Table of Contents

General Methods

Materials
All materials were purchased from Fisher Scientific, Alfa Aesar, Sigma-Aldrich, or Chem-Impex.

Bacterial growth selection
All bacteria transformed with pET28 were grown in the presence of the corresponding antibiotic kanamycin (50 mg/L).

NMR
NMR spectroscopic data was collected on a Bruker Avance III spectrometer (600 MHz) using either a 1.7 mm inverse detection triple resonance (H-C/N/D) cryoprobe or a 5 mm inverse detection triple resonance (H-C/N/D) cryoprobe or a JEOL spectrometer (500 MHz). All samples were dissolved in D_2O and supplemented with 0.1% methanol to serve as a reference standard.

Molecular Biology/Biochemical Methods

DNA extraction
Digenea simplex was collected from Onna, Okinawa Prefecture, Japan. The sample was dried by lyophilization and finely ground. Five hundred milligrams of powder was resuspended in 10 mL of CTAB buffer (3% CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl pH 8.0, 0.2% polyvinylpolypyrrolidone, 0.2% β-mercaptoenthanol, 1.5 mg/mL RNase A, and 0.2 mg/mL proteinase K) and incubated at 55 °C for 2 hours with gentle mixing every 15 minutes. The cell debris was removed by centrifugation at 14,000 xg for 10 minutes at room temperature. The supernatant was transferred and 2 mL of 5 M potassium acetate pH 8.0 was added. The mixture was gently mixed and cooled on ice for 30 minutes then centrifuged at 14,000 xg for 15 minutes at 4 °C. The supernatant was removed and extracted with one equivalent of phenol:chloroform:isoamyl alcohol 25:24:1 saturated with 10 mM Tris, pH 8.0, 1 mM EDTA. The mixture was centrifuged at 12,000 xg for 5 minutes at 4 °C. The aqueous layer was transferred to a new tube and extracted with one equivalent of chloroform and centrifuged at before. The aqueous layer was transferred and one equivalent of cold isopropanol was added and carefully mixed. The mixture was centrifuged at 12,000 xg for 15 minutes at 4 °C. The supernatant was removed and the pellet was carefully washed with 75% ethanol. The pellet was dried to remove the ethanol and carefully resuspended in 250 μL of elution buffer (10 mM Tris-HCl pH 8.0). To further remove RNA, 300 μg of RNase A was added and incubated at room temperature for 1 hour prior to incubation at 4 °C overnight. The RNase A treated DNA was extracted with phenol:chloroform:isoamyl alcohol 25:24:1 saturated with 10 mM Tris, pH 8.0, 1 mM EDTA followed by chloroform extraction as before. Sodium acetate (pH 5.2) was added to a final concentration of 0.3 M followed by three equivalents of cold ethanol and carefully mixed. The mixture was incubated at -20 °C for 2 hours and centrifuged at 16,000 xg for 20 minutes at 4 °C. The pellet was washed with 75% ethanol and dried. The pellet was resuspended with 100 μL elution buffer and further purified by a PippinHT with > 8,000 bp size selection.

Palmaria palmata was collected from Wallace Cove Lighthouse, Bay of Fundy, New Brunswick, Canada and dried in silica gel. The sample was further dried by lyophilization and ground into a

powder. DNA was extracted in a similar manner to the previous method, except 20 mL of CTAB buffer was used for 500 mg of *P. palmata* powder and RNase A was omitted. Because RNA was not completely digested after the overnight incubation an additional 100 μg of RNase was added and further incubated as before. DNA was directly used for sequencing without a PippinHT purification step.

Genome sequencing

The high molecular weight (HMW) *D. simplex* and *P. palmata* DNA was sequenced by running on an Oxford Nanopore MinION or GridION sequencer (Oxford Nanopore Technologies, Oxford, UK). A one-dimensional (1d) library was prepared with 1.2 μg DNA using the Ligation Sequencing Kit (SQK-LSK109, Oxford Nanopore Technologies, Oxford, UK). The libraries were loaded on a R9.4 flowcell and run for 48 hr resulting in 6.7 Gb of sequencing with a read length N50 of 7.2 kb and 4.7 Gb of sequencing with a read length N50 of 8 kb for *D. simplex* and *P. palmata* respectively. Since the HMW DNA generated both long reads (good distribution) and throughput, 1 μg of DNA was used to make 1d library for the PromethION (Oxford Nanopore Technologies, Oxford, UK) for *D. simplex*. The resulting sequencing run produced 47 Gb of sequence with a read length N50 of 7 kb.

Illumina sequencing

NEBnext sequencing libraries were generated to polish the genome assemblies (New England Biolabs, Beverly, MA, USA). NEBnext sequencing libraries were created with 100 ng of DNA and quality controlled on a bioanalyzer. The resulting libraries were sequenced on an Illumina MiSeq 2x150 bp to check quality and quantity (Illumina, San Diego, CA). The libraries were then sequenced on an Illumina NovaSeq6000 2x150 bp run that resulted in 19 and 34 Gb of sequence for *D. simplex* and *P. palmata* respectively.

Genome assembly

Sequencing reads resulting from the PromethION (47 Gb) were assembled using a correction-less overlap-layout-consensus strategy was taken to assemble the genome[1]. The resulting raw reads in fastq format were aligned (overlap) with minimap[2] and an assembly graph (layout) was generated with miniasm[3]. The resulting graph was inspected using Bandage[4]. A consensus sequence was generated by mapping reads to the assembly with minimap, and then Racon[5] three times. Finally, the assembly was polished with pilon[6] three times using the Illumina paired-end 2x150 bp sequence; the Illumina reads were mapped to the consensus assembly using minimap2.

***P. palmata kab* cluster**

The 4.7 Gb of ONT sequence generated on the GridION for *P. palamta* was used to search for long reads containing the *ppkab* cluster. The fastq file was transformed to a fasta file and made into a blast database. The DsKab genes were then used to search the long-read blast database by tblastn and a 44 kb read was identified with the *ppkab* cluster. A consensus sequence was generated by mapping reads to the single 44 kb read with minimap[2], and then Racon[5] two times. Polishing with Illumina reads was attempted on the 44 kb consensus read but the *ppkab* cluster was flanked by repeat elements with 60,000 and 100,000-fold coverage, which resulted in poor polishing. Therefore, only the *kab* cluster was polished with pilon[6] five times using the Illumina paired-end 2x150 bp sequence.

PCR and cloning

PCR of the Kab genes was performed with PrimeSTAR HS DNA Polymerase (Takara) according to the manufacturer's instructions using either genomic DNA or codon optimized synthetic genes as a template. pET28-MBP (Maltose Binding Protein)-TEV and pET28 vectors were amplified using PrimeSTAR MAX DNA Polymerase (Takara). PCR products were purified by agarose gel extraction and assembled into expression vectors using NEBuilder HiFi DNA Assembly mix and transformed into *E. coli* DH5α cells. Construction of vectors was confirmed by Sanger sequencing. The pET28-MBP-TEV vector produced proteins with a His_6-MBP-His_6 N-terminal tag and the pET28 vector produced proteins with a His_6 N-terminal tag. See Table S1 and S2 for a list of primers and synthetic genes respectively.

Phylogenetic tree

Ribulose-1,5-bisphosphate carboxylase large subunit (RbcL) DNA sequences were selected from every red alga for which RNA-seq or genomic information is available. Nucleotide sequences were aligned with the online MAFFT v7.0 server[7] using the Auto strategy and were converted to phylip format with Mesquite v3.51[8]. A maximum likelihood (ML) tree was generated using raxmlGUI v1.5b2[9] employing the GTRGAMMA strategy. Five hundred thorough bootstrap calculations followed by ten iterations of ML tree search to identify the best ML tree were utilized. The tree was visualized with FigTree v1.4.3.

***Digenea simplex* kainic acid extraction**

Twenty milligrams of *D. simplex* were finely ground and soaked in 300 mL of distilled water for 2 h at 55 °C. The sample was filtered and the flow through was analyzed by an Agilent Technologies 1200 Series system with a diode-array detector coupled to an Agilent Technologies 6530 accurate-mass Q-TOF LCMS using the following method on a Synergi Polar-RP 4μ. 250 x 4.6 mm column at 0.75 mL/min: 0% B (4.5 min), 0 to 5% B (0.5 min), 5 to 26% B (9 min), 26 to 80% B (9 min), 80 to 100% B (1 min), 100% B (1.5 min), 100 to 0% B (2.5 min), 0% B (2 min), wherein A = 0.1% aqueous formic acid, and B = 0.1% formic acid in acetonitrile. HRMS (ESI) calculated for prekainic acid ($C_{10}H_{16}NO_4^-$) 214.1085, found 214.1078 $(M-H)^-$; kainic acid ($C_{10}H_{14}NO_4^-$) 212.0928, found 212.0936 $(M-H)^-$; and kainic acid lactone ($C_{10}H_{14}NO_4^-$) 212.0928, found 212.0922 $(M-H)^-$.

Protein purification

E. coli BL-21 co-transformed with the protein expression plasmid were shaken 37 °C in 1 L of terrific broth to an OD_{600} ~0.8 and then cooled at 18 °C for 1 h before addition of 0.5 mM IPTG for pET28 vectors or 0.2 mM IPTG for MBP vectors. The flasks were shaken overnight and harvested by centrifugation the following morning to remove the media. The cells were resuspended in 25 mL of suspension buffer (500 mM NaCl, 20 mM Tris pH 8.0, and 10% glycerol) and lysed by sonication with a Qsonica 6 mm tip at 40% amplitude for 12 cycles of 15 seconds on and 45 seconds off. The cell lysate was centrifuged at 14,000 xg for 25 minutes to removed insoluble debris. The cleared lysate was loaded at 2 mL/min onto a 5 mL HisTrap FF column (GE Healthcare Life Sciences) that was pre-equilibrated with buffer A (1 M NaCl, 30 mM imidazole, and 20 mM Tris pH 8.0) using an ÄKTA FPLC (GE Healthcare Life Sciences). After loading, the column was washed with 40 mL of buffer A. Protein was eluted using a linear gradient of 0 – 100% buffer B (1 M NaCl, 250 mM imidazole, and 20 mM Tris pH 8.0) over 40 mL and 5 mL fractions were collected. The fractions were analyzed by SDS-PAGE and fractions with at least 90% purity

were combined. For the KabC proteins, 2 mM EDTA was added to remove any metals bound in the active site. If the His_6 tag was to be removed, 60 units of thrombin were added and incubated overnight at 4 °C. Complete cutting was confirmed by SDS-PAGE. Protein was concentrated using Amicon Ultra Centrifugal Filters and further purified by gel filtration using a either a HiLoad 16/60 Superdex 75 or 200 prep grade gel filtration column (GE Healthcare Life Sciences) pre-equilibrated with 300 mM KCl and 50 mM HEPES pH 8.0. If the protein was to be frozen following gel filtration, 10% glycerol was also added.

Enzymatic activity assays

To test DsKabA activity and substrate specificity, a 50 μL reaction containing 5 mM $MgCl_2$, 5 mM amino acid (L-Glu, D-Glu, L-Gln, L-Asp, L-Asn, or Gly), 1 mM prenyl donor (DMAPP or GPP), and 20 μM DsKabA in a buffer of 300 mM KCl and 50 mM HEPES pH 8.0 was incubated at 23 °C for 5 h. One equivalent of 0.1% formic acid in H_2O was added and filtered with a VWR PES 3k 500 μL spin column. The flow through was analyzed by an Agilent Technologies 1200 Series system with a diode-array detector coupled to an Agilent Technologies 6530 accurate-mass Q-TOF LCMS using the following method on a Synergi Polar-RP 4μ 250 x 4.6 mm column at 0.75 mL/min: 0% B (4.5 min), 0 to 5% B (0.5 min), 5 to 26% B (9 min), 26 to 80% B (9 min), 80 to 100% B (1 min), 100% B (1.5 min), 100 to 0% B (2.5 min), 0% B (2 min), wherein A = 0.1% aqueous formic acid, and B = 0.1% formic acid in acetonitrile. For KabC activity and substrate specificity assays, 50 μL reactions containing 5 mM αKG, 1 mM L-ascorbic acid, 1 mM substrate (prekainic acid, NGG, or cNGG) 30 μM ferrous sulfate, and 20 μM KabC in a buffer of 300 mM KCl and 50 mM HEPES pH 8.0 were incubated at 23 °C for 5 h. One equivalent of 0.1% formic acid in H_2O was added and filtered as before. The flow through was analyzed using the same methodology as DsKabA.

DsKabA prekainic acid reaction scaleup and purification

A 5 mL reaction of 30 μM His_6-DsKabA, 5 mM L-Glu, 10 mM $MgCl_2$, and 5 mM DMAPP in a buffer of 300 mM KCl and 50 mM HEPES pH 8.0 was incubated at 22 °C for 18 h. The protein was separated from the reaction using a 30 kDa Amicon Ultra Centrifugal Filter and washed to recover all small molecule products. The collected flow through was lyophilized and resuspended in 0.1% aqueous formic acid before being purified by preparative RP-HPLC (Agilent PrepStar and ProStar 410 HPLC) using a Phenomenex Luna 5μ C18(2), 100 x 21.2 mm column at a flow rate of 10 mL/min with the following method: 2% B (5 min), 2 to 20% B (10 min), 20 to 95% B (1 min), 95% B (4 min), 95 to 2% B (2 min), 2% B (3 min), wherein A = 0.1% aqueous formic acid, and B = 0.1% formic acid in acetonitrile. Absorbance at 210 nm was monitored and the major peak was collected. Acetonitrile was removed by rotary evaporation before lyophilization to afford prekainic acid (3.7 mg, 56% yield) as a white powder. ^{1}H NMR (599 MHz, D_2O) δ 5.22 (t, *J* = 7.3 Hz, 1H), 3.64 (t, *J* = 6.7 Hz, 2H), 3.61 (t, *J* = 6.4 Hz, 1H), 2.55 – 2.42 (m, 2H), 2.12 (ddt, *J* = 13.7, 7.3, 7.3 Hz, 1H), 2.03 (dtd, *J* = 14.3, 7.3, 6.9 Hz, 1H), 1.76 (s, 3H), 1.68 (s, 3H). ^{13}C NMR (151 MHz, D_2O): δ177.6, 173.3, 144.8, 113.0 60.2, 44.5, 30.5, 25.4, 25.4, 17.7; HRMS (ESI) calculated for $C_{10}H_{16}NO_4^-$ 214.1085, found 214.1088 $(M-H)^-$.

DsKabC kainic acid reaction scaleup and purification

A 7.5 mL reaction of 30 μM His_6-DsKabC, 5 mM prekainic acid, 10 mM αKG, 2 mM L-ascorbate, and 50 μM ferrous sulfate in a buffer of 300 mM KCl and 50 mM HEPES pH 8.0 was incubated

at 22 °C for 18 h. The protein was separated from the reaction using a 30 kDa Amicon Ultra Centrifugal Filter and washed to recover all small molecule products. The collected flow through was lyophilized and resuspended in 0.1% aqueous formic acid before being purified by preparative RP-HPLC (Agilent PrepStar and ProStar 410 HPLC) using a Phenomenex Luna 5μ C18(2), 100 x 21.2 mm column at a flow rate of 10 mL/min with the following method: 2% B (5 min), 2 to 20% B (10 min), 20 to 95% B (1 min), 95% B (4 min), 95 to 2% B (2 min), 2% B (3 min), wherein A = 0.1% aqueous formic acid, and B = 0.1% formic acid in acetonitrile. Absorbance at 210 nm was monitored and the major peak was collected. Acetonitrile was removed by rotary evaporation before lyophilization to afford kainic acid (4.5 mg, 46% isolated yield) as a white powder. ^{1}H NMR (599 MHz, D_2O) δ 5.01 (s, 1H), 4.73 (s, 1H), 4.07 (d, *J* = 3.0 Hz, 1H), 3.60 (dd, *J* = 11.9, 7.3 Hz, 1H), 3.40 (dd, *J* = 11.5, 11.5 Hz, 1H), 3.09 – 3.02 (m, 1H), 3.02 – 2.95 (m, 1H), 2.43 (dd, *J* = 16.6, 6.4 Hz, 1H), 2.34 (dd, *J* = 16.6, 8.4 Hz, 1H), 1.73 (s, 3H). ^{13}C NMR (151 MHz, D_2O): δ 176.9, 173.6, 140.3, 113.6, 66.5, 46.8, 46.2, 40.8, 33.5, 22.1; HRMS (ESI) calculated for $C_{10}H_{14}NO_4^-$ 212.0928, found 212.0929 $(M-H)^-$.

