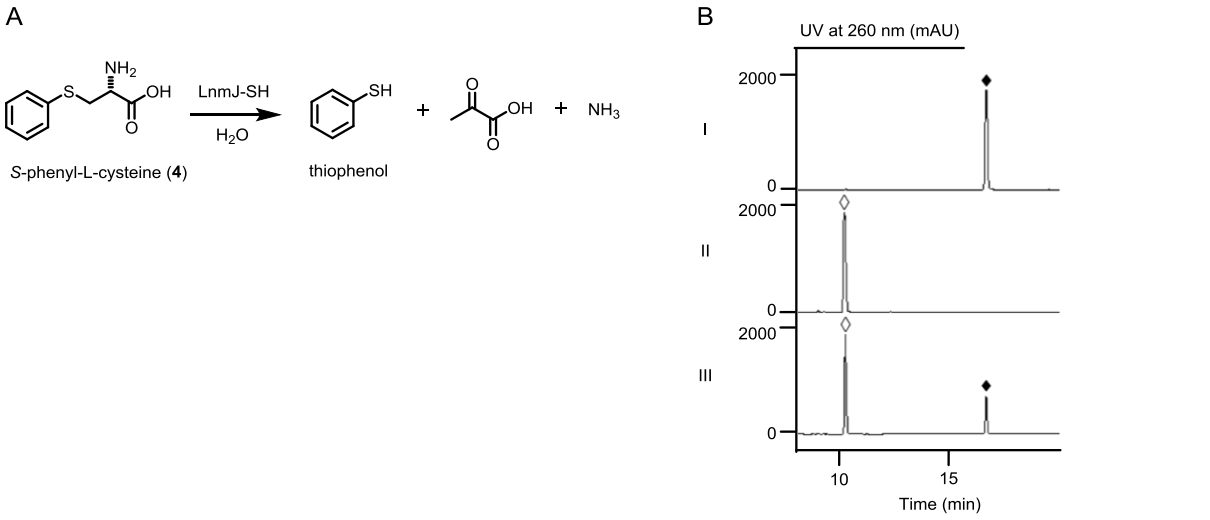


Fig. S8. Chemical synthesis of compound **9** and **10**. *(A)* Chemical synthesis of compound **9**. (a) Dimethylacryl acid (5.0 equiv), EDC (5.0 equiv), DMAP (0.5 equiv), CH_2Cl_2, rt, 3 h; (b) L-cysteine (10.0 equiv), H_2O, rt, 3 h. *(B)* Chemical synthesis of compound **10**. (a) CSA (0.02 equiv), 2-methoxypropene (10.0 equiv), THF, rt, 18 h; (b) DTT (10.0 equiv), MeOH/H_2O (1/1), rt, 1 h; (c) Dimethylacryl acid (5.0 equiv), EDC (5.0 equiv), DMAP (0.5 equiv), CH_2Cl_2, rt, 3 h; (d) TFA/MeCN/H_2O (5/1/1), rt, 20 min; (e) L-cysteine (10.0 equiv), H_2O, rt, 3 h.