GfKabC kainic acid lactone reaction scaleup and purification

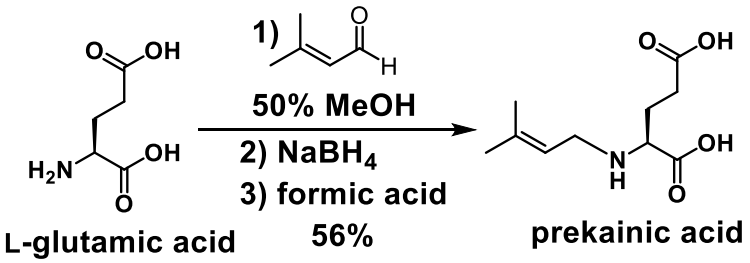
MBP-GfKabC was buffer exchanged into 300 mM KCl, 50 mM HEPES pH 8.0 to remove glycerol. Six 1 mL reactions of GfKabC were set up as follows: 8 mM αKG, 5 mM prekainic acid, 1 mM L-ascorbic acid, 50 μM ferrous sulfate, 50 μM MBP-GfKabC in buffer containing 300 mM KCl, 50 mM HEPES pH 8.0. The reaction was incubated at 22 °C for 30 hours and the protein was removed with 30 kDa Amicon Ultra Centrifugal Filter. The protein was washed with distilled water and the flow throughs were collected and lyophilized. The dried material was resuspended in 500 μL of 0.1% aqueous formic acid and purified by semipreparative HPLC (Agilent 1200 Infinity HPLC) using a Synergi Polar-RP 4μ 250 x 10.0 mm column with the following method: 0% B (18 min, 2 mL/min), 0 to 95% B (2 min, 3 mL/min), 95% B (3 min, 3 mL/min), 95 to 0% B (4 min, 3 mL/min), 0% B (3 min, 3 mL/min), wherein A = 0.1% aqueous formic acid, and B = 0.1% formic acid in acetonitrile. The major peak (~9.5 min retention time) was collected and lyophilized to afford kainic acid lactone (3.3 mg, 26% isolated yield) as a white powder. ^{1}H NMR (600 MHz, D_2O) δ 3.97 (d, *J* = 3.9 Hz, 1H), 3.73 (dd, *J* = 12.2, 8.2 Hz, 1H), 3.24 (dd, *J* = 11.9, 11.9 Hz, 1H), 3.22 – 3.16 (m, 1H), 3.15 (dd, *J* = 18.0, 9.8 Hz, 1H), 2.79 (ddd, *J* = 11.5, 8.1, 8.1 Hz, 1H), 2.55 (dd, *J* = 17.5, 3.9 Hz, 1H), 1.48 (s, 3H), 1.39 (s, 3H). ^{13}C NMR (151 MHz, D_2O): δ174.3, 172.9, 82.8, 68.3, 45.3, 42.8, 36.5,30.9, 27.9, 26.6; HRMS (ESI) calculated for $C_{10}H_{14}NO_4^-$ 212.0928, found 212.0930 $(M-H)^-$.

Kainic acid biotransformation and purification

E. coli BL-21 transformed with the pET28 DsKabC vector were grown overnight in 20 mL of lysogeny broth. The culture was pelleted to remove the media and added to 1 L of M9 minimal media supplemented with 100 μM $CaCl_2$ in a 2.8 L baffled flask. The flasks were placed in a shaking incubator (200 rpm, 37 °C) and the cells were grown to an OD_{600} of ~1.3. The flasks were cooled at 18 °C for 45 minutes followed by the addition of 0.5 mM IPTG. Fifteen minutes after IPTG was added, 4 g of glucose and the crude prekainic acid synthetic reaction product (pg. S10) was added with continuous shaking incubation at 18 °C. Additional glucose (4 g) was added after 16 h. Time points were taken and analyzed for the presence of kainic acid using the LCMS methodology described in the kainic acid fermentation protocol. After 40 h, the cultures were centrifuged to remove cells and the media was collected. Thirty grams of activated carbon (Darco G-60) per 1 L of media was washed with distilled water and added to the cell free media. The

activated carbon was stirred with the media for 30 min before filtering to remove the spent media. The activated carbon was washed with distilled water twice and then rapidly stirred in 300 mL of 80% acetic acid for 15 min to elute the kainic acid. Three additional 300 mL 80% acetic acid elutions were used to maximize kainic acid recovery. The eluent was dried by rotary evaporation to afford a reddish oil which was further purified by preparative RP-HPLC (Agilent PrepStar and ProStar 410 HPLC) using a Phenomenex Luna 5μ C18(2), 100 x 21.2 mm column at a flow rate of 10 mL/min with the following method: 2% B (5 min), 2 to 20% B (10 min), 20 to 95% B (1 min), 95% B (4 min), 95 to 2% B (2 min), 2% B (3 min), wherein A = 0.1% aqueous formic acid, and B = 0.1% formic acid in acetonitrile. Absorbance at 210 nm was monitored and the major peak was collected. The kainic acid containing fractions were combined and the volume was reduced in vacuo before lyophilization. Using this method, 6 L of culture were used to afford 1.07 g of kainic acid (32% overall isolated yield) as a white powder. NMR (Figure S12) indicates >95% purity with the major contaminants consistent with acetic acid (~2%) and formic acid (~2%).

Synthetic Methods



Prekainic acid chemical synthesis

This molecule was synthesized using a modified published procedure[10]. An aqueous solution of L-glutamic acid (0.147 g, 1 mmol) and crushed NaOH (0.083 g, 2.1 mmol) in 2.5 mL of H_2O was stirred at RT. MeOH (2.5 mL) was added. 3-methyl-2-butenal (0.135 mL, 1.4 mmol) was added and the reaction was immediately cooled in an ice bath while stirring. $NaBH_4$ (0.048 g, 1.3 mmol) was added and stirred for 20 minutes. The reaction was brought to room temperature over 20 minutes and was then acidified to ~pH 4.0 with formic acid. The methanol was removed in vacuo and the reaction was purified by injection onto preparative RP-HPLC (Agilent 1260 Infinity HPLC) using a Phenomenex Luna 5μ C18(2), 100 x 21.2 mm column at a flow rate of 10 mL/min with the following method: 2% B (5 min), 2 to 20% B (10 min), 20 to 95% B (1 min), 95% B (4 min), 95 to 2% B (2 min), 2% B (3 min), wherein A = 0.1% aqueous formic acid, and B = 0.1% formic acid in acetonitrile. Absorbance at 210 nm was monitored and the major peak was collected. Acetonitrile was removed by rotary evaporation before lyophilization to afford prekainic acid (147 mg, 56% isolated yield) as a white powder. ^{1}H NMR (599 MHz, D_2O) δ 5.22 (t, *J* = 7.8 Hz, 1H), 3.64 (t, *J* = 7.9, 6.4 Hz, 2H), 3.61 (dd, *J* = 7.7, 5.5 Hz, 1H), 2.56 – 2.42 (m, 2H), 2.13 (ddt, *J* = 13.8, 7.0, 7.0 Hz, 1H), 2.03 (dtd, *J* = 14.4, 7.3, 6.7 Hz, 1H), 1.76 (s, 3H), 1.68 (s, 3H). ^{13}C NMR (151 MHz, D_2O): δ 177.5, 173.8, 145.3, 113.2, 60.4, 44.5, 30.1, 25.5, 25.4, 17.9; HRMS (ESI) calculated for $C_{10}H_{16}NO_4^-$ 214.1085, found 214.1088 $(M-H)^-$.

Prekainic acid chemical synthesis for biotransformation

An aqueous solution of L-glutamic acid (1.18 g, 8.0 mmol) and crushed NaOH (0.672 g, 16.8 mmol) in 20 mL of H_2O was stirred at room temperature. MeOH (20 mL) was slowly added. 3-methyl-2-butenal (1.08 mL, 11.2 mmol) was added and the reaction was immediately cooled in an ice bath while stirring. $NaBH_4$ (0.393 g, 10.4 mmol) was added and stirred for 20 minutes. The reaction was brought to room temperature over 20 minutes and was then acidified with formic acid to ~pH 4.0. The methanol was removed in vacuo. The crude reaction product was neutralized with NaOH to ~pH 8.0 and 1/3 of the reaction was added to a 1 L biotransformation (pg. S8).

DMAPP chemical synthesis

This synthesis was adapted from a literature procedure[11]. To a solution of tris (tetrabutylammonium) hydrogen pyrophosphate trihydrate[11] (3.00 g, 3.32 mmol) in dry acetonitrile (15 mL) at -35 °C was added prenyl bromide (0.19 mL, 0.25 g, 1.66 mmol) over 2

minutes. The reaction mixture was stirred at -35 °C for 10 minutes, then warmed to room temperature and stirred for an additional 2 hours. The solvent was removed in vacuo and the crude residue was resuspended in minimal ion exchange buffer (25 mM aqueous NH_4HCO_3, 2% isopropanol) and passed through a column of DOWEX AG50W-X8 (NH_4^+ form, 40 mL (~70 milli-equivalents)). The first two column volumes (80 mL) were collected and lyophilized to generate an off-white solid. This crude mixture was further purified via resuspension in 50 mM aqueous NH_4HCO_3 (2 mL), addition of 1:1 acetonitrile:isopropanol (8 mL) with thorough mixing, centrifugation (2000 x *g*, 5 minutes) and pipetting the resultant supernatant into a new flask, repeating this procedure 5 times in total. The pooled supernatants were concentrated in vacuo then further purified by cellulose chromatography prepared in an analogous way to that previously described[11]. Purified DMAPP was eluted using a gradient of (1:2:1) to (2:2:1) water:acetonitrile:isopropanol (50 mM NH_4HCO_3). Pooled fractions were concentrated in vacuo and lyophilized, generating trisammonium DMAPP (0.452 g, 91%) as a white powder. ^{1}H NMR (500 MHz, D_2O) δ 5.31 (t, *J* = 7.2 Hz, 1H), 4.31 (dd, *J* = 7.0, 6.8 Hz, 2H), 1.62 (s, 3H), 1.58 (s, 3H); LRMS (ESI) calculated for $C_5H_{11}O_7P_2^-$ 245.00, found 245.02 $(M-H)^-$.

DsKabA p28 For	GGTGCCGCGCGGCAGCCAT ATG AGT ATT ACC ACT ATG AAA GGA GTT ACA TC
DsKabA p28 Rev	GGTGGTGGTGGTGCTCGAG TTA ATT TGT CGC CTT GAG TTT CTT AAA G
DsKabC p28 For	GGTGCCGCGCGGCAGCCAT ATG ACA GTA AGT AAG TAC AAC GGT GTT G
DsKabC p28 Rev	GGTGGTGGTGGTGCTCGAG TTAAAGGTAGTATCCTTCATGAAATTTG
GfKabC MBP For	CTGTACTTCCAATCCGGATCC ATG ACT TCT TTC GAT TTC GTC GCT AAG ACT TTC
GfKabC MBP Rev	GGTGGTGGTGGTGCTCGAG TTA GAC GTA GTA CCC CGA GTG AAA CTT TGC
PpKabC MBP For	CTGTACTTCCAATCCGGATCC ATG CCT GTG TAT AAT GGT ACG TCA CAA GC
PpKabC MBP Rev	GGTGGTGGTGGTGCTCGAG TTA CAG GTA GTA ACC TTC ATG GAA TTT GTT CTT G
ReKabC MBP For	CTGTACTTCCAATCCGGATCC ATG GCG GTG TAC AAT GGC GTC TCG CAG G
ReKabC MBP Rev	GGTGGTGGTGGTGCTCGAG CTA TAG ATA ATA GCC CTC GTG GAA CTT ATT TTT CAC G
pET28 Up	CATATGGCTGCCGCGCGGCACC
pET28 Down	CT CGA GCA CCA CCA CCA CCA CCA CTG AG

Table S1: Primers used in this study

PpKabC	ATGCCTGTGTATAATGGTACGTCACAAGCATTCAATTTTGCGCCGCTGCACC CAGATAGTTTGAATTGGTGCCCAAAGTCGAGTCTTCCTCCAGAAATTCCTGT AGTGGATCTGTCAAAACTTAATTCTGAAGCCGAGCTGGCACAATTTCTTGTG GACATTCGTAAGAGCGGGTTGTTTTATGTAGTCAATCATGGTGTCCCGGAAG AACTTTCTATTGGGTGTTACAATCGCTTCCGTCAATTCTGTAACATTCCTGAG GACCAACGTGCGGAATACTCGACCGACCATCATTTTGTAAACGGGGGCTAC ATGCCGTATAAGAGTTCCAGCATCGGCAAGGCAAATAAAGGTAAATCTCAAA AGGACTTTGTGGTTAAATACTTCTGGCGTGGACCGCGCGTTGAGAACCGTA GTCCGGATGCTGATTTCAAGAGCTATCACGACGAATACCACCGTCGCACAG CTGATATCGCTAACAGCGTGATTACTAAAATTCTTCAGGCCCTGACGACGCG TTTTCCTGACTTTGACCCTGCTGAATATAAAGATAACATTAATTCCCATCATAT GTTCTTCTCGAACCGCCTTTACCCAGACAACGAACCTGAGAAAGGCGAGAAT GTAGAATACCGTTTAGTACCGCACCGTGACCTGAGCTTCGTGACCCTTGCTC ACCAAATCCCGGCGGACAACGGTTACCAAGGCCTTTTCGTTTTTACCGGCGA TGGGAAGAAGGTTTGCGTGCCTCCGATTCGTAACAGCTATCTTGTCTTCATG GGACAGGCAATGTCGTATCTGACGAATAAGTACTTACCTGCAGGGTTGCATG GCGTGGATTTTCCTGAAAAGAACTGTTTCGAGGGGTCTGAACGCTCCTCTCT TATCTCATTTTACGAACCTCACGATCATATGATGCCCAGCAAAGCACTTACAC CTAAGGATGATGAGATCTTCGATCGTGGTTGTTCGTTCTTTGATGACATCGG CGTGGACAAAACAGGGACGACCTATATGTATGTCAAGAACAAATTCCATGAA GGTTACTACCTGTAA
GfKabC	ATGACTTCTTTCGATTTCGTCGCTAAGACTTTCGATTTTGCTCCCGTCAAAGC TGACGAGTTGAAGTGGTATCCACGTTCCGCGTTACCACCAGAAATTCCTGTA ATTGATTTGAACAACGTCAATACGGAGGAGGAGCTTGCACAGTTCCTGGTG GACATCCGCAAGTCTGGACTTTTCTACGTTGTCAACCACGATATCCCGGAGG AAATCTCAATCCAGACGTACAACGAATTTCGTGAGTTTTGTAAATTGCCAGAA GAGGCCCGCCAAAAATACAATACTGATCGCTGCTTCTTCAATGGGGGATATG TTCCATTCAAAGCCACGAGCATTAATGGGGGTAATAAGAACAAACAGCAACG CGACTTTGTCGTGAAGTTTTTCTGGCGCGGCCCACACGTGGTCAATCGTTCC CCCAATGAACGCTTTGCCAAATGGCACCATCAGTATCATACACGCACTGCTG AATTAGGTGAAAAGGTCATGACCACAATCGCAAAAGCTTTGAAGACGCGCTT TCCTGACTTCGACCCGGCCGAACTGGAGGACAATGTGAACCCTCACAACAT GATTTTTTCTAACCGCATTTATCCGGAATTCCCACCCAGTGAAGGGGAGGAT GCTGAGTATCGCCTGACCCCCCACCGTGATATTTCCTACATTACCCTTGTGA ATCAGACTCCTGCCAATAACGGATTCAAAGCATTATTTATCGTAACCGGGGA CGGCGAACGCGTTTACGTCCCTCCTATCCGCAACTCTTACTTGGTATTCATC GGGCAGAGCTTGAGCTACCTTACCAATAAATACCTTCCAGCGGCACTTCACG GGGTCGCTTTTCCGGACGGCAATCTTGAGGGTTGCGAGCGTGCCTCACTGG TGTCTTTCTATGAACCGTATGATCGCATGGTCCCATCGAAAAATATCAAACCA ACGGCAGAAGAGATCGCTCCAAAATCCTGCTCCTTCTATGATTCGATTGGGT GTGACAAGACAGGAACTACCTTTACTGATGTTCGTGCAAAGTTTCACTCGGG GTACTACGTCTAA

Table S2: Synthesized genes used in this study were purchased as codon optimized gBlocks from IDT.

	Organism	NCBI Accession Number
Digenea simplex bioproject		PRJNA509898
Digenea simplex hologenome		RXNZ00000000
Digenea simplex SRA		SRR8325638, SRR8325637, SRR8325636
Digenea simplex kainic acid cluster		MK312630
Palmaria palmata bioproject		PRJNA509900
Palmaria palmata SRA		SRR8325558, SRR8325559
Palmaria palmata kainic acid cluster		MK312631
PpKabA from published transcriptomics[12]	*Palmaria palmata*	QCC62375
PpKabC from published transcriptomics[12]	*Palmaria palmata*	QCC62376
DsKabA	*Digenea simplex*	QCC62379
DsKabC	*Digenea simplex*	QCC62383
PpKabA	*Palmaria palmata*	QCC62384
PpKabC	*Palmaria palmata*	QCC62385
ReKabA	*Rhodophysema elegans*	QCC62371
ReKabC	*Rhodophysema elegans*	QCC62372
PhKabA	*Palmaria hecatensis*	QCC62373
PhKabC	*Palmaria hecatensis*	QCC62374
GfKabA (partial sequence)[13]	*Grateloupia filicina*	QCC62377
GfKabC[13]	*Grateloupia filicina*	QCC62378

Table S3: NCBI accession codes for deposited sequencing data and previously published RNASeq data.

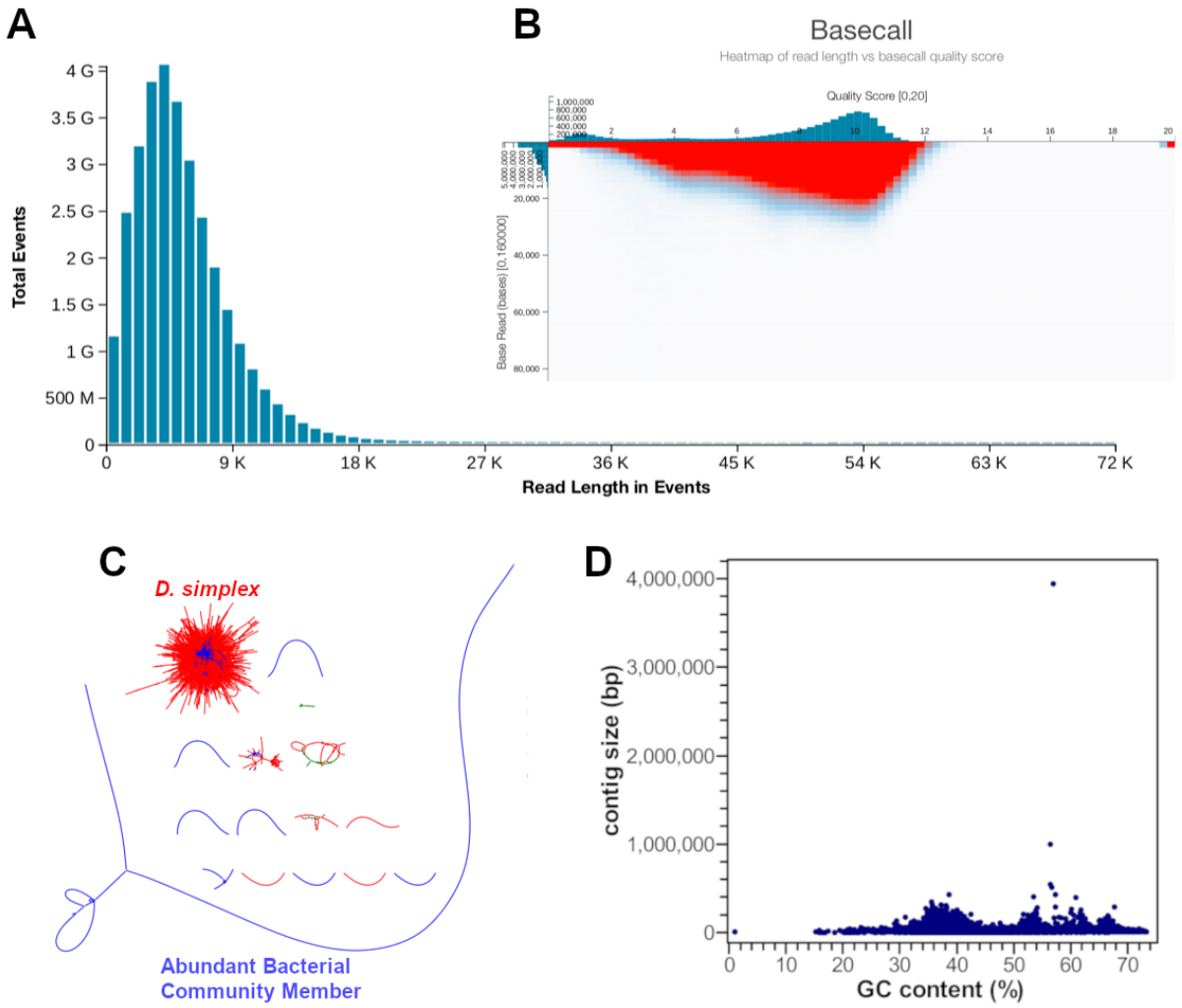


Figure S1: (A) High Molecular Weight (HMW) DNA was sequenced on the PromethION platform producing 47 Gb of long read sequencing data with an N50 read length of 7 kb. (B) Nanopore sequence has an average quality score of 10. (C) Zoomed in view of assembly graph color coded by GC% reveals *D. simplex* genome contigs (red), bacterial contigs (blue), and low, unclassified contigs (green). *D. simplex* genome contigs are clustered in a “hairball” due to the repeat structures in this genome. One bacterium appeared as a 6 Mb assembly. (D) GC% by length plot reveals *D. simplex* genome and bacterial genomes.

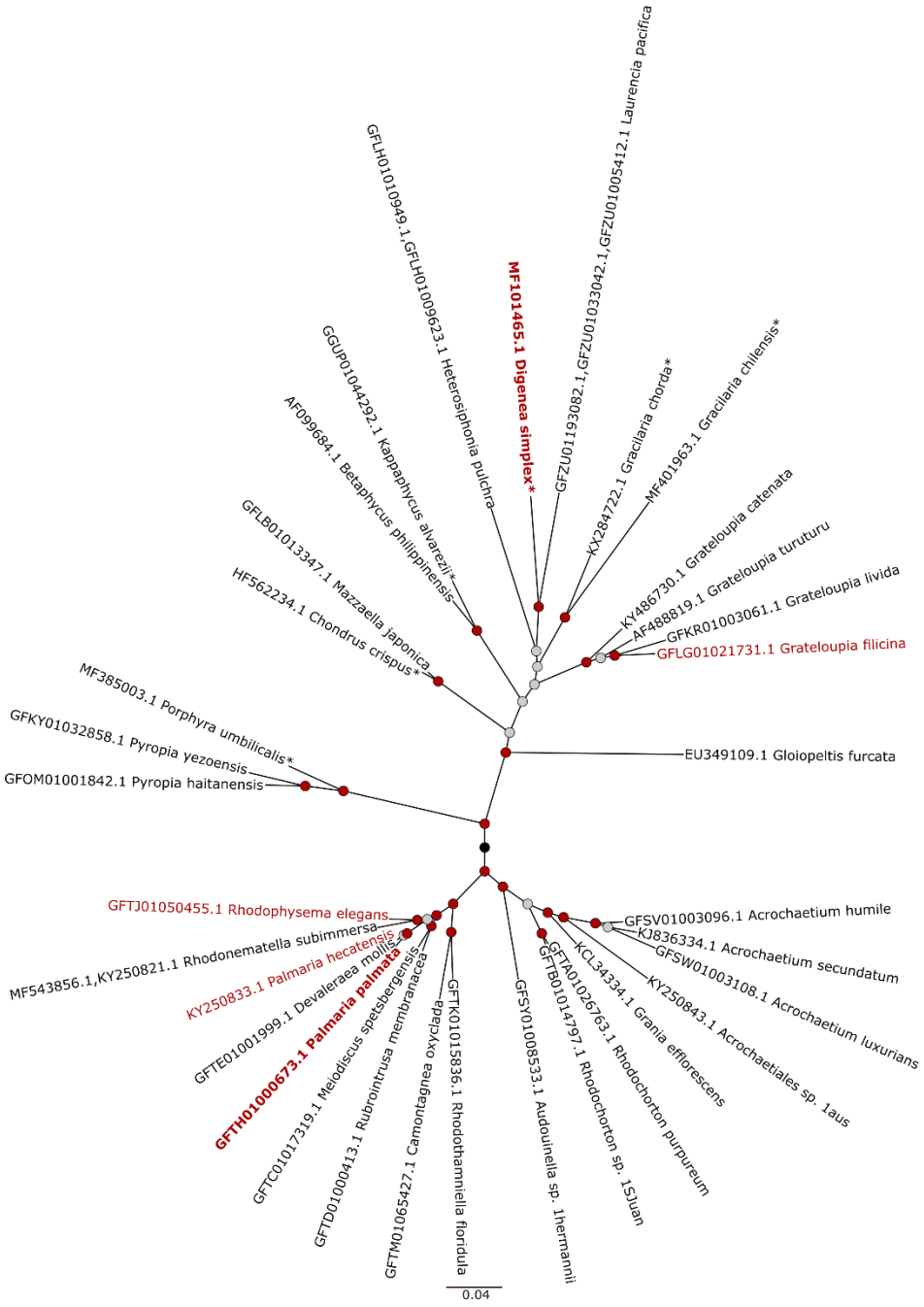


Figure S2: Unrooted maximum likelihood tree of ribulose-1,5-bisphosphate carboxylase (RuBisCO) large subunit of red macroalgae with RNASeq or genome data available (indicated by *). Those in indicated in red have the *kabA*/*C* pair. *P. palmata* and *D. simplex* are known kainic acid producers (bold). NCBI accession codes are listed. Nodes in red are supported with standard bootstrap values > 0.75. Previous phylogenetic work indicates that the plastid of heterokonts, of which diatoms are members, arose from a red algal endosymbiotic event.[14,15] This indicates that diatoms and red algae are likely evolutionarily related and suggests a mechanism of gene transfer to domoic acid producing diatoms.

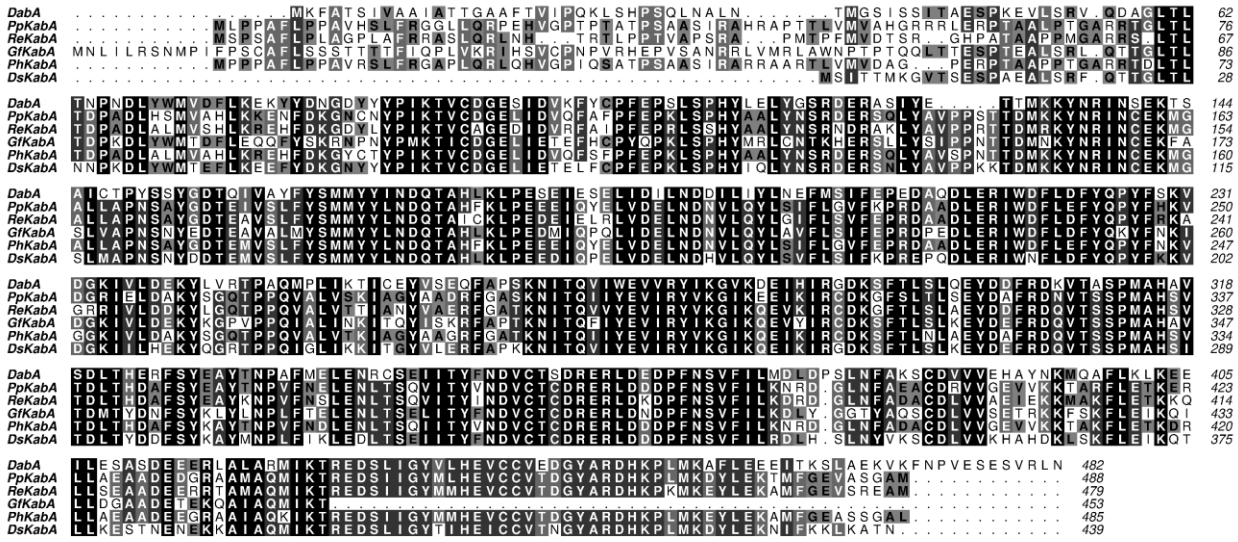


Figure S3: Amino acid sequence alignment of the domoic acid DabA and kainic acid KabA proteins. The accession numbers for proteins from this study can be found in Table S3 and the *Pseudo-nitzschia multiseries* DabA NCBI accession number is AYD91073.1.