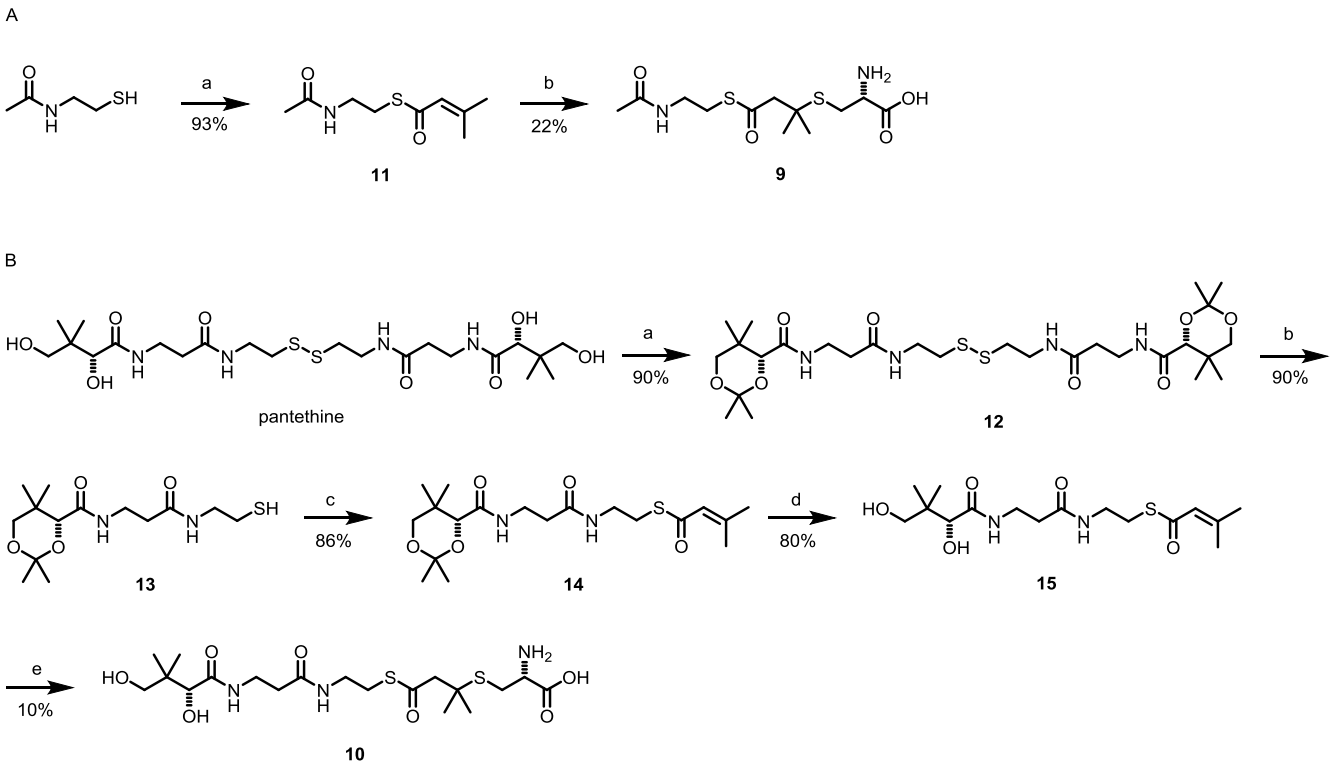


Fig. S9. The ^{1}H NMR (400 MHz) spectrum of compound **11** in methanol-d_4

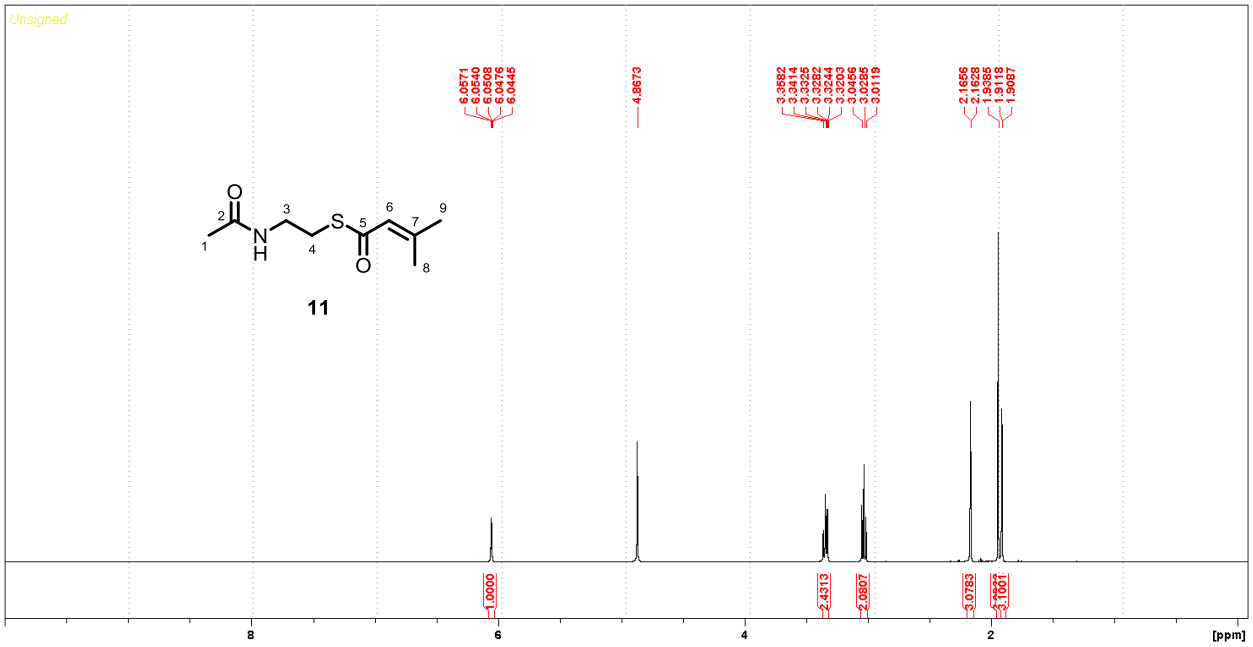


Fig. S10. The HR-ESI-MS of compound **11**

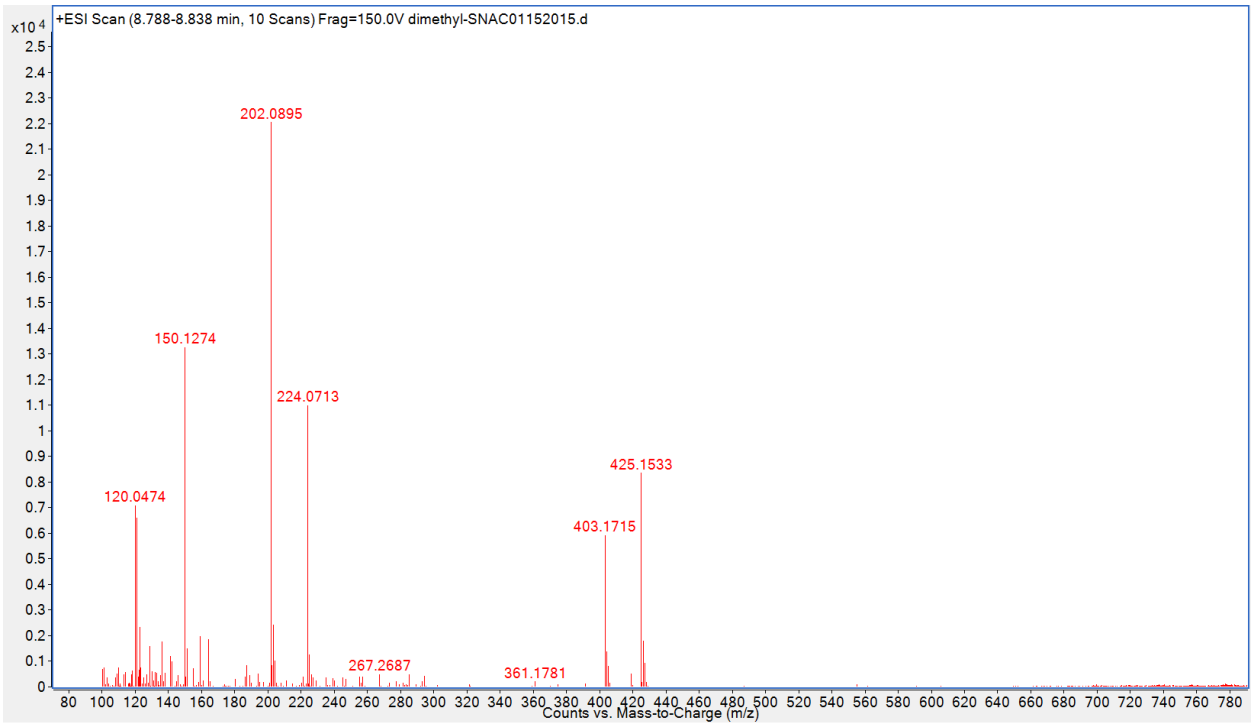


Fig. S11. The ^{1}H NMR (400 MHz) spectrum of compound **9** in methanol-d_4