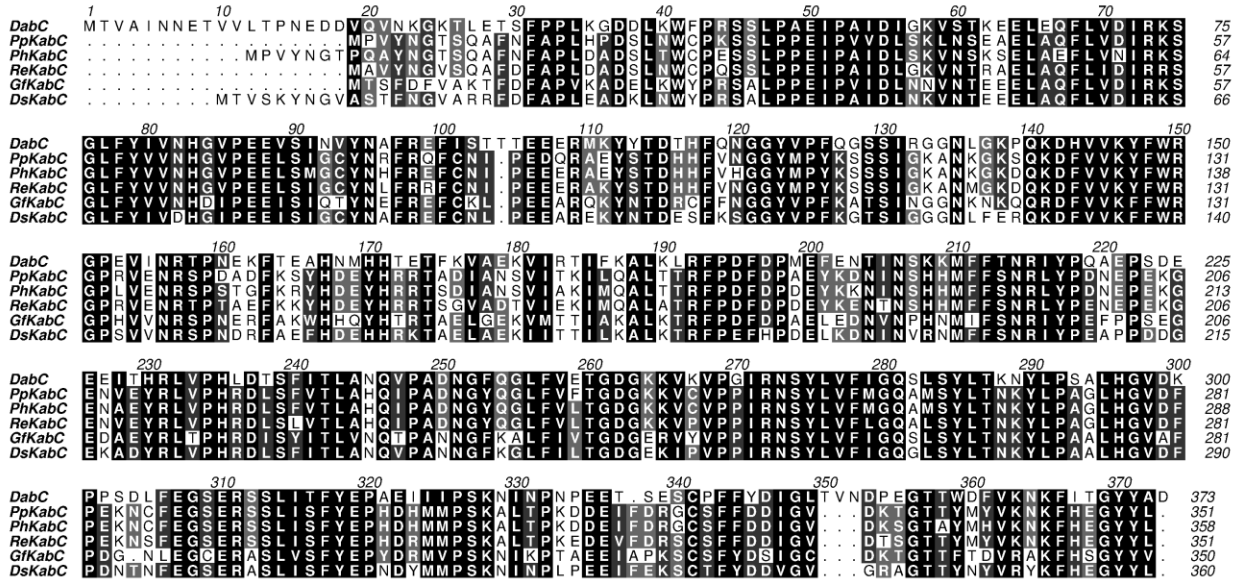


Figure S4: Amino acid sequence alignment of the domoic acid DabC and kainic acid KabC proteins. The accession numbers for proteins from this study can be found in Table S3 and the *Pseudo-nitzschia multiseries* DabA NCBI accession number is AYD91075.1.

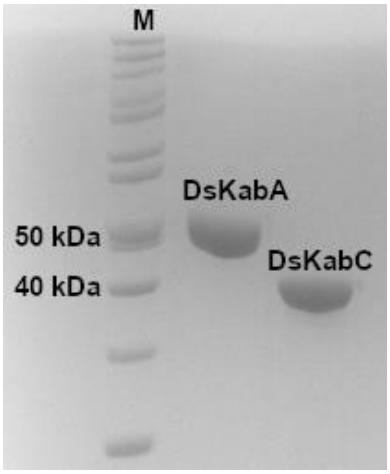


Figure S5: 12% SDS-PAGE loaded with EZ-Run *Rec* Protein Ladder (Fisher Bioreagents), DsKabA, and DsKabC.

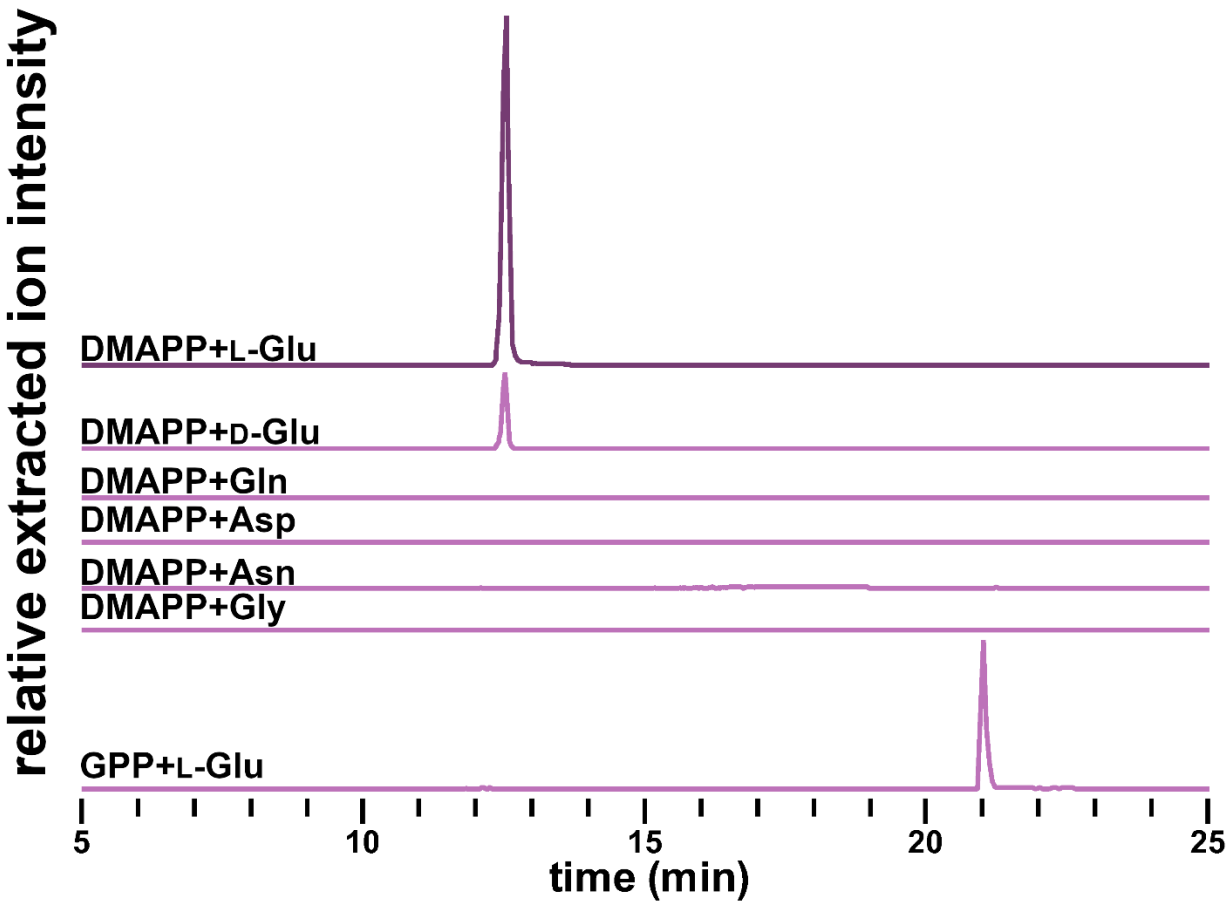


Figure S6: DsKabA substrate specificity was analyzed by LCMS. Each trace represents the extracted ion chromatogram for the expected $(M-H)^-$ product mass ± 0.5 *m/z*. DMAPP+L-Glu: 214.1 *m/z*, DMAPP+D-Glu: 214.1 *m/z*, DMAPP+L-Gln: 213.1 *m/z*, DMAPP+L-Asp: 200.1 *m/z*, DMAPP+L-Asn: 199.1 *m/z*, DMAPP+Gly: 142.2 *m/z*, and GPP+L-Glu: 282.2 *m/z*.

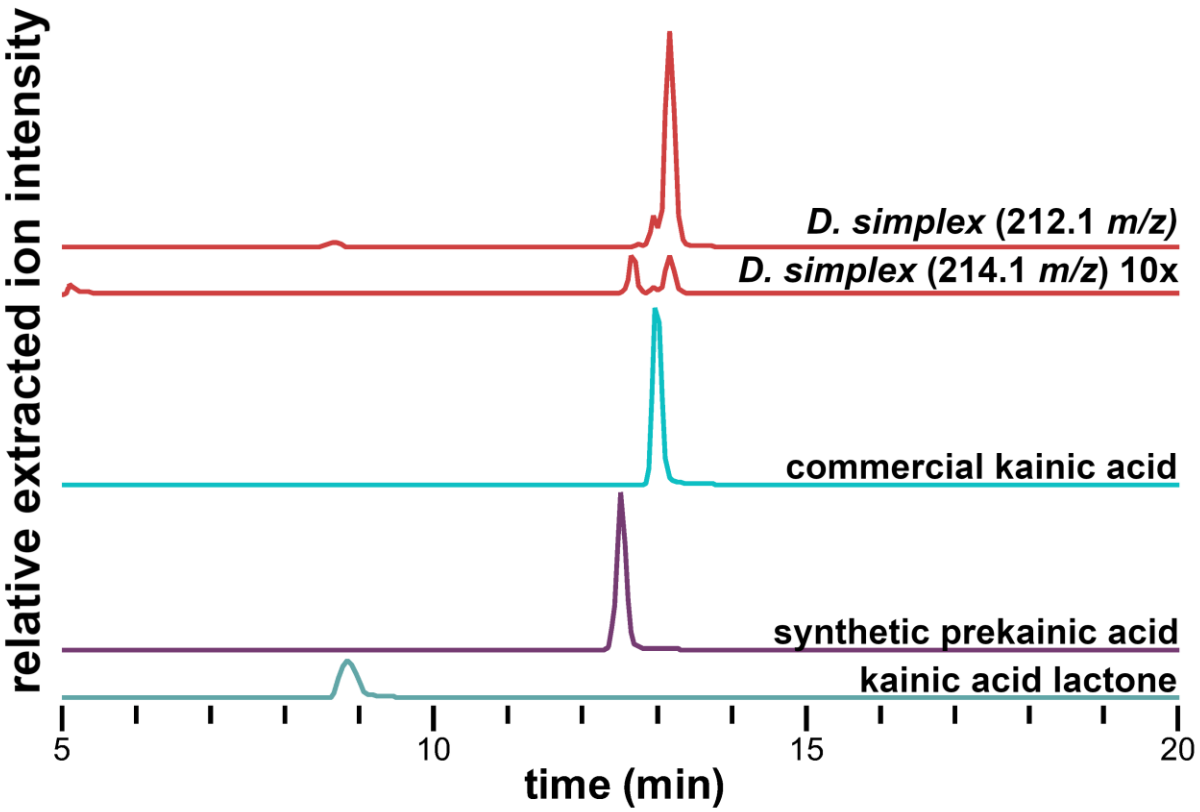


Figure S7: An aqueous *D. simplex* extract was analyzed by LCMS and compared to synthetic/enzymatic standards to verify the presence of kainic acid, kainic acid lactone, and prekainic acid. Each trace represents the $(M\text{-}H)^-$ extracted ion chromatogram for prekainic acid (214.1 ± 0.1 *m/z*) and kainic acid/kainic acid lactone (212.1 ± 0.1 *m/z*). The *D. simplex* trace extracted for the prekainic acid mass (214.1 ± 0.1 *m/z*) was magnified tenfold.

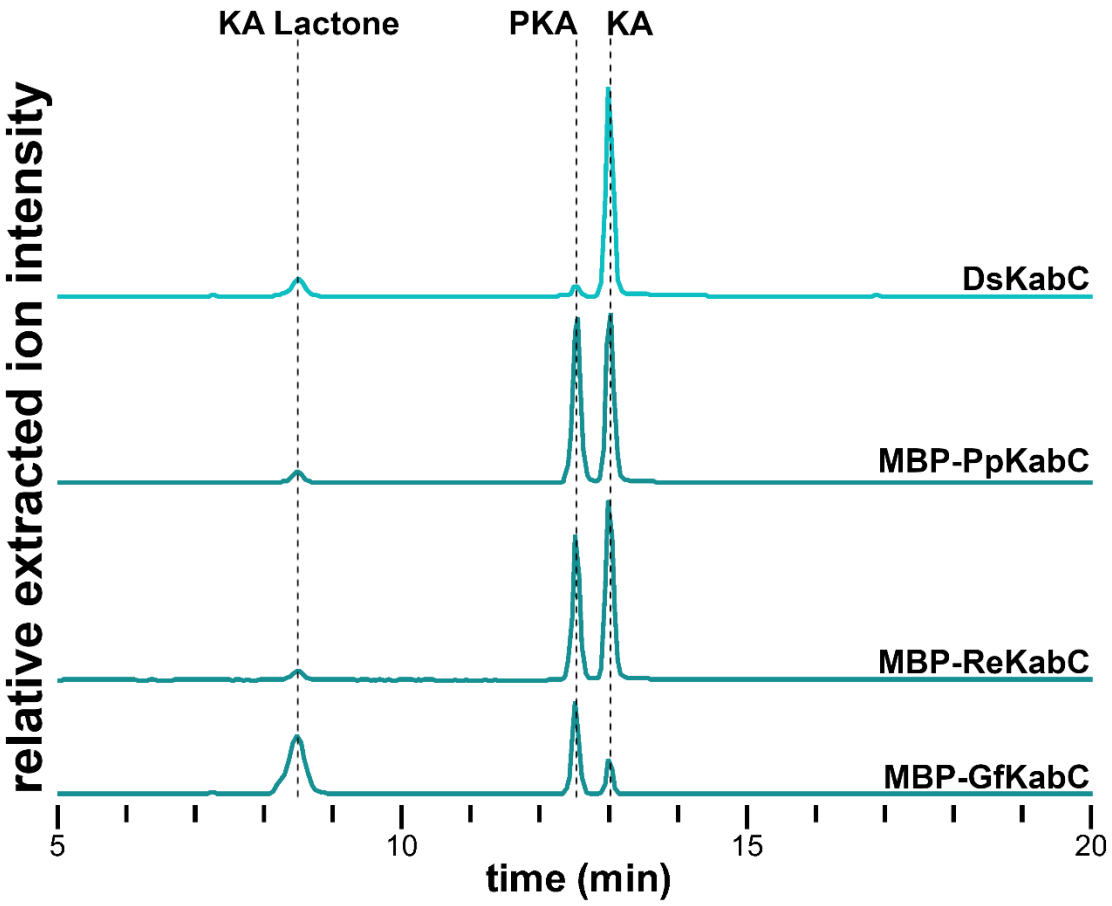


Figure S8: The reaction products of KabC enzymes from *D. simplex*, *P. palmata*, *R. elegans*, and *G. filicinia* were analyzed by LCMS following a 5 h incubation. DsKabC was used in the His_6 tag cleaved form while the other three homologs were tested as an MBP-fusion. Each trace represents the $(M\text{-}H)^-$ extracted ion chromatogram for both prekainic acid (214.1 ± 0.5 *m/z*) and kainic acid (212.1 ± 0.5 *m/z*).

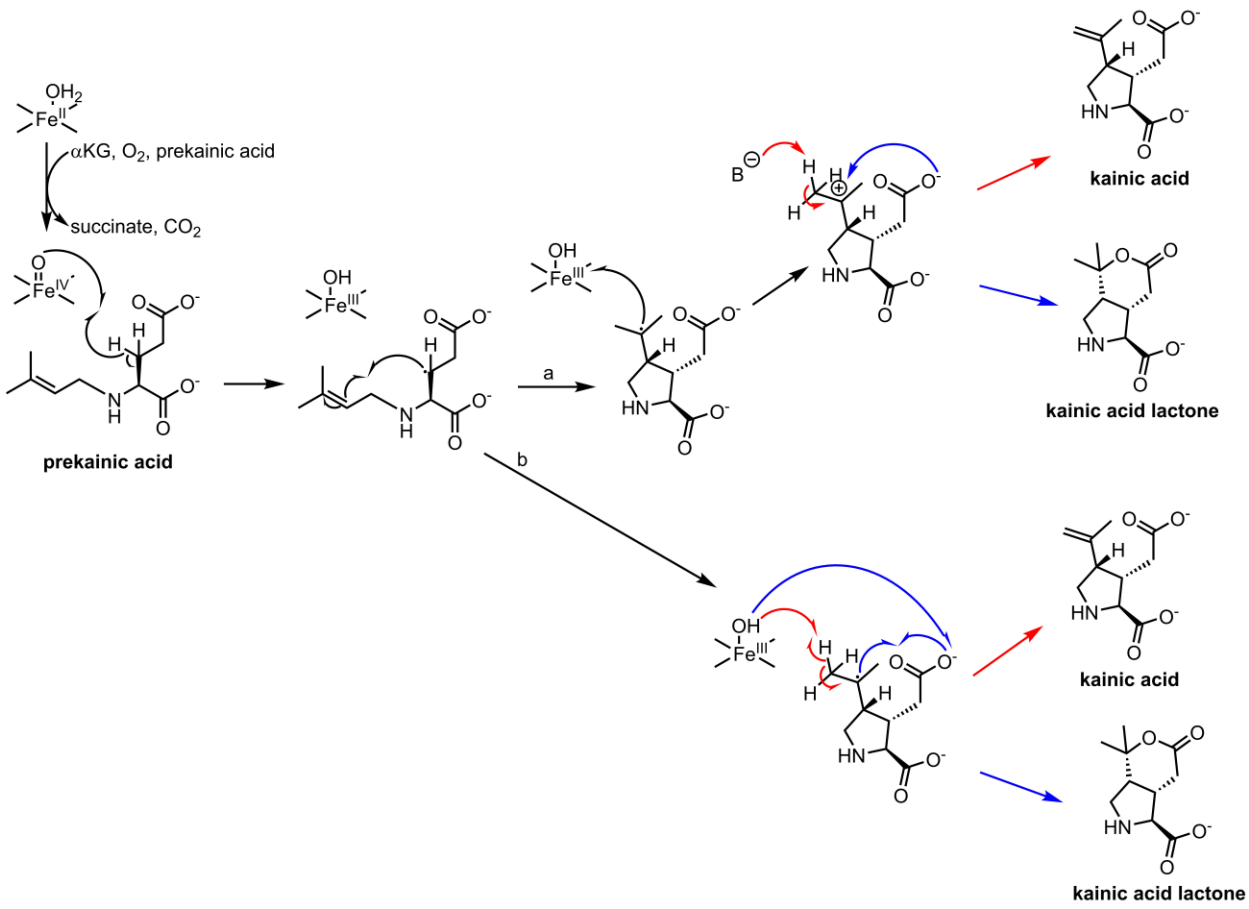


Figure S9: Proposed KabC mechanism can explain the presence of both kainic acid and kainic acid lactone. The reaction could procedure through two distinct mechanisms: **a** electron transfer (ET) or **b** hydrogen atom transfer (HAT)[16].

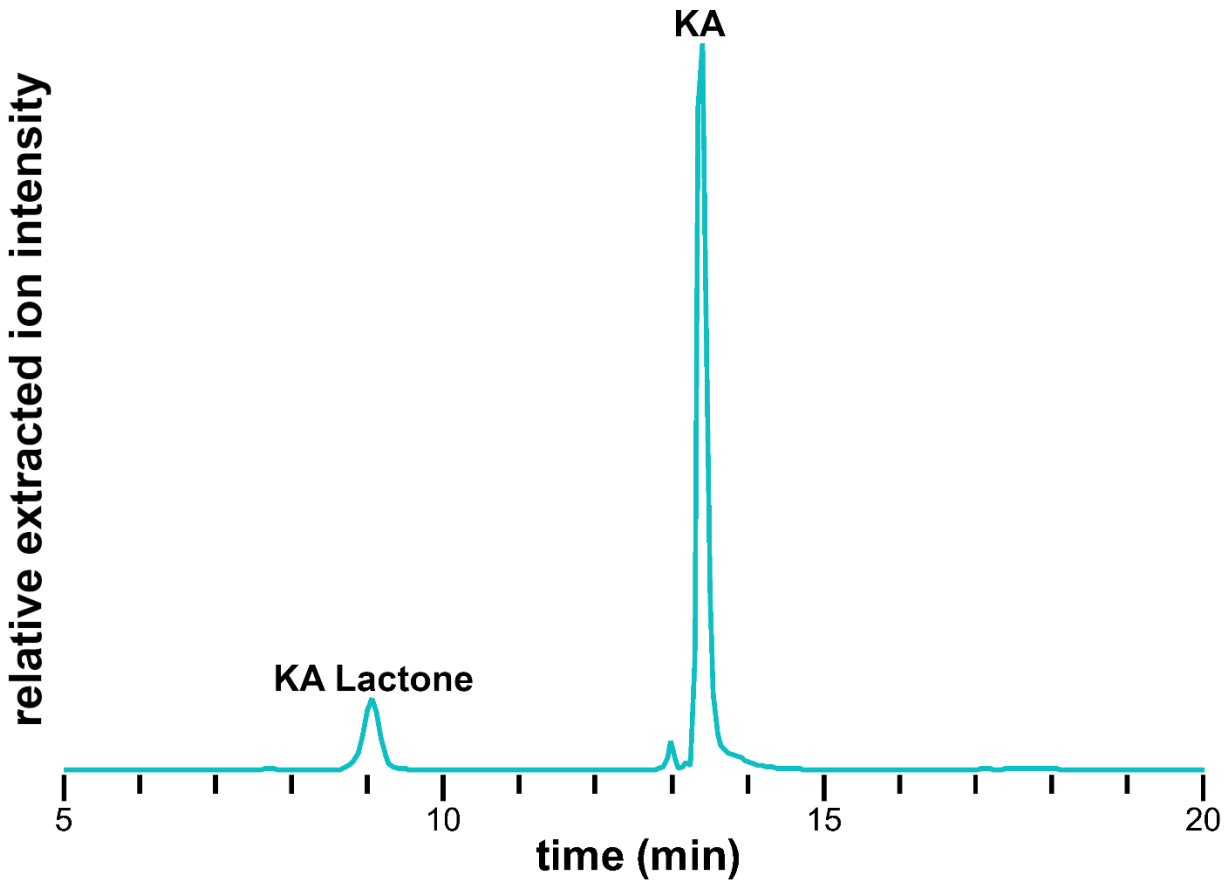


Figure S10: Kainic acid production by in vitro conversion of synthetic prekainic acid by DsKabC. The trace represents the $(M-H)^-$ extraction ion chromatogram for prekainic acid (PKA, 214.1 ± 0.1 *m/z*) and kainic acid/kainic acid lactone (KA/KA Lactone, 212.1 ± 0.1 *m/z*). The prekainic acid peak is <0.3% of the kainic acid intensity and is not visible.

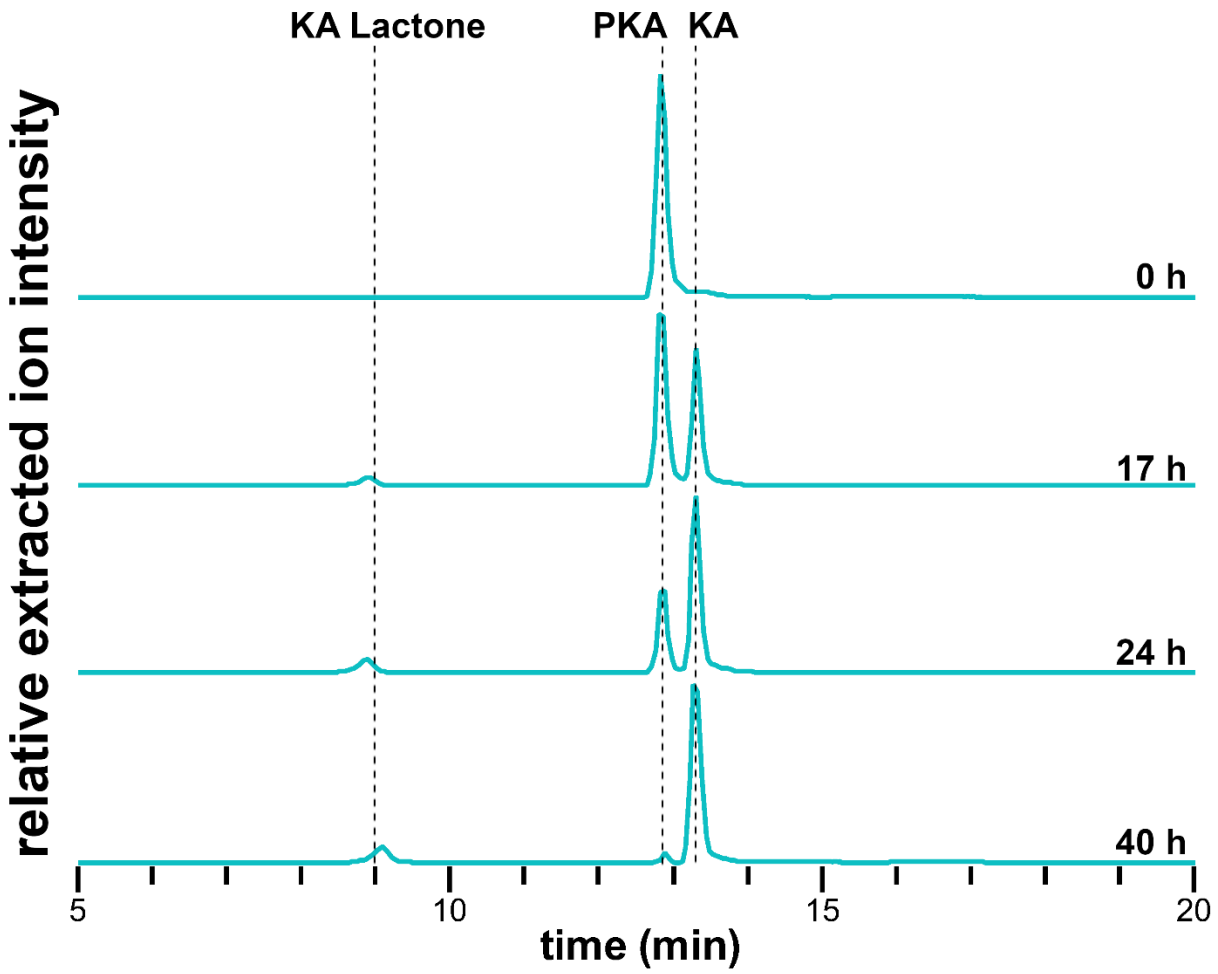


Figure S11: Time course of the conversion of prekainic acid to kainic acid in the media of *E. coli* expressing the *dskabC* gene on a pET28 vector. Each trace represents the $(M\text{-}H)^-$ extracted ion chromatogram for both prekainic acid (214.1 ± 0.1 *m/z*) and kainic acid (212.1 ± 0.1 *m/z*).

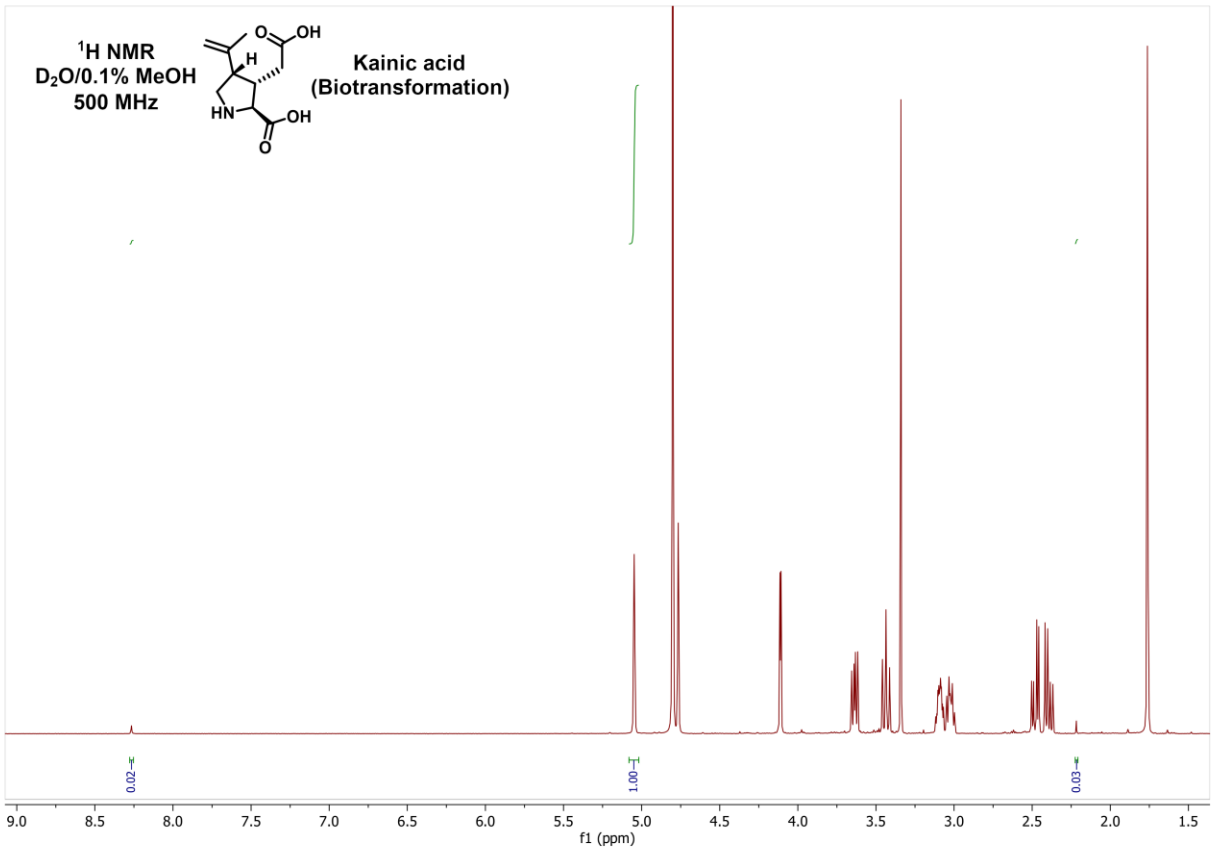
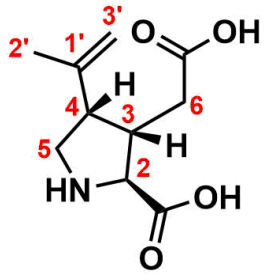
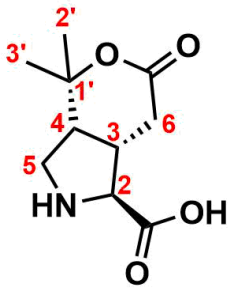


Figure S12: ^{1}H NMR was used to confirm the purity of kainic acid isolated from biotransformation by *E. coli*.

NMR Analysis



kainic acid



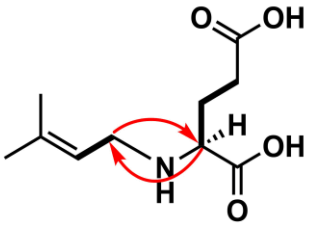
kainic acid lactone

NMR Comparison (Chemical shifts)

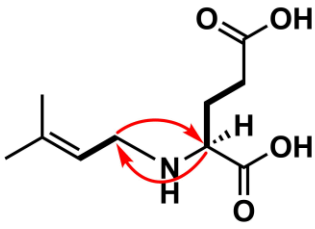
Proton	Literature Kainic Acid[17]	Literature *allo*-kainic acid[17]	Purified Kainic Acid	Proton	Literature Kainic acid lactone[18]	Purified Kainic acid lactone
2	4.01	3.37	4.07	2	3.98	3.97
3	3.00	3.10	3.05	5a	3.76	3.73
4	3.00	2.90	2.97	3, 5b, 6a*	3.1-3.4	3.13-3.24
5a	3.54	3.54	3.60	5b		3.24
5b	3.34	3.30	3.40	3		3.20
6a	2.3	2.35	2.34	6a		3.15
6b	2.38	2.41	2.44	4	2.82	2.79
2’	1.67	1.68	1.73	6b	2.60	2.55
3’	4.95	4.97	4.98, 4.73	2’	1.50	1.48
				3’	1.42	1.39

*Literature kainic acid lactone shifts of these protons are reported as a range.