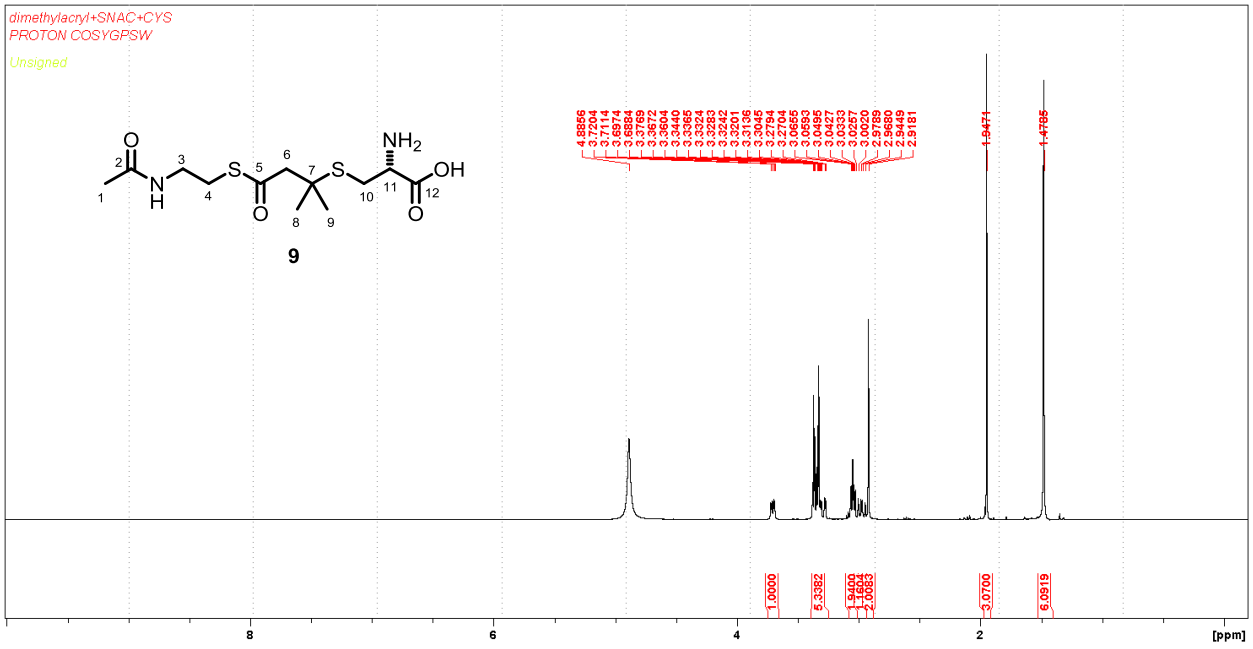


Fig. S12. The HR-ESI-MS of compound **9**

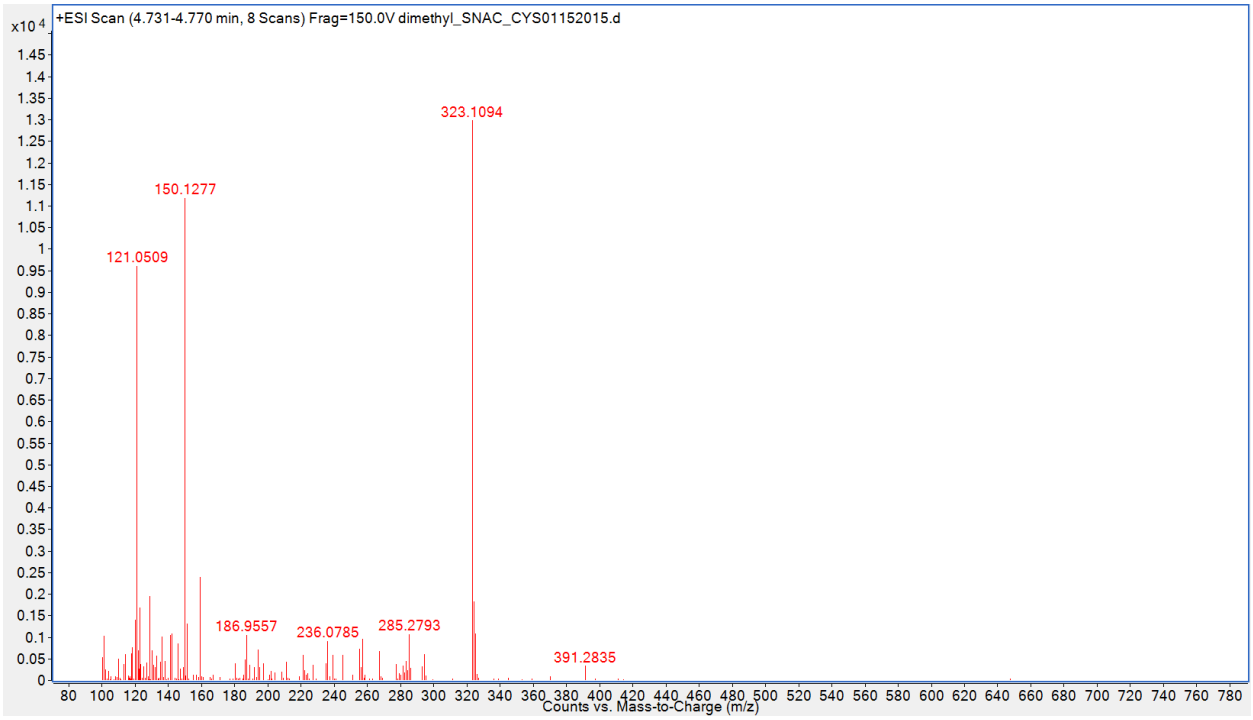


Fig. S13. The ^{1}H NMR (400 MHz) spectrum of compound **15** in methanol-d_4

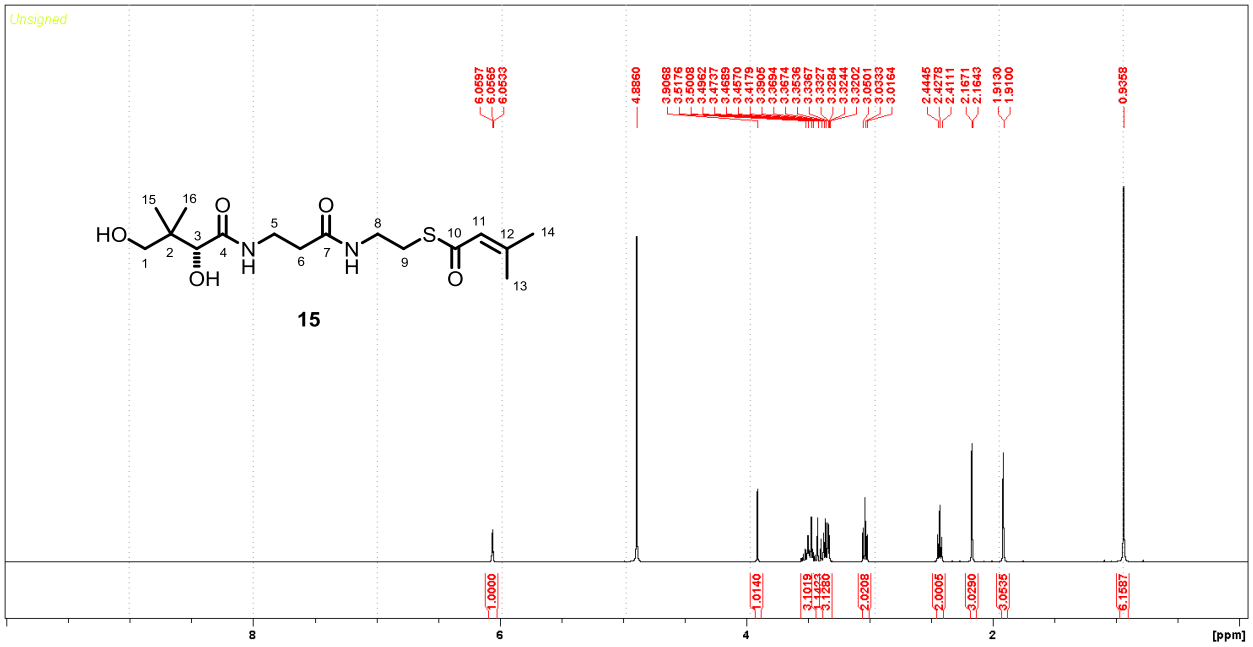


Fig. S14. The HR-ESI-MS of compound **15**

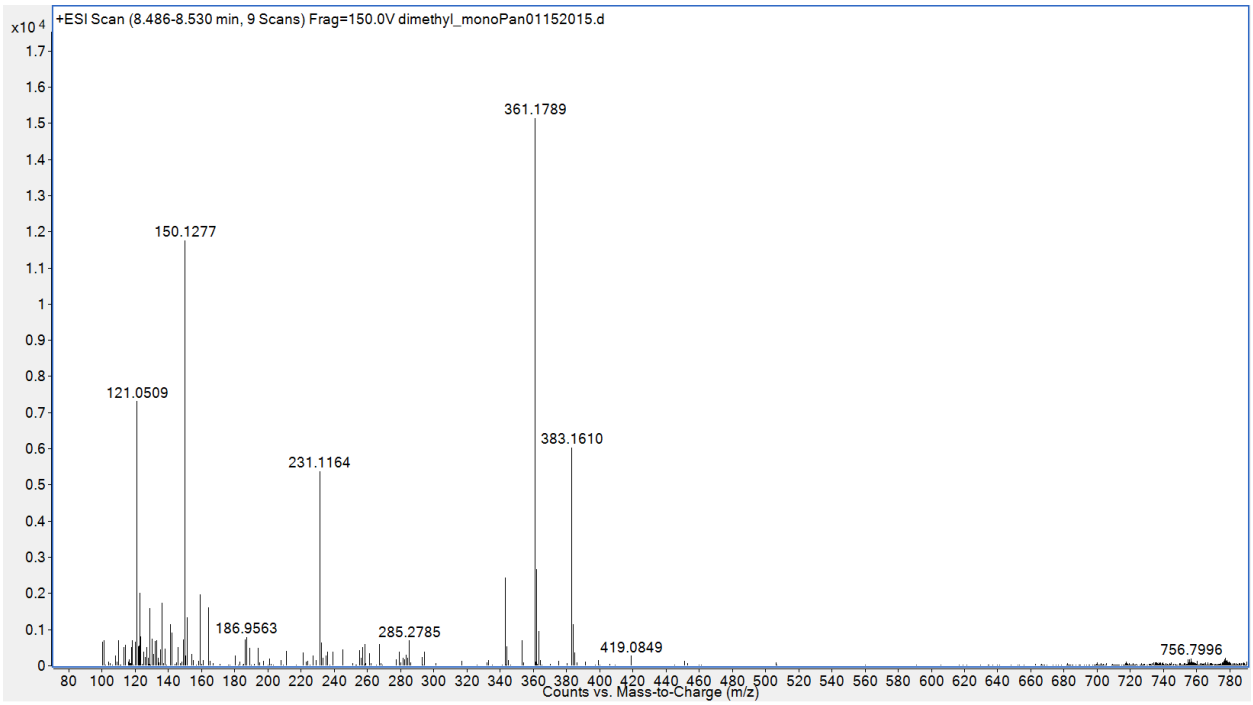