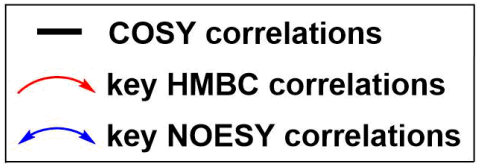
NMR Correlations

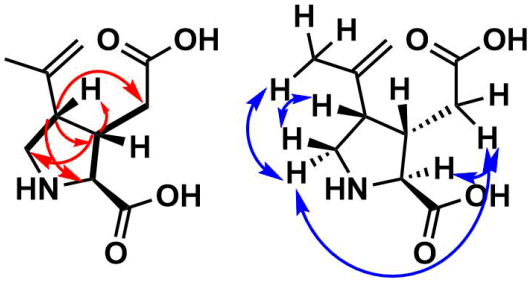


prekainic acid
(enzymatic)

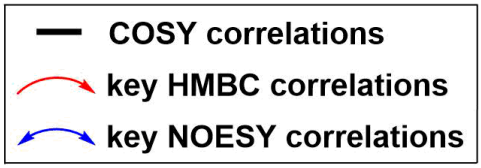


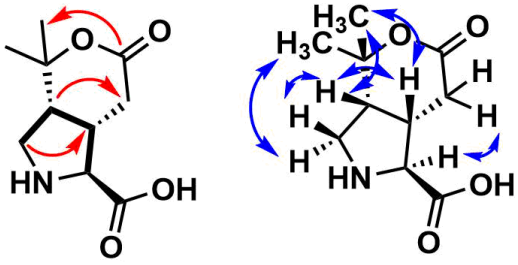
prekainic acid
(synthetic)





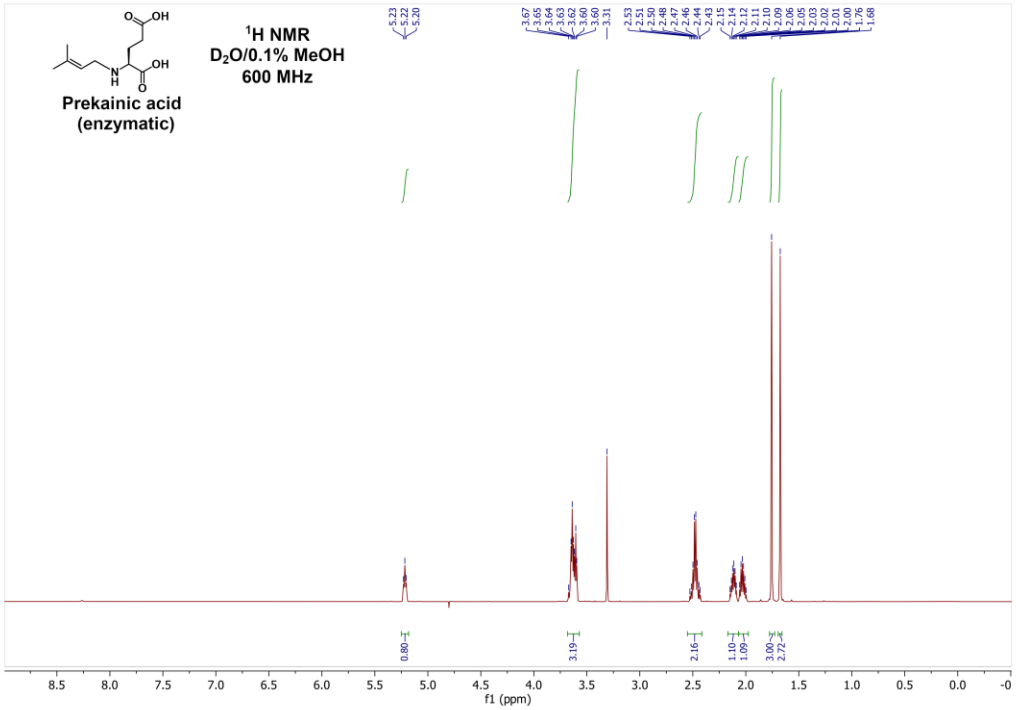
kainic acid

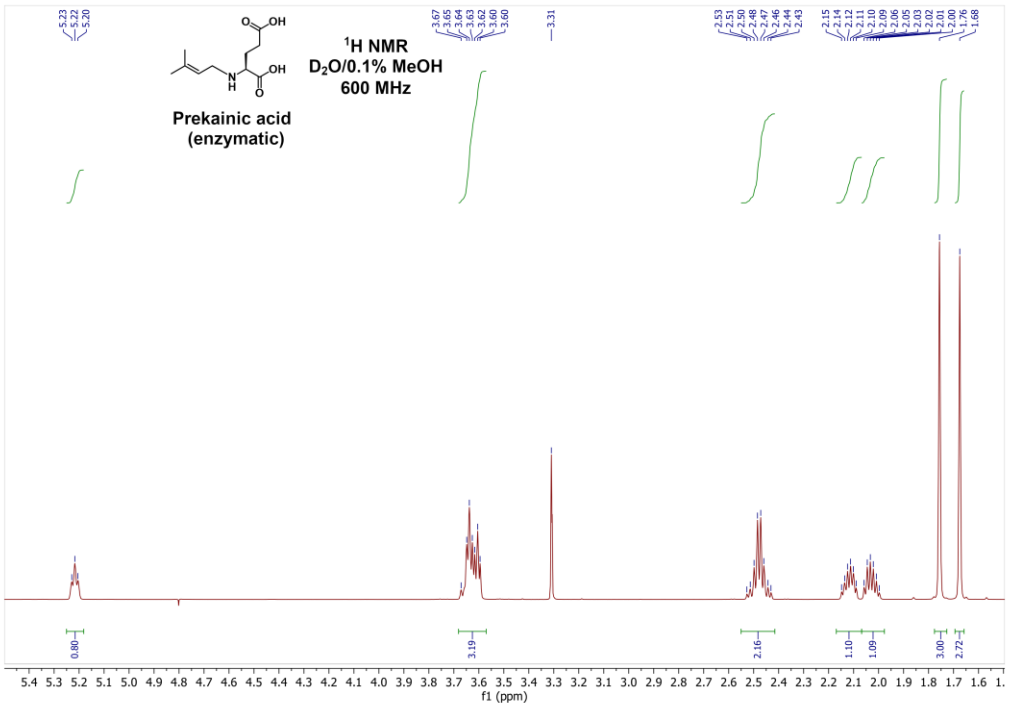


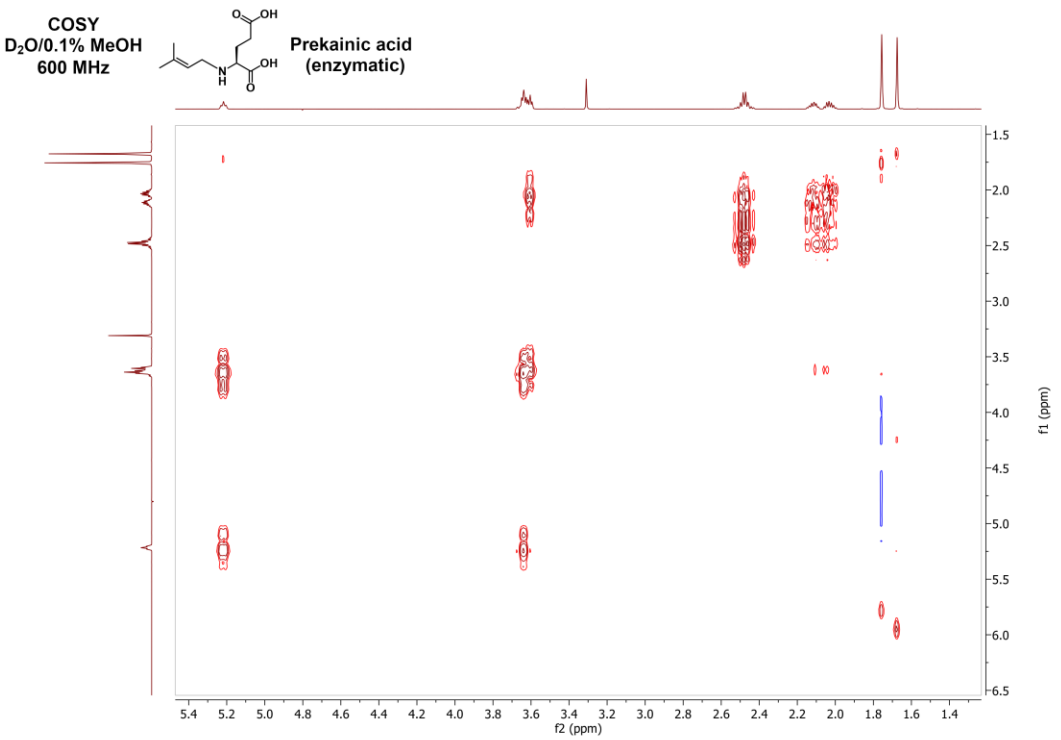


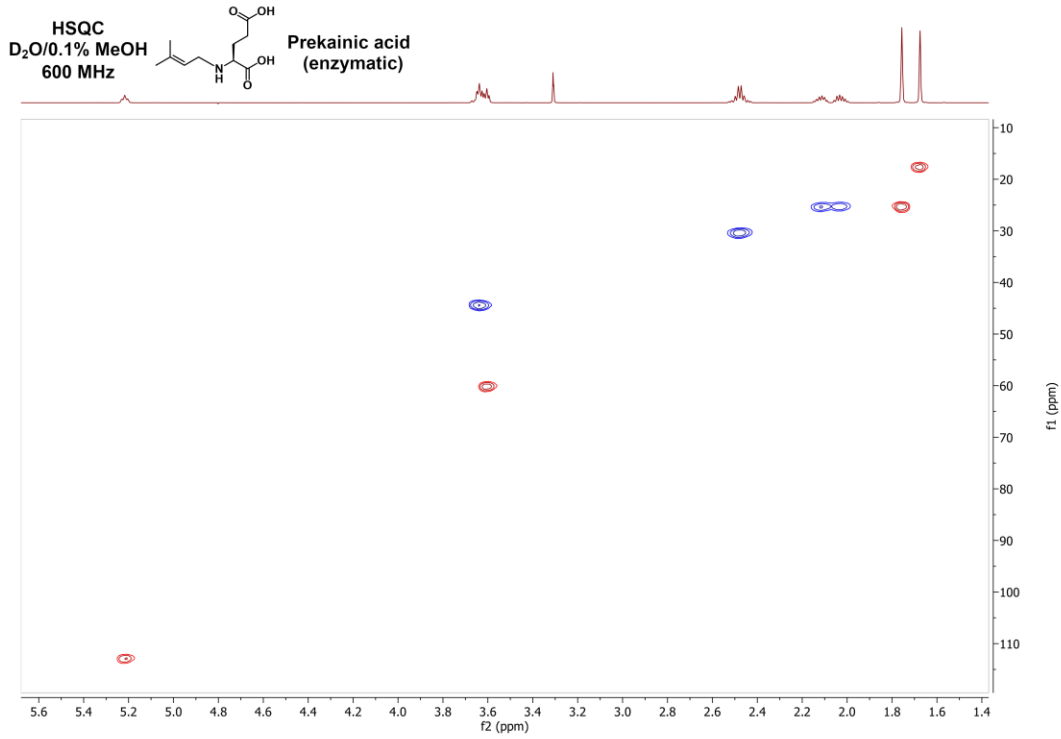
kainic acid lactone

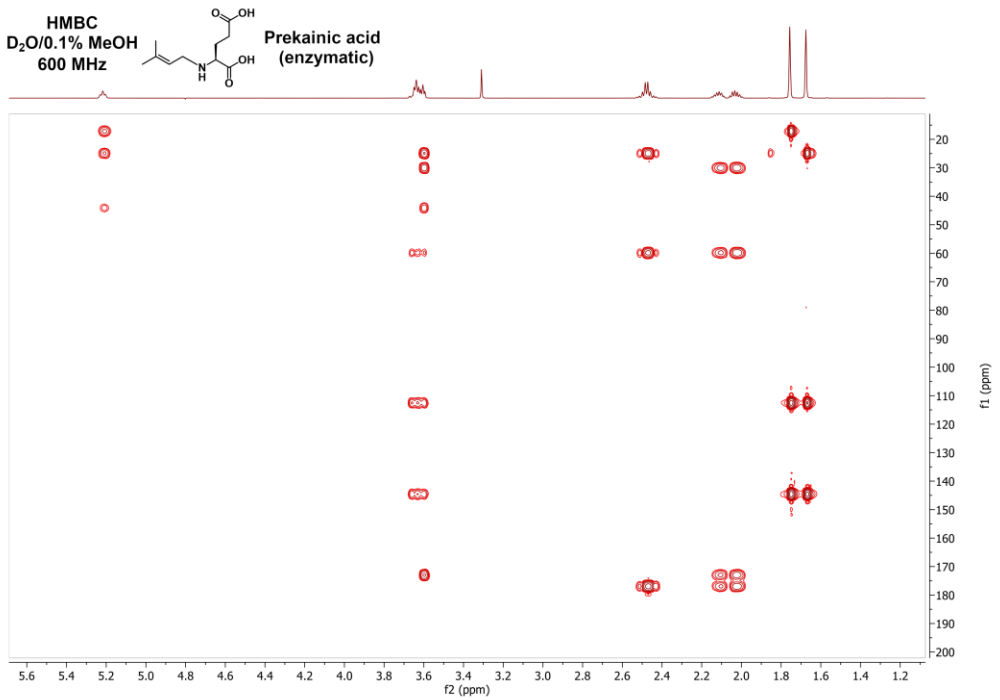
Prekainic acid (DsKabA product):

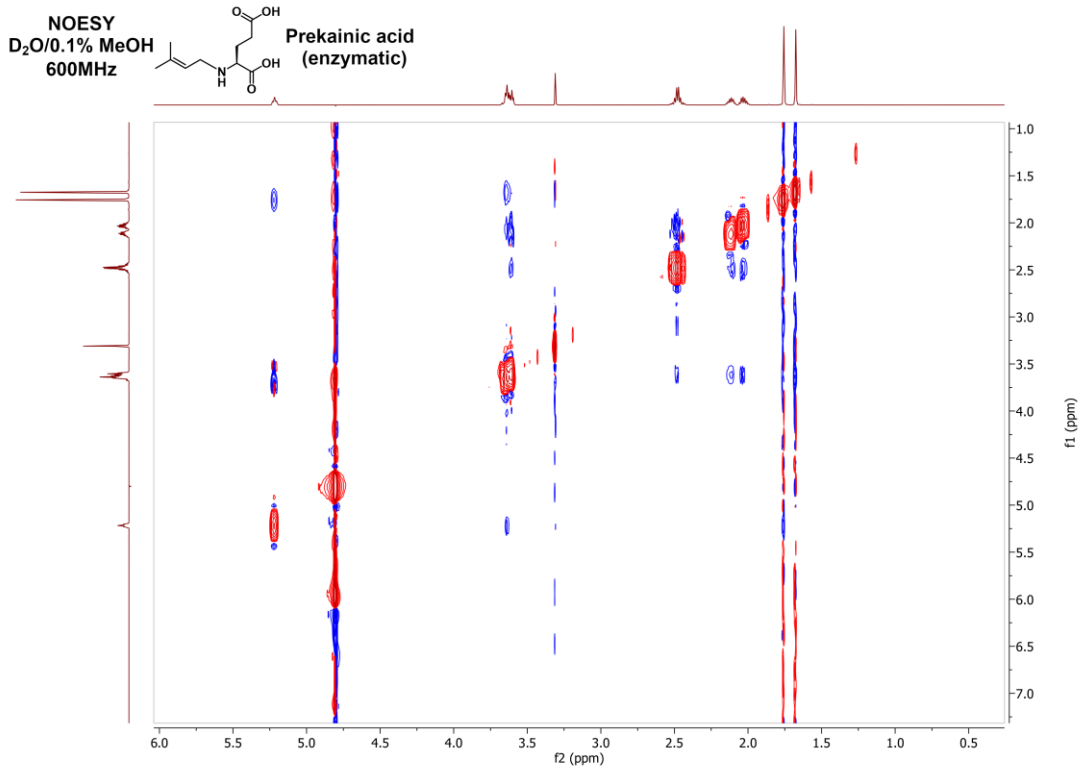




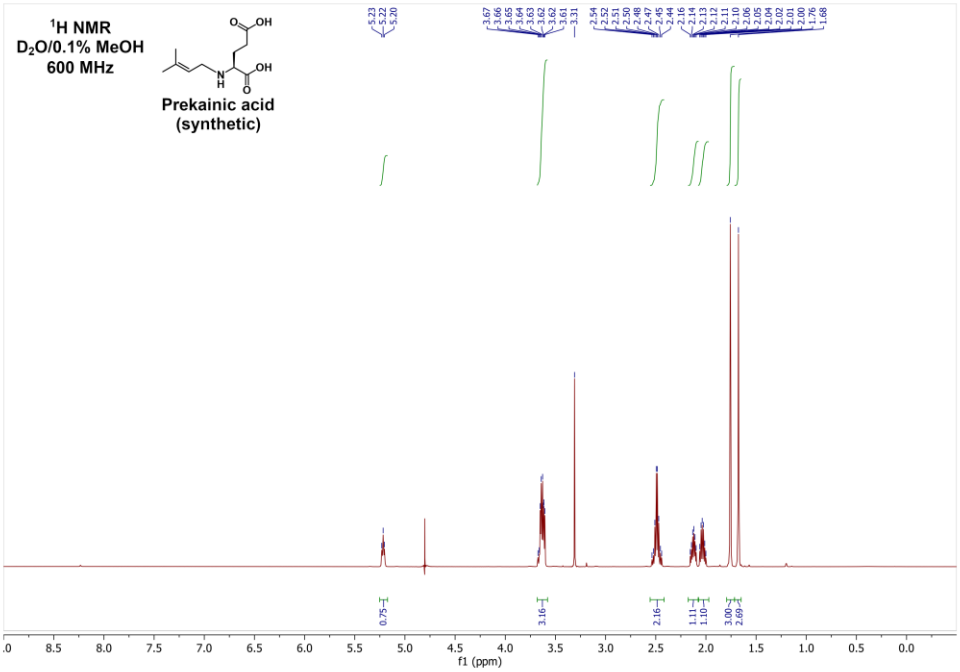

COSY
D2O/0.1% MeOH
600 MHz
Prekainic acid
(enzymatic)
f2 (ppm)
f1 (ppm)

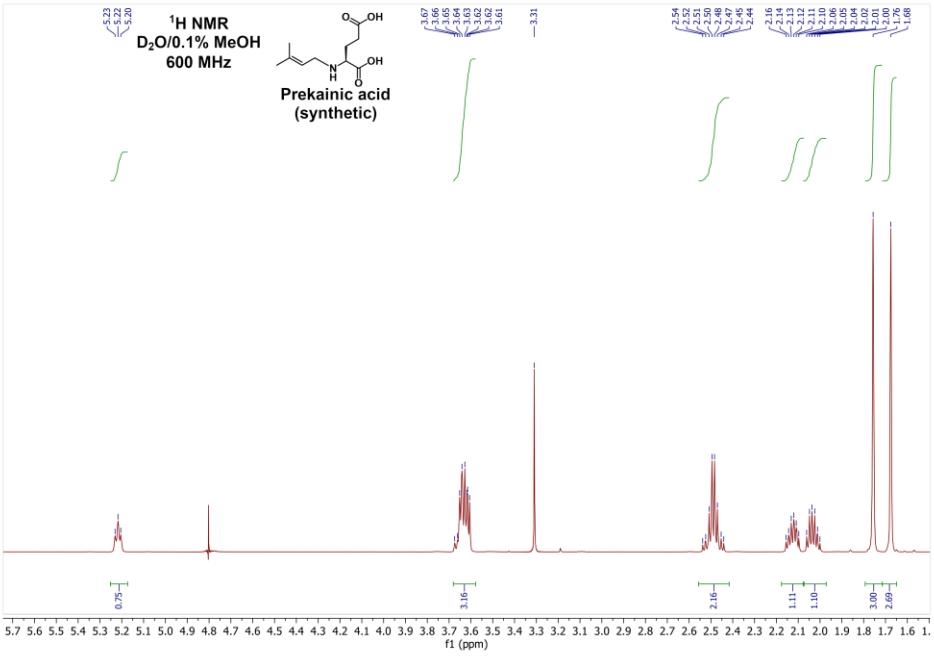

HSQC
D2O/0.1% MeOH
600 MHz
Prekainic acid
(enzymatic)
f2 (ppm)
f1 (ppm)

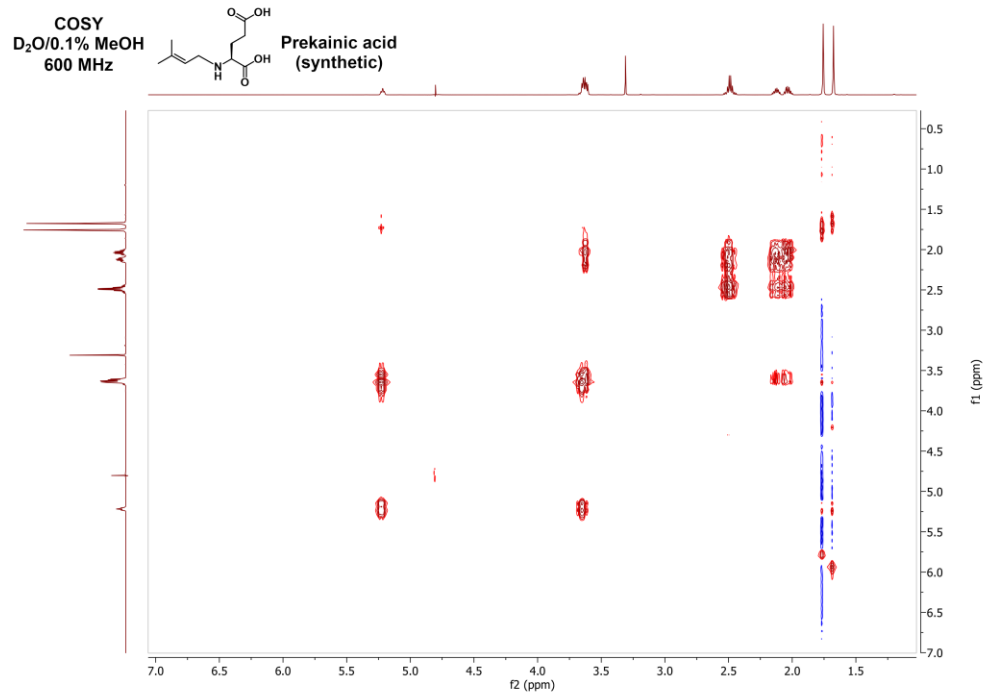

HMBC
D2O/0.1% MeOH
600 MHz
Prekainic acid
(enzymatic)
f2 (ppm)
f1 (ppm)

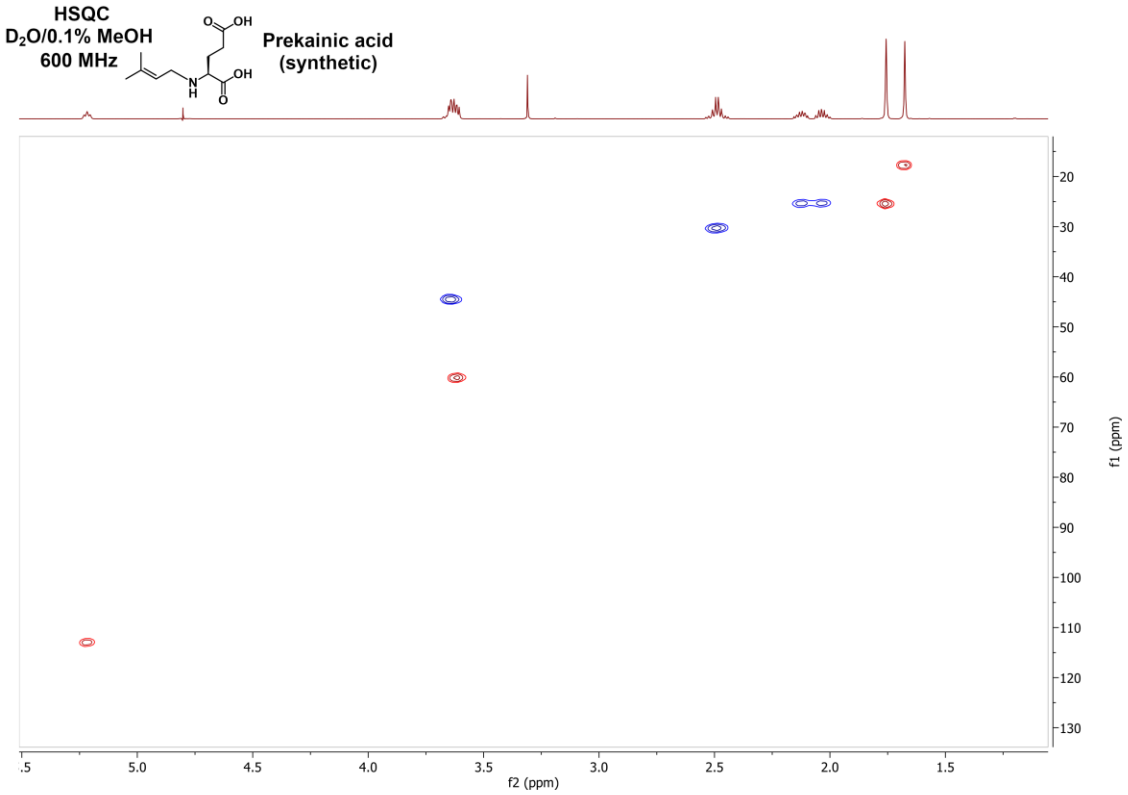

NOESY
D2O/0.1% MeOH
600MHz
Prekainic acid
(enzymatic)
f2 (ppm)
f1 (ppm)

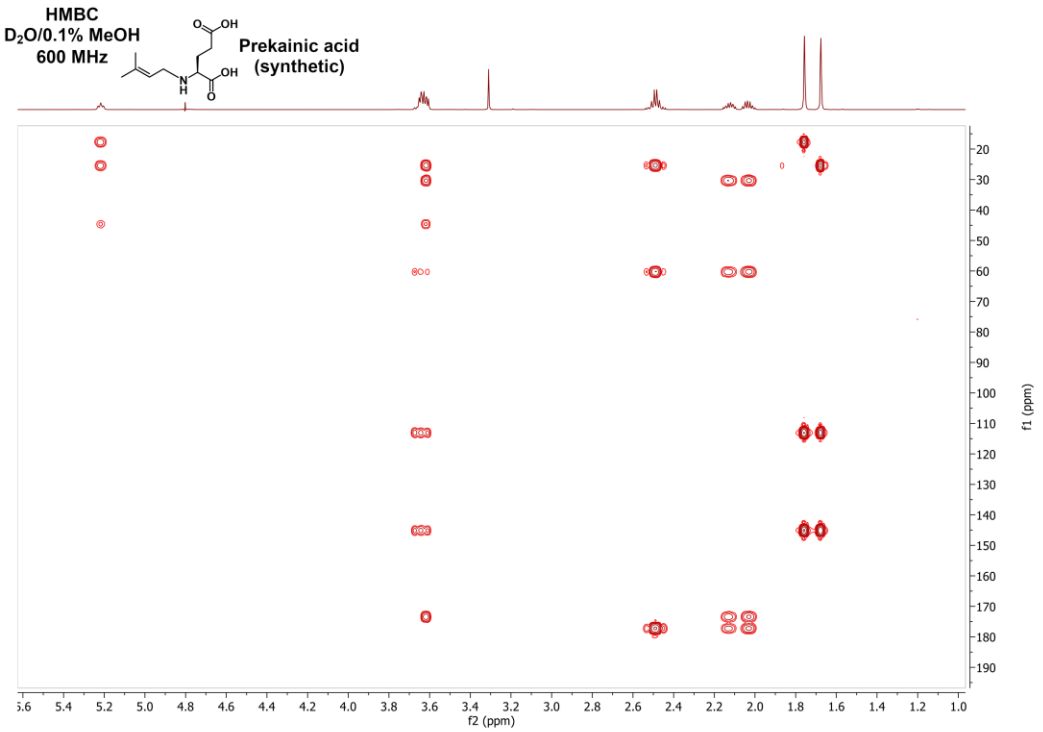
Prekainic acid (Synthetic)

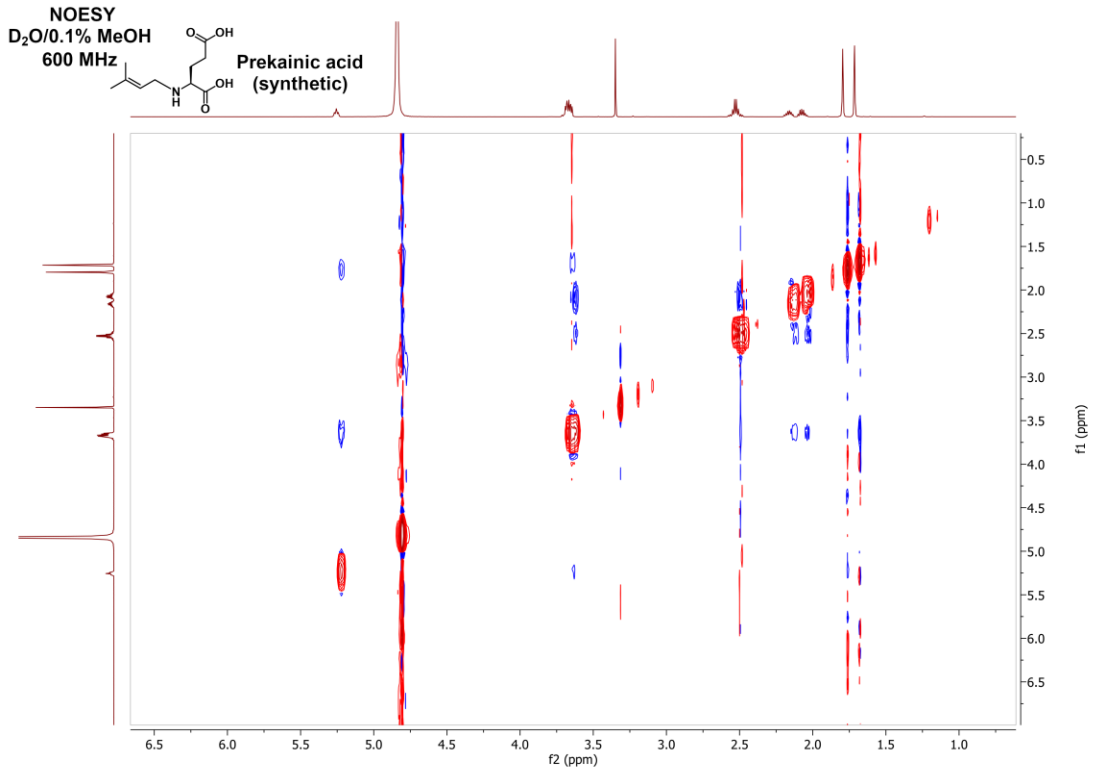




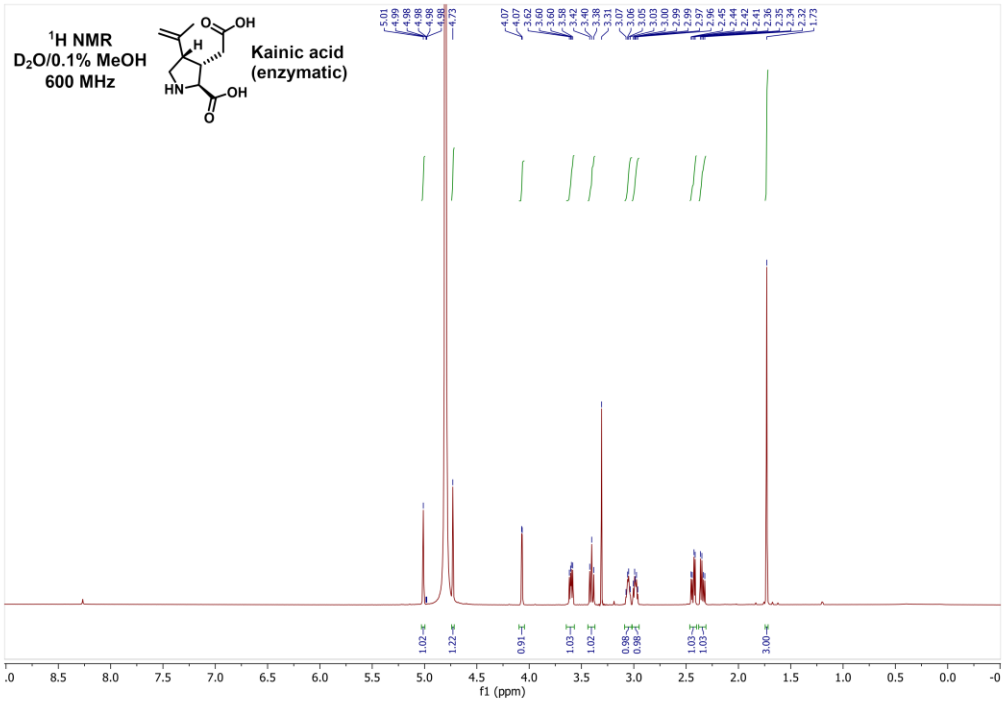

COSY
D2O/0.1% MeOH
600 MHz
Prekainic acid
(synthetic)
f2 (ppm)
f1 (ppm)