Fig. S15. The ^{1}H NMR (700 MHz) spectrum of compound **10** in methanol-d_4

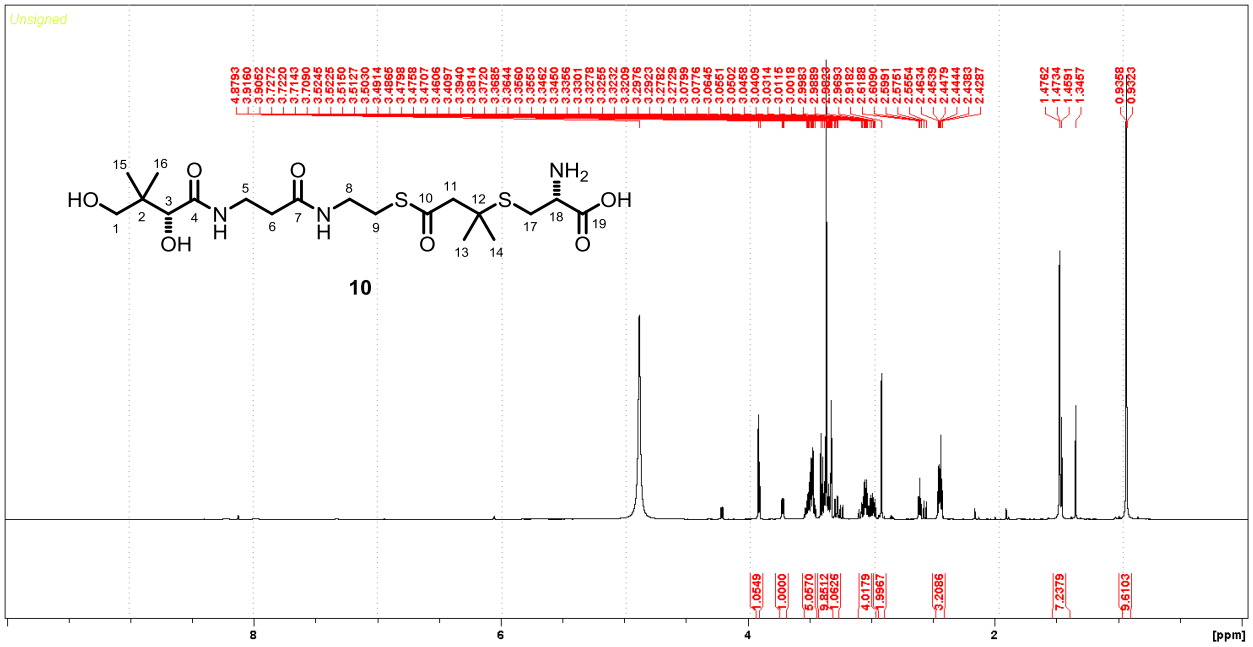


Fig. S16. The ^{1}H-^{1}H COSY spectrum of compound **10**

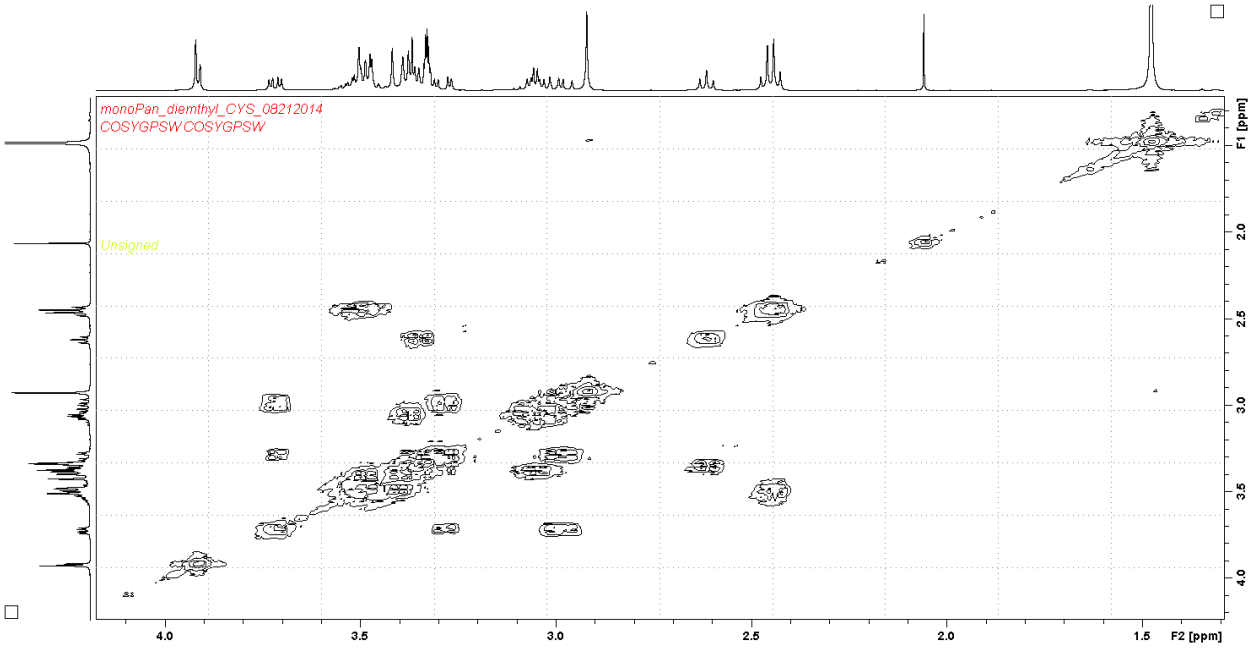


Fig. S17. The ^{13}C NMR (175 MHz) spectrum of compound **10** in methanol-d_4

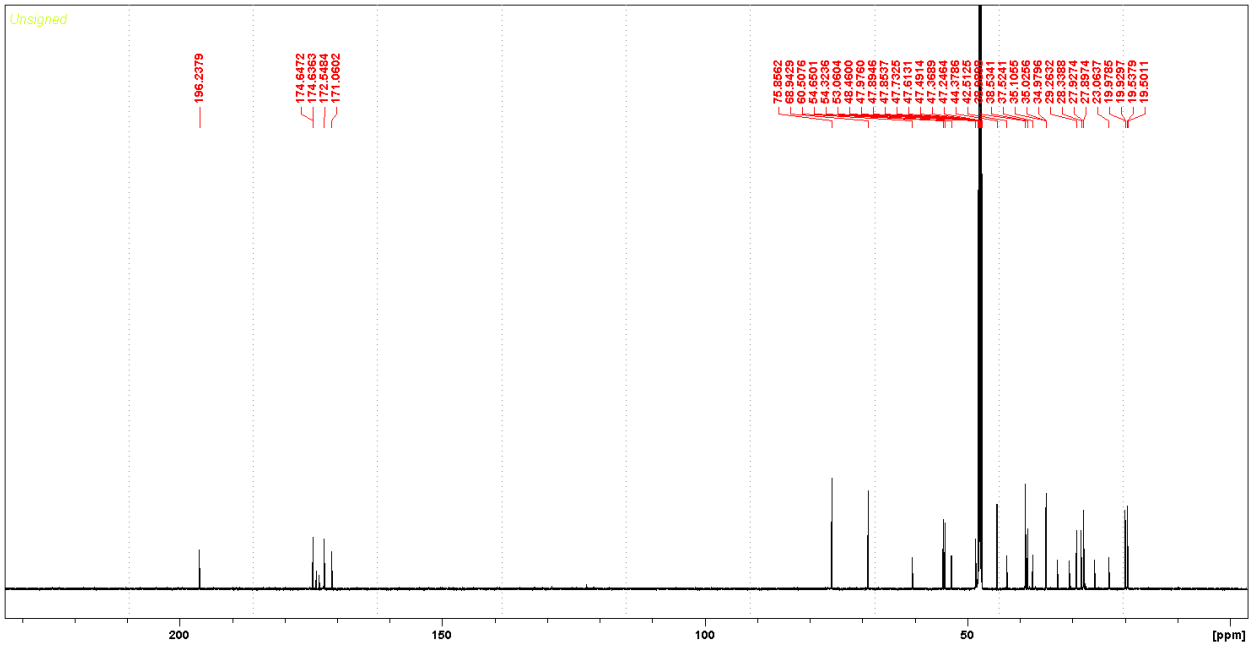


Fig. S18. The HSQC spectrum of compound **10**

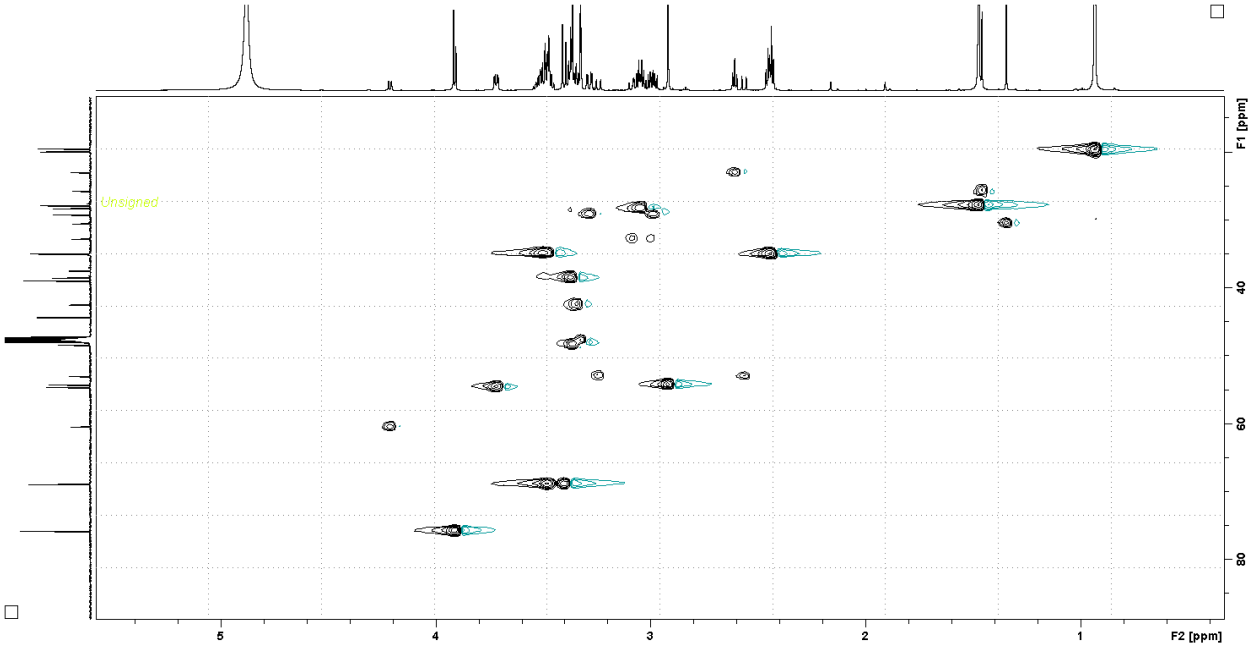


Fig. S19. The HMBC spectrum of compound **10**

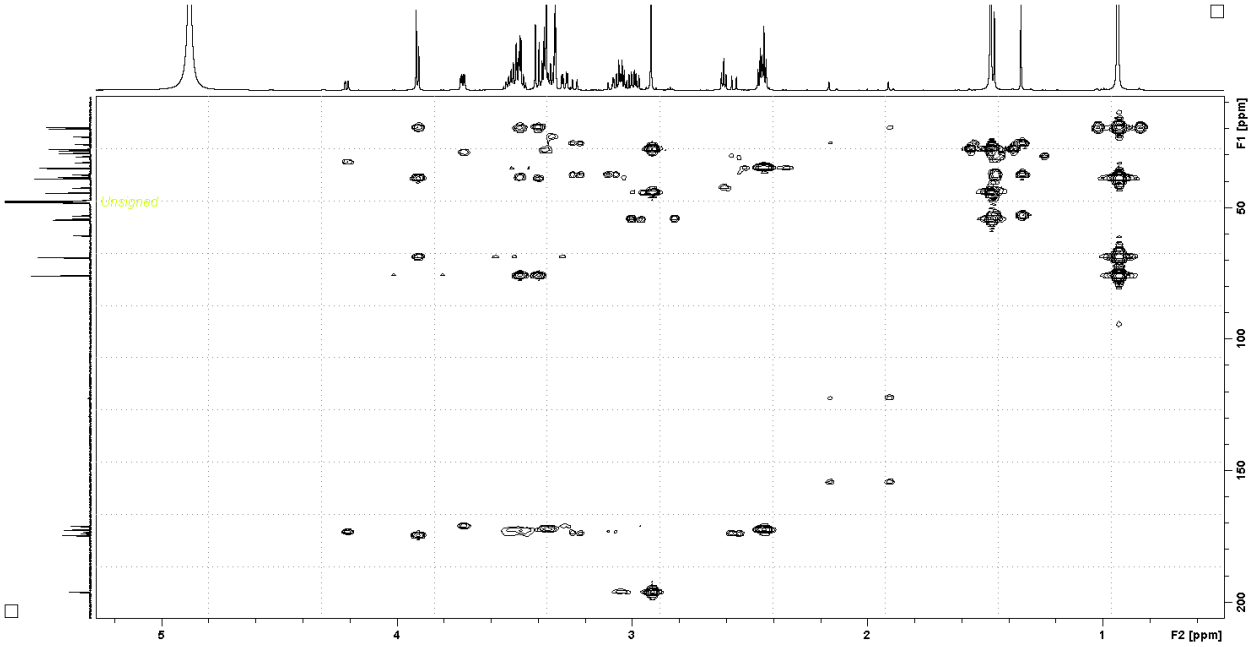


Fig. S20. The HR-ESI-MS of compound **10**

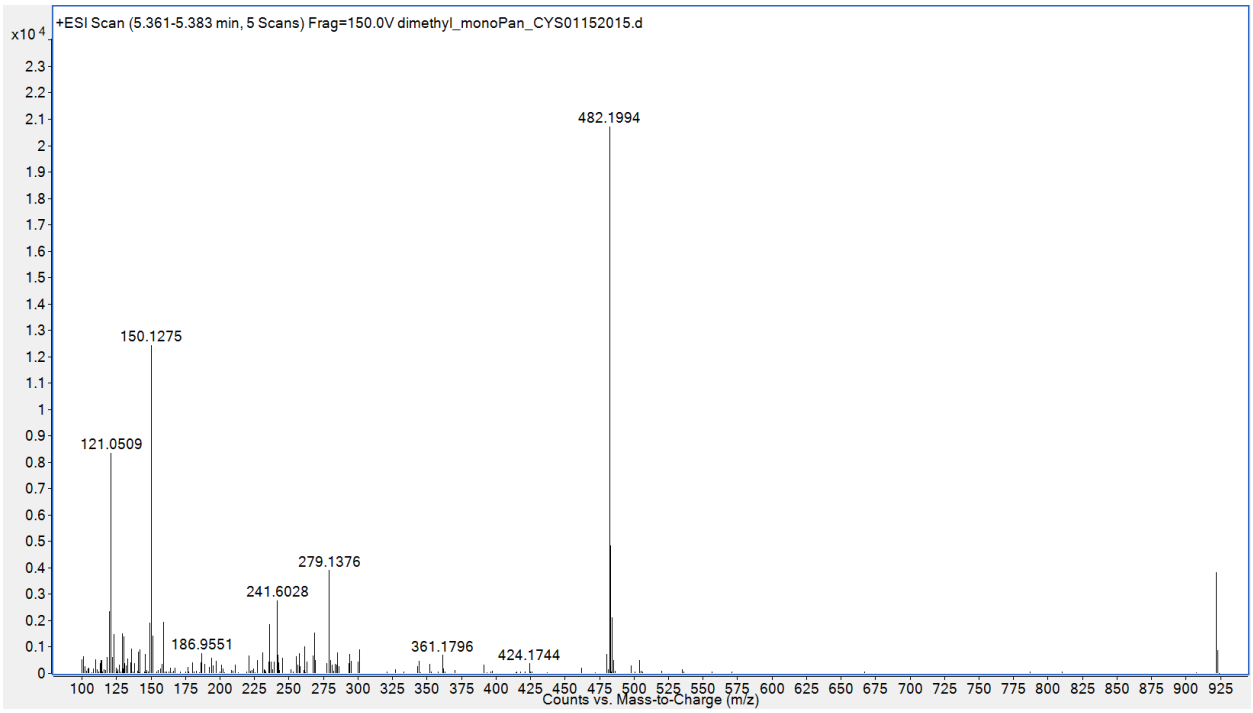
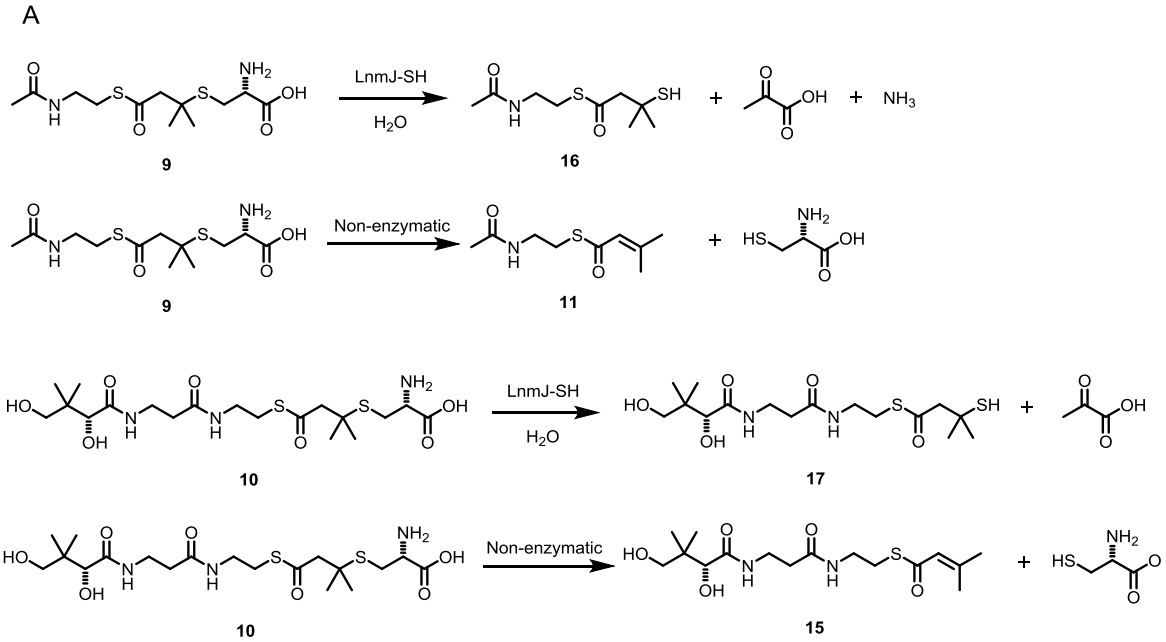