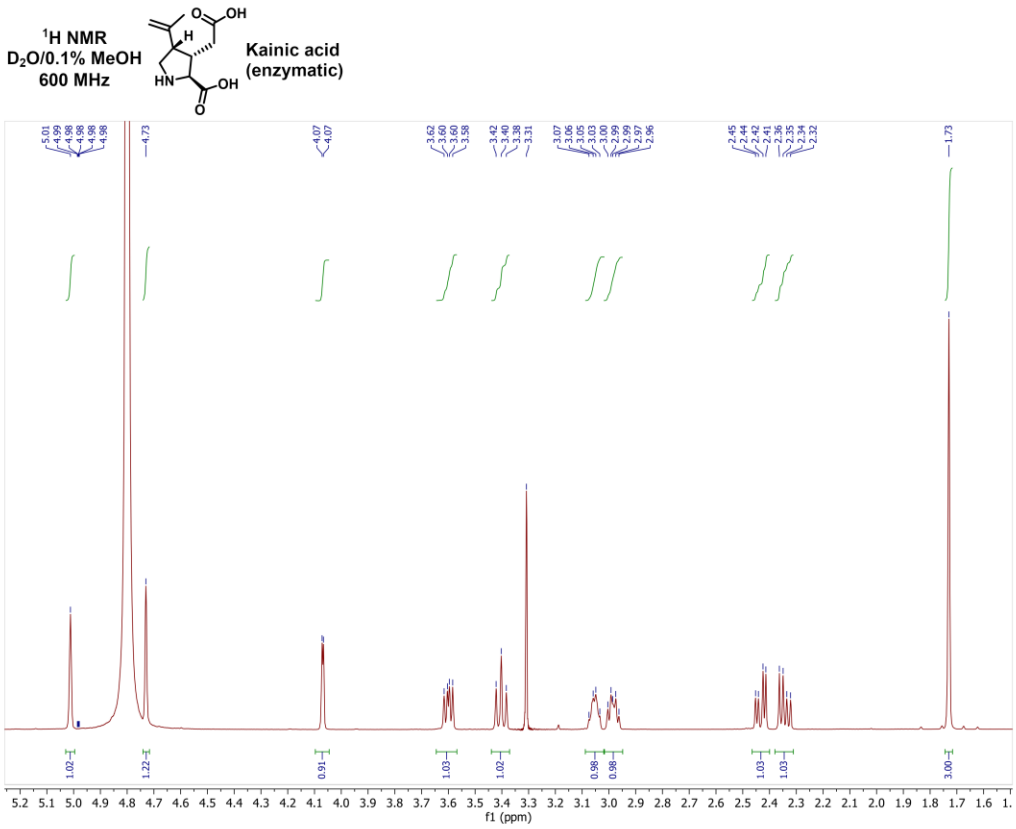

HSQC
D2O/0.1% MeOH
600 MHz
Prekainic acid
(synthetic)
f2 (ppm)
f1 (ppm)

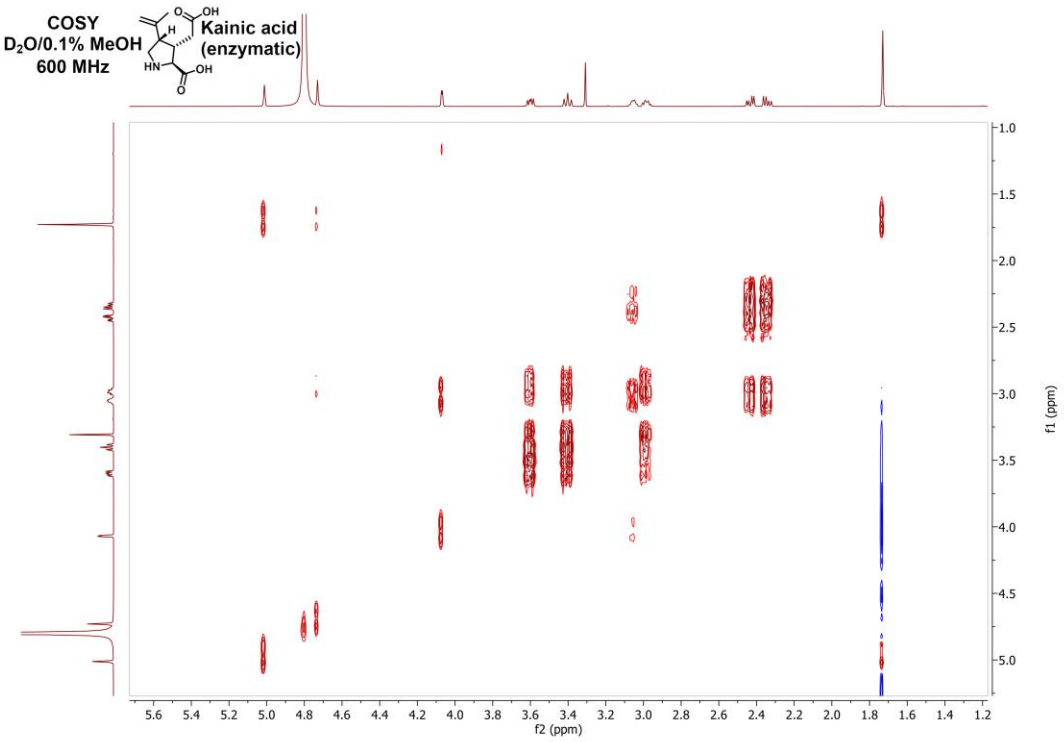

HMBC
D2O/0.1% MeOH
600 MHz
Prekainic acid
(synthetic)
f2 (ppm)
f1 (ppm)

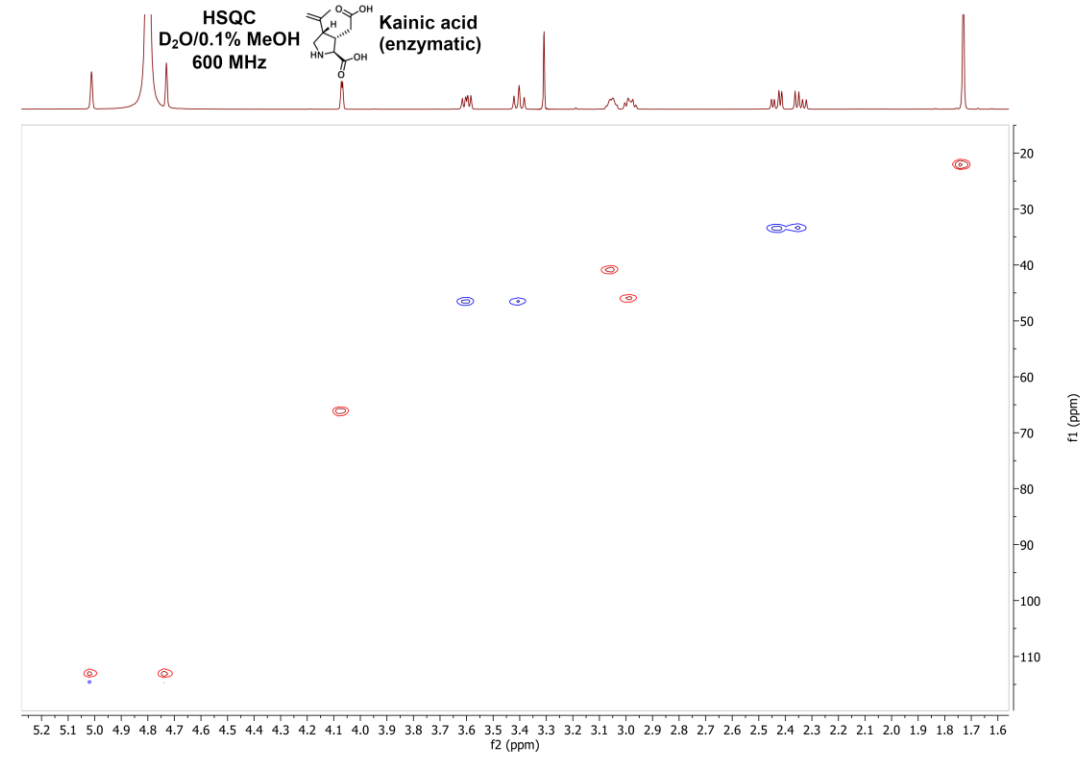

NOESY
D2O/0.1% MeOH
600 MHz
Prekainic acid
(synthetic)
f2 (ppm)
f1 (ppm)

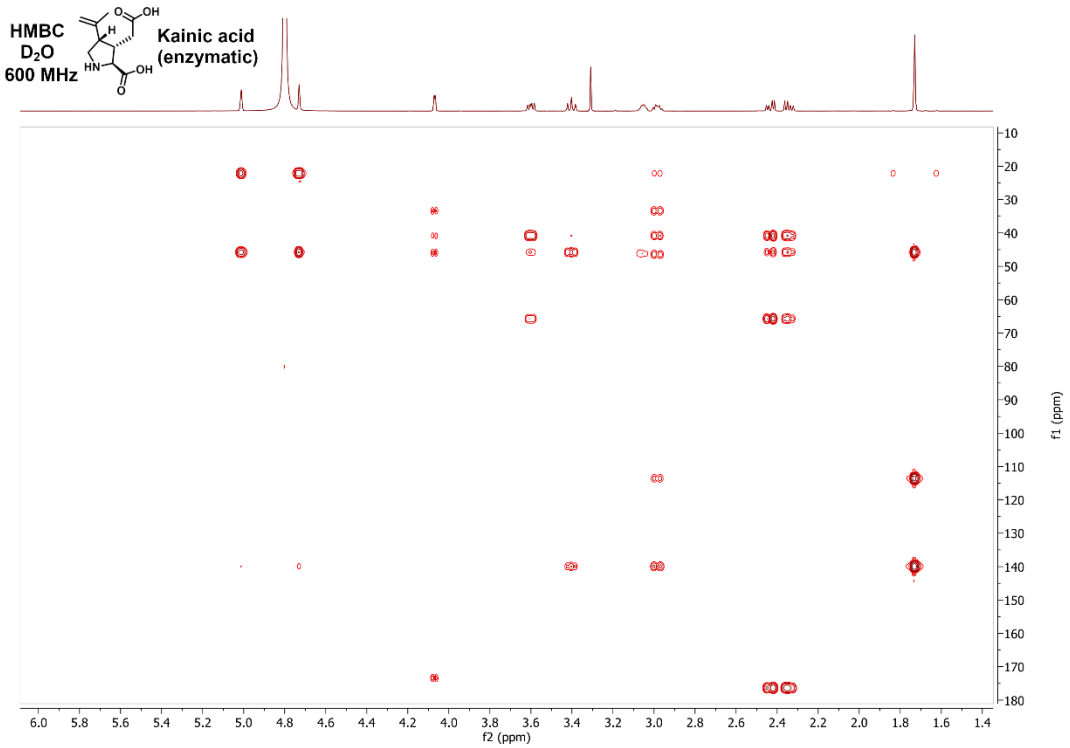
Kainic acid (DsKabC product)

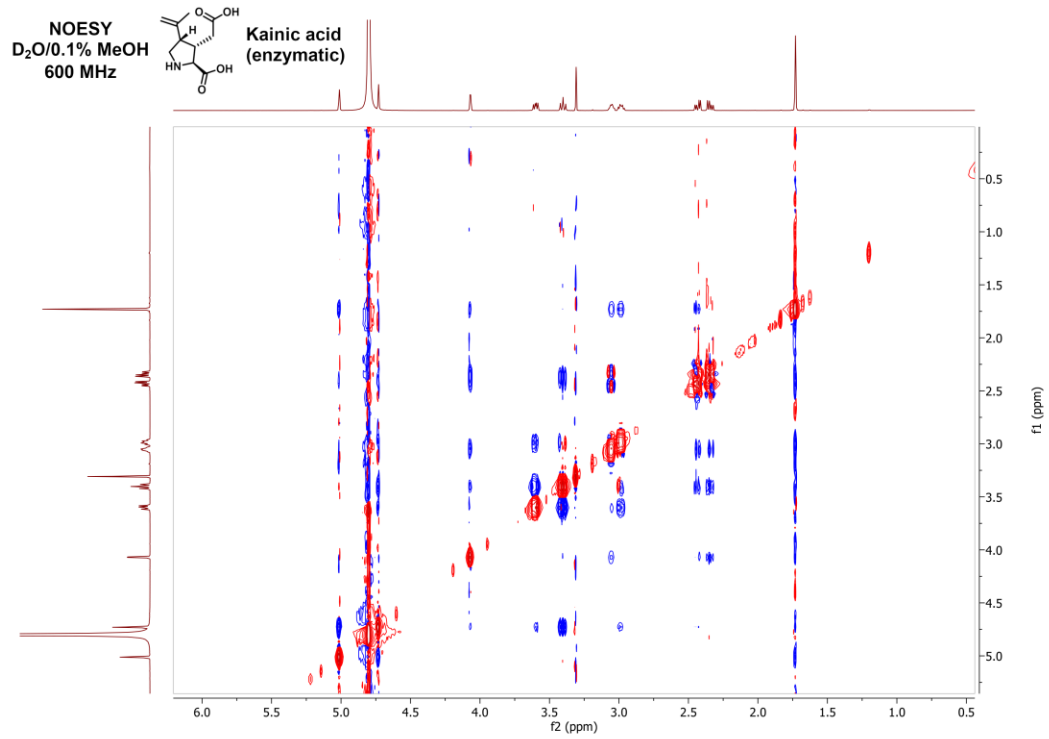