Fig. S21. The detection of compound **16** and **17** produced from the reaction of LnmJ-SH with compound **9** and **10**, respectively. *(A)* The reaction of LnmJ-SH with **9** and **10** produced the product **16** and **17**, respectively. *(B)* The HPLC analysis of the reaction. (I) Compound **9** standard (●); (II) the reaction of boiled LnmJ-SH with **9** as the negative control, with about 1% **9** non-enzymatically degraded into **11** (○); (III) the reaction of LnmJ-SH with **9** produced **16** (▼) as the product; (IV) Compound **10** standard (□); (V) the reaction of boiled LnmJ-SH with **10** as the negative control, with about 1% **10** non-enzymatically degraded into **15** (♦); (VI) the reaction of LnmJ-SH with **10** produced **17** (◊) as the product.



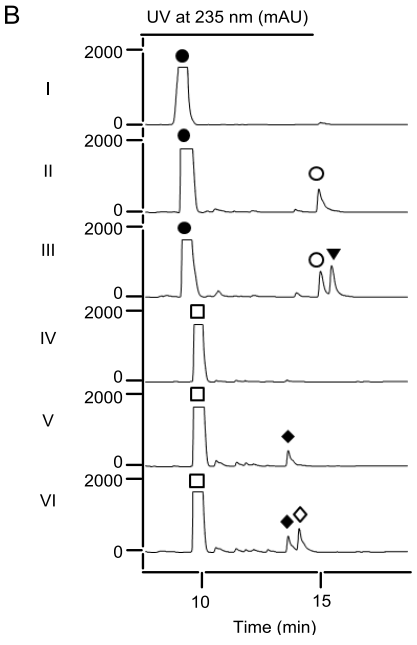


Fig. S22. The HR-ESI-MS of compound **16**

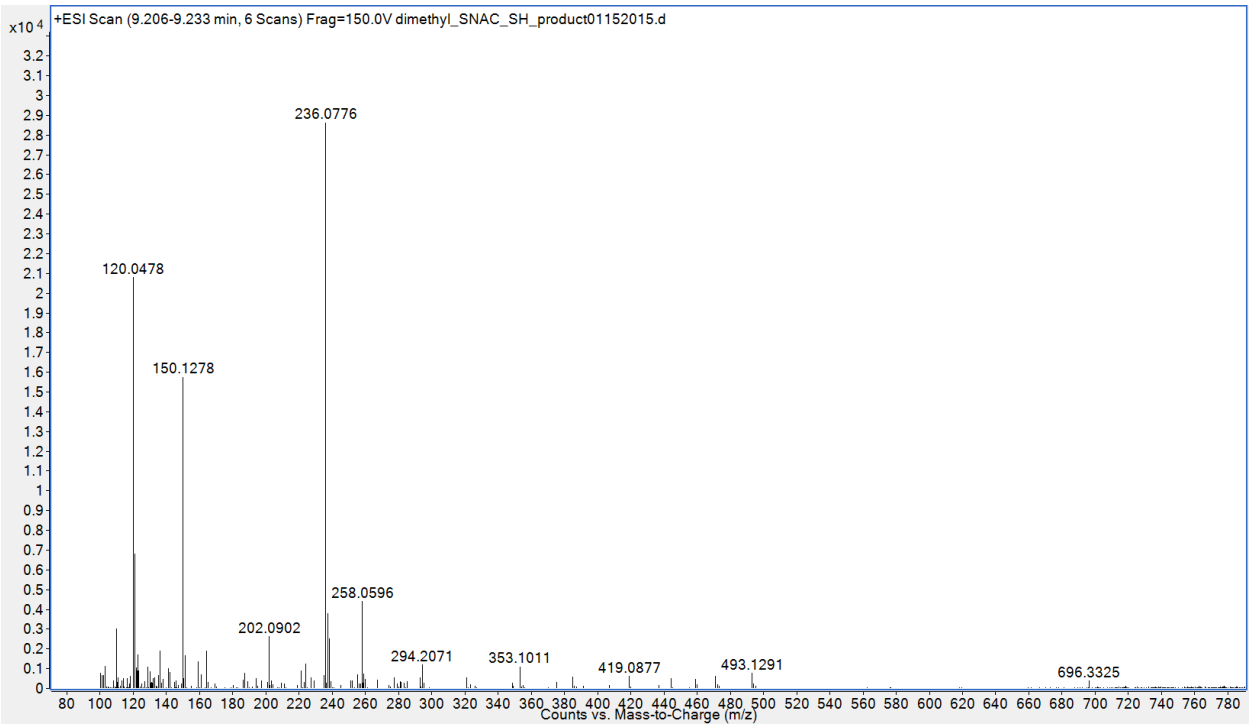


Fig. S23. The ^{1}H NMR (700 MHz) spectrum of compound **16** in methanol-d_4

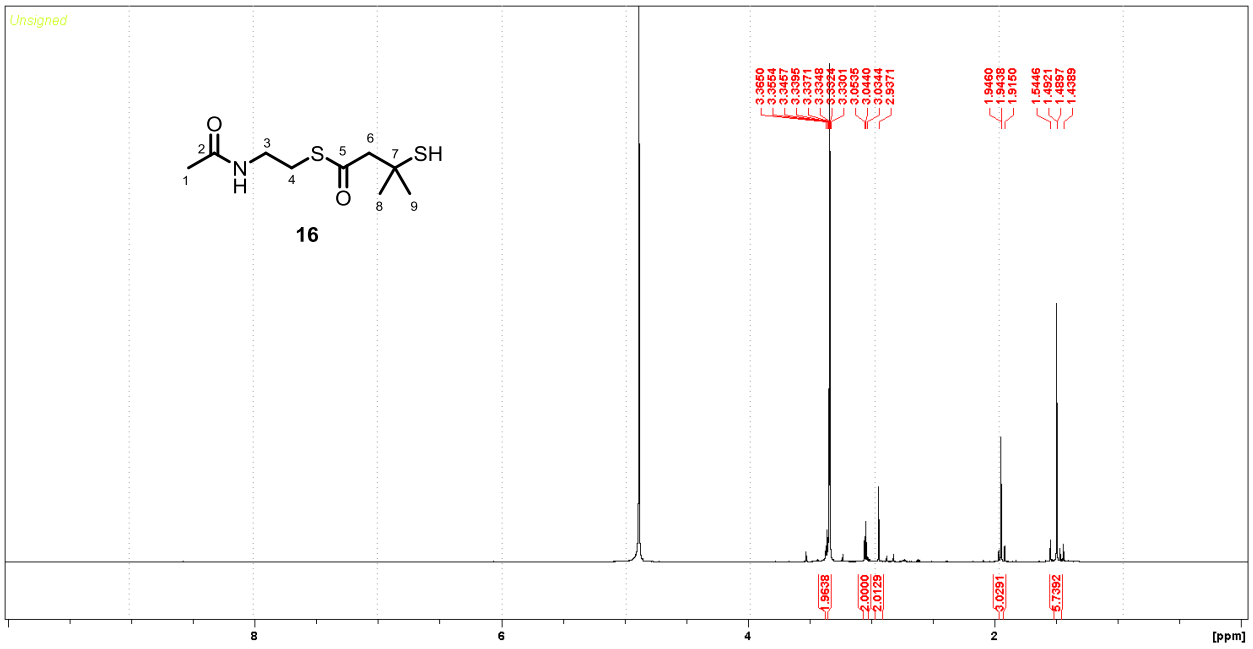


Fig. S24. The ^{1}H-^{1}H COSY spectrum of compound **16**

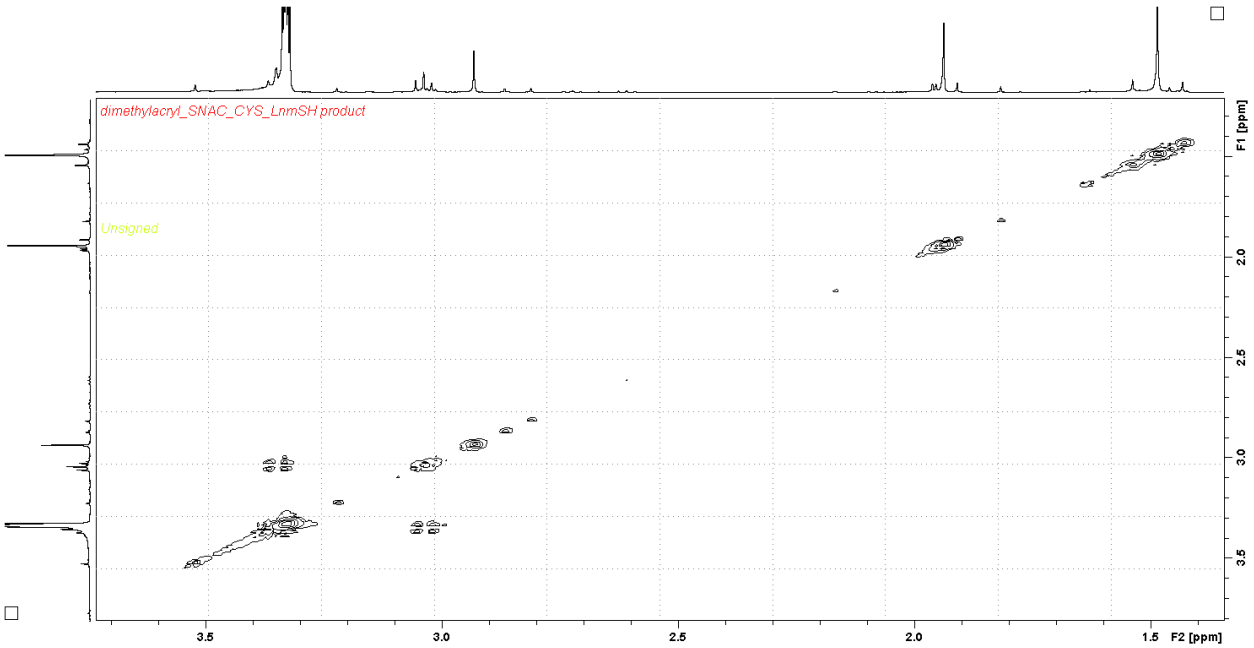


Fig. S25. The ^{13}C NMR (175 MHz) spectrum of compound **16** in methanol-d_4

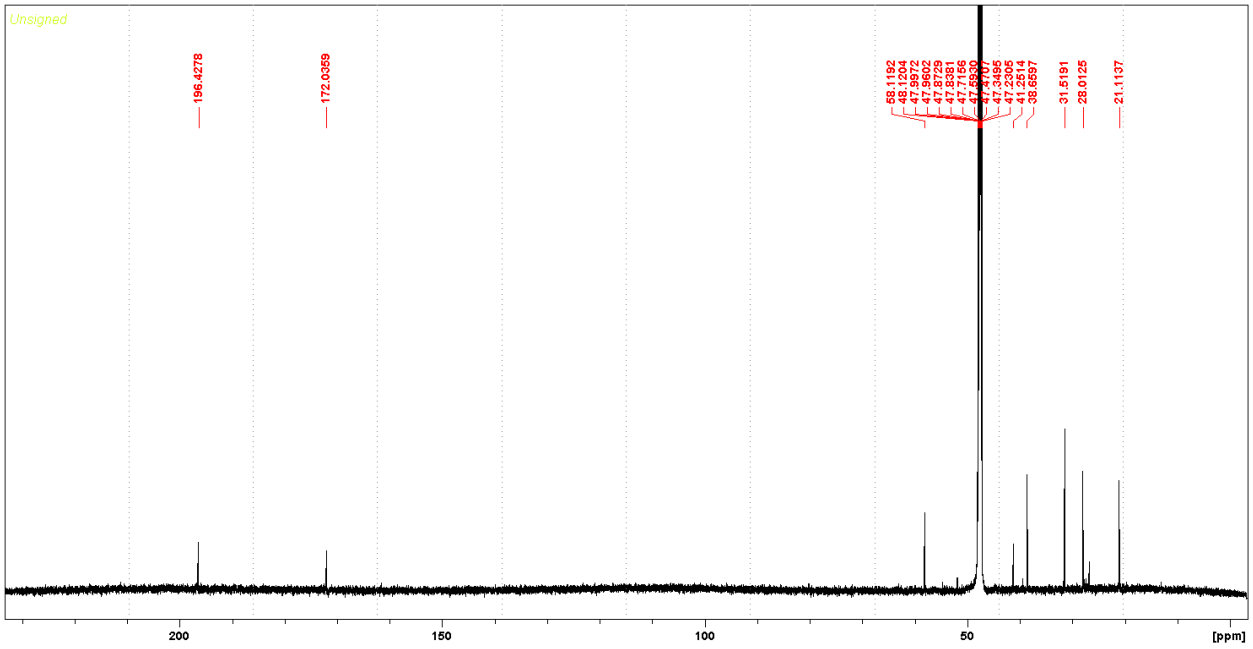


Fig. S26. The HSQC spectrum of compound **16**

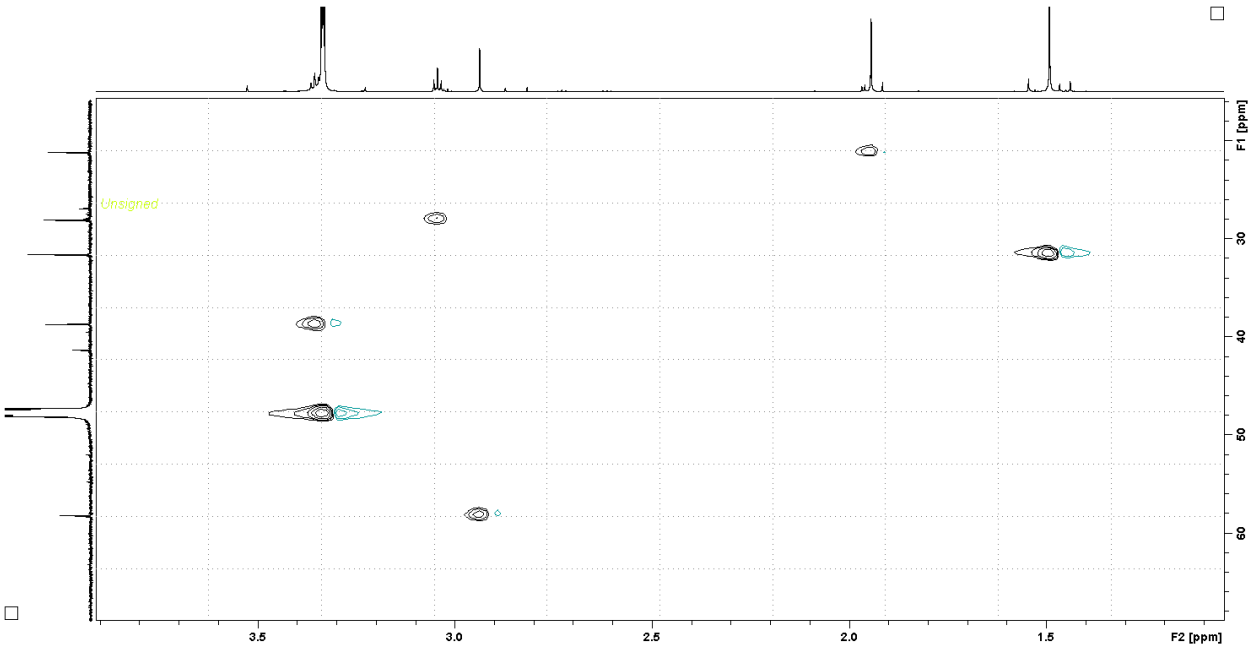


Fig. S27. The HMBC spectrum of compound **16**

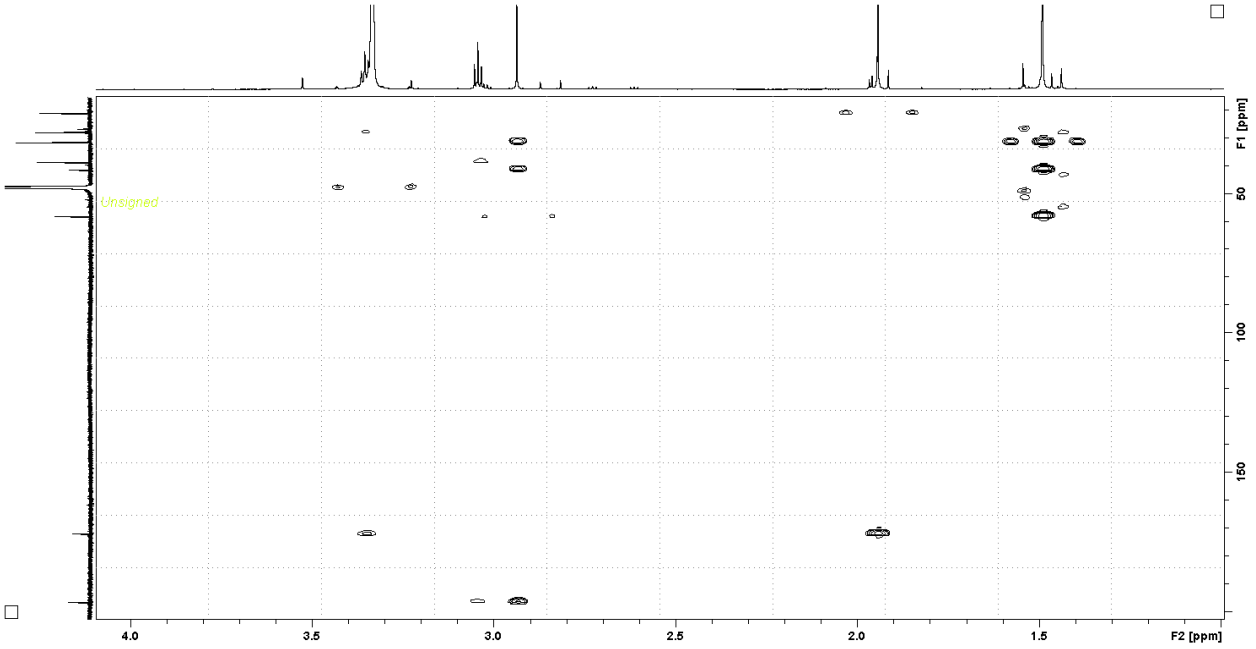


Fig. S28. The HR-ESI-MS of compound **17**

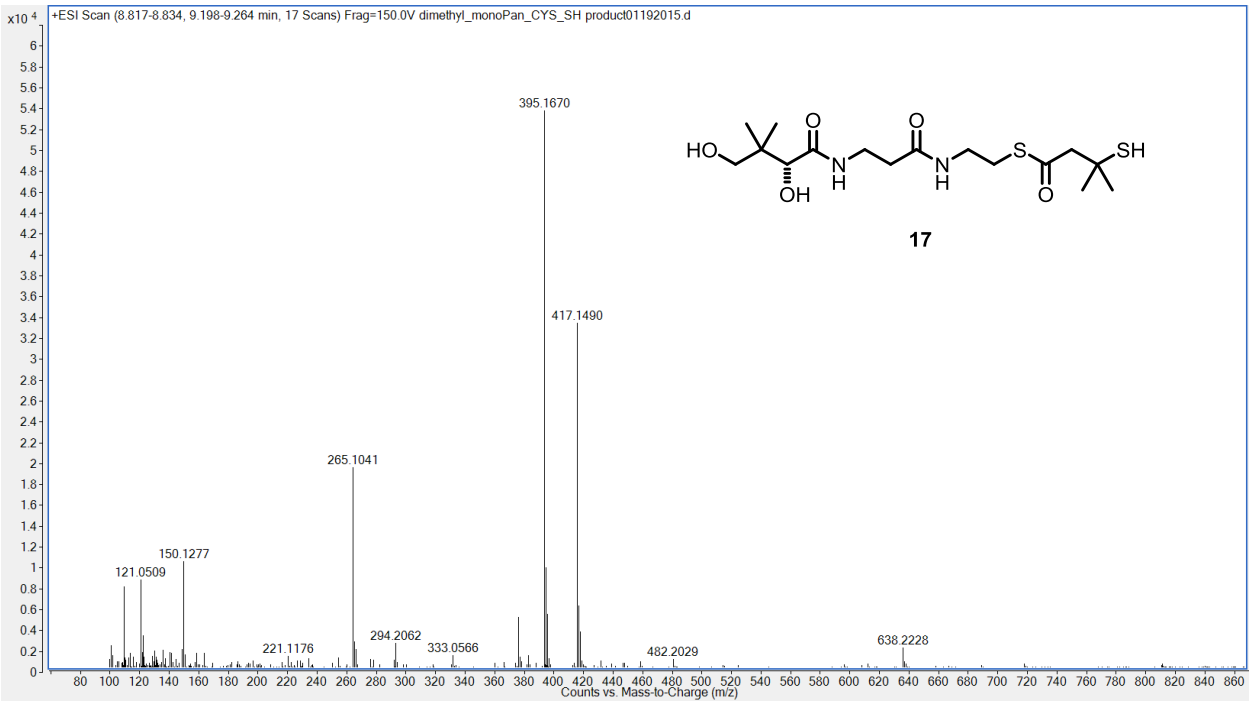


Fig. S29. Optimization of LnmJ-SH activity in vitro assay conditions. *(A)* The effect of pH, phosphate buffer (■, 6.0-8.0), Tris-HCl (●, 7.5-8.5). *(B)* The effect of KCl. *(C)* The effect of PLP.

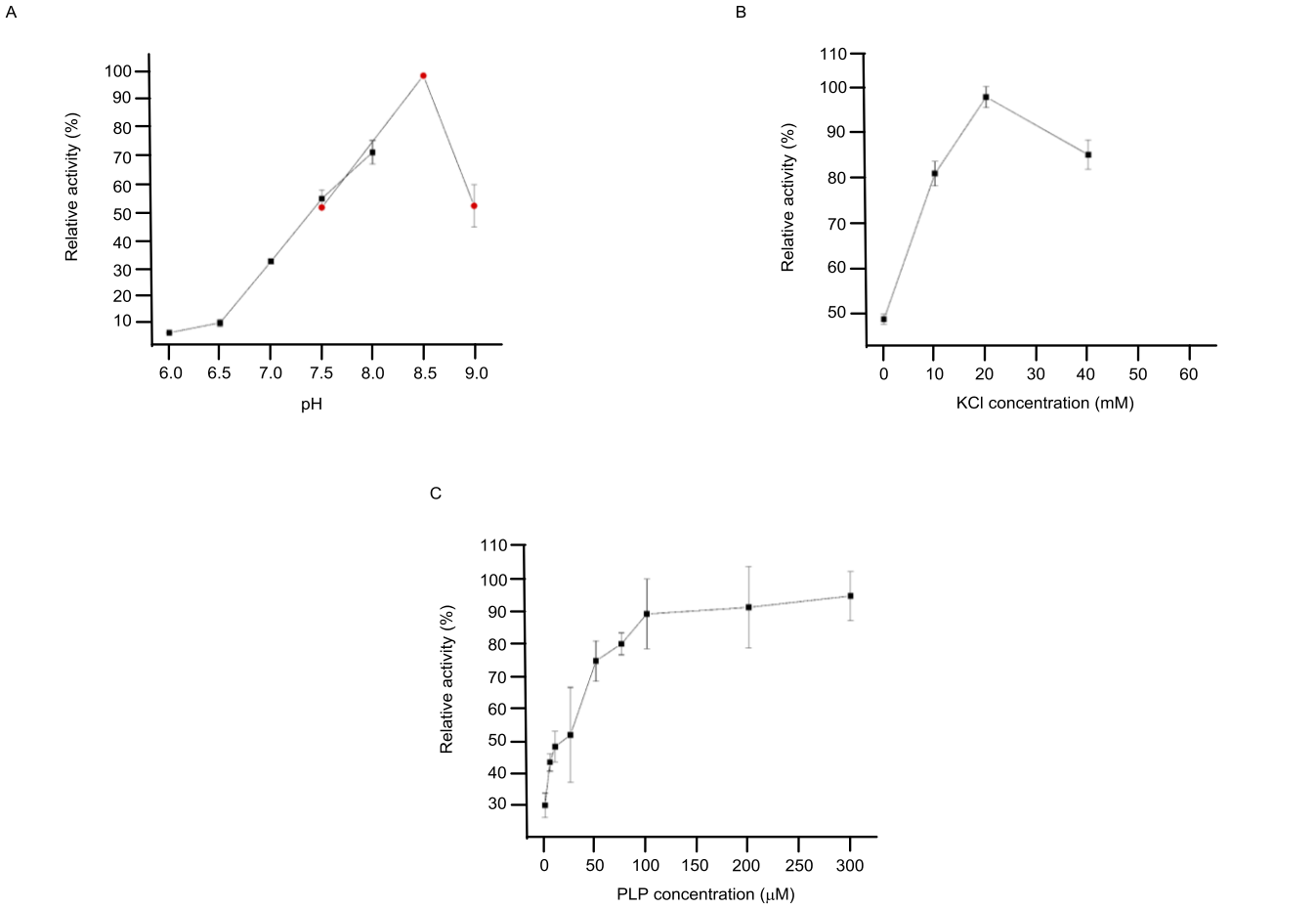


Fig. S30. Kinetic analysis of SH domains with various substrates at pH 7.2. Michaelis-Menten curves (solid lines) were generated by nonlinear regression analysis using KaleidaGraph (Synergy Software) for the reactions of LnmJ-SH (red circles), CaJ-SH (blue squares) and MaJ-SH (green diamonds) with compounds **1**, **4**, **9** and **10** (the velocity values have been divided by the enzyme concentration). Due to the limited solubility of **4**, the k_{cat}/K_m was determined by the slope of the line through the four lowest concentration data points, as represented by dashed lines.

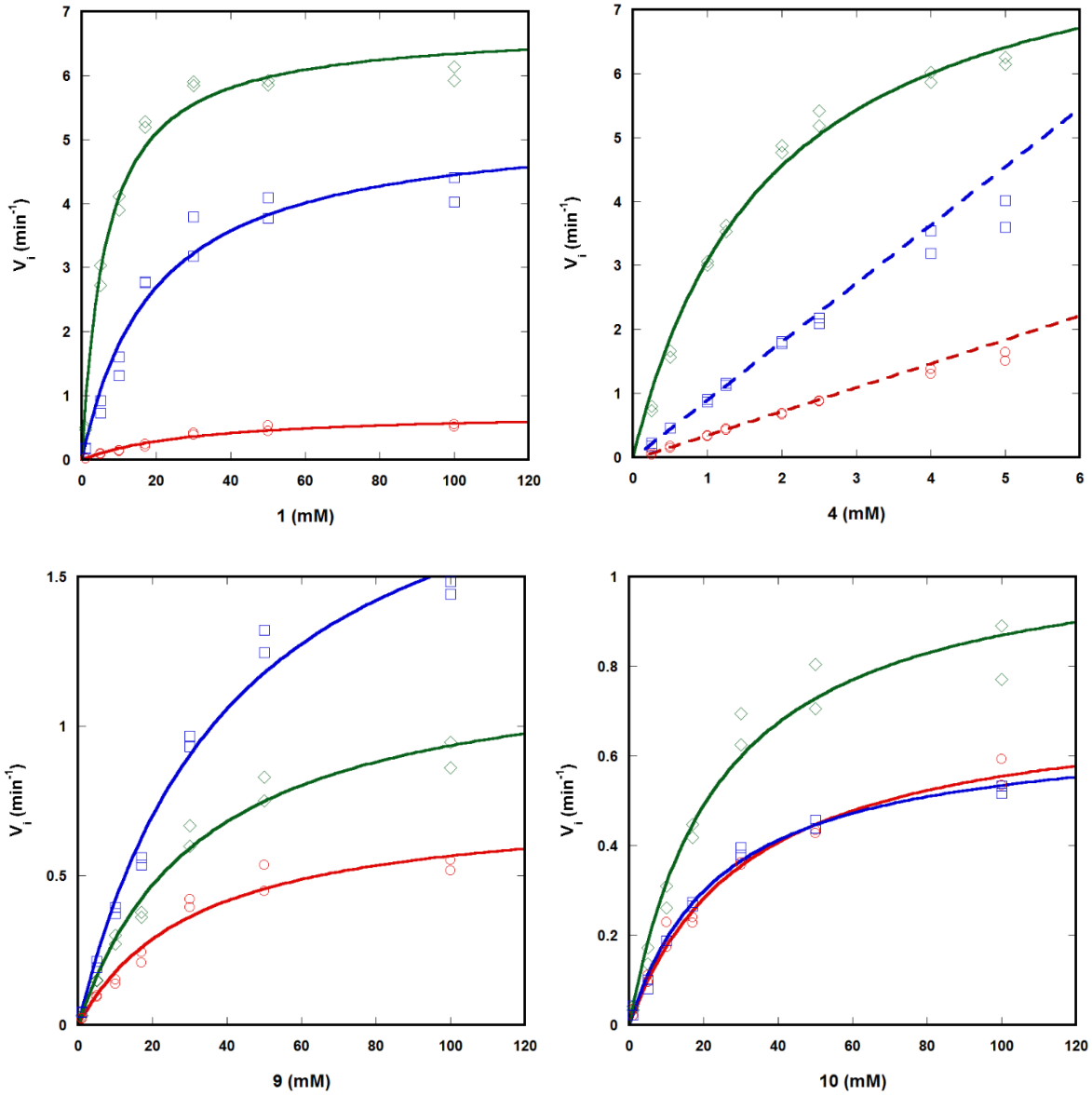


Fig. S31. Natural product biosynthesis featuring (*A*) C-S bond formation or (*B*) C-S bond cleavage steps known to date (see main text for details and references).

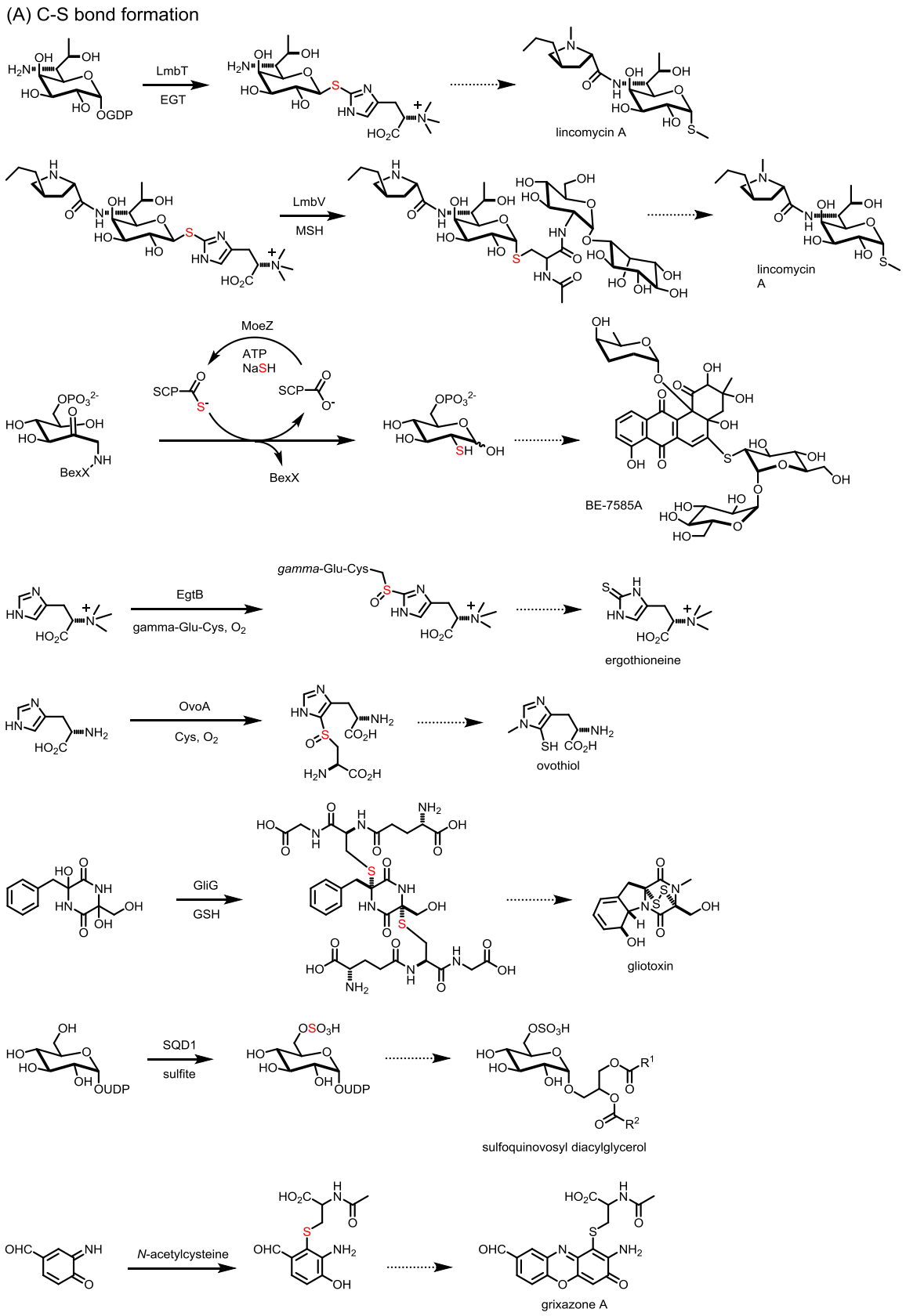


Fig. S31 continued.

(B) C-S bond cleavage

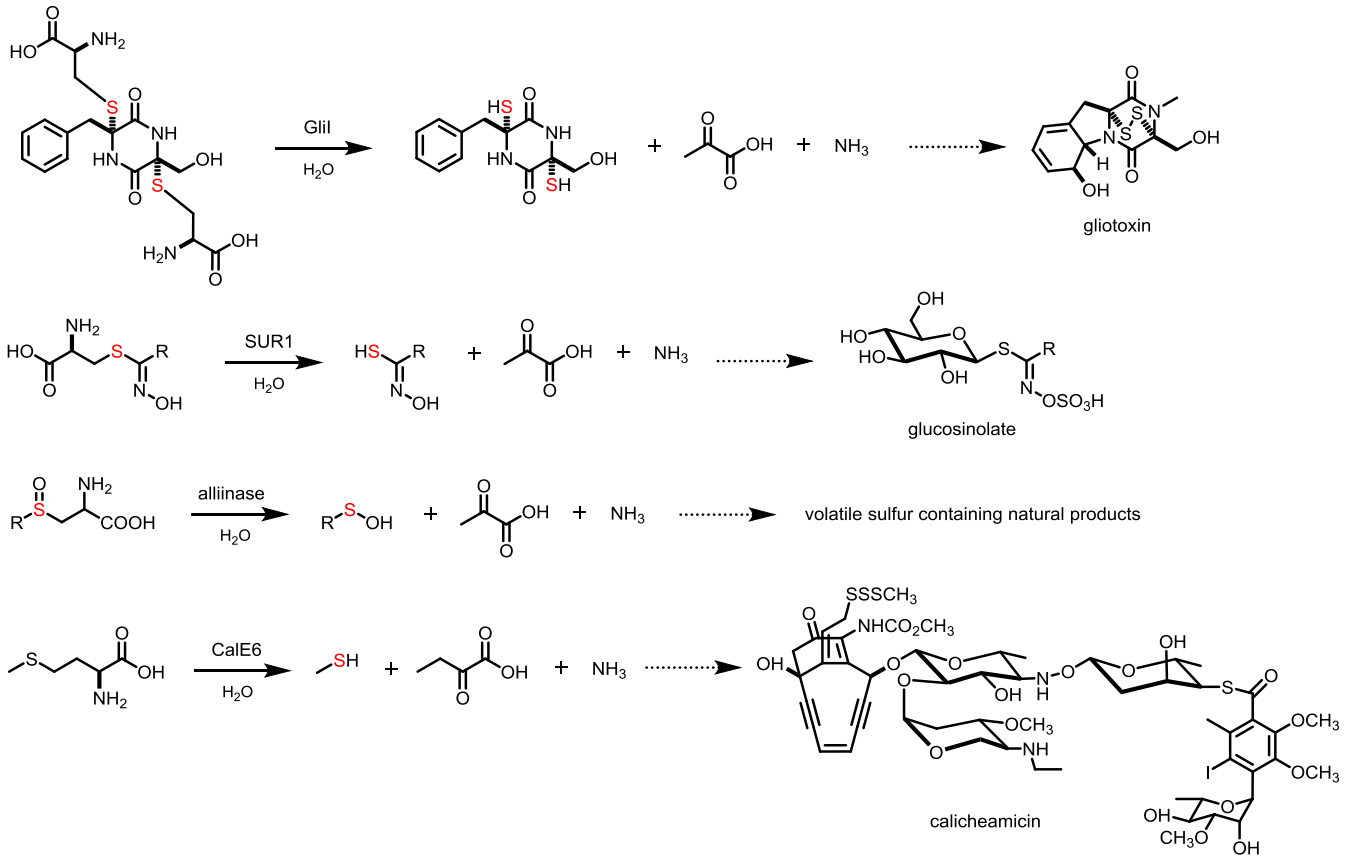


Fig. S32. A structural model of LnmJ-SH generated by SWISS MODEL (11). *(A)* LnmJ-SH shown in color from blue (N-terminal) to red (C-terminal), TPL (PDB 2EZ2) is colored black, a model of the PLP-aminoacrylate intermediate is shown as spheres. *(B)* Coloring similar to A, with catalytic dimer of TPL shown in gray and potassium cation shown in magenta. (C) model of PLP-aminoacrylate intermediate in active site, displaying common residues interacting with intermediate, TPL in black and LnmSH in white. (D) Potassium cation (purple) binding site, with same coloring scheme as in B, demonstrating interference of cation binding.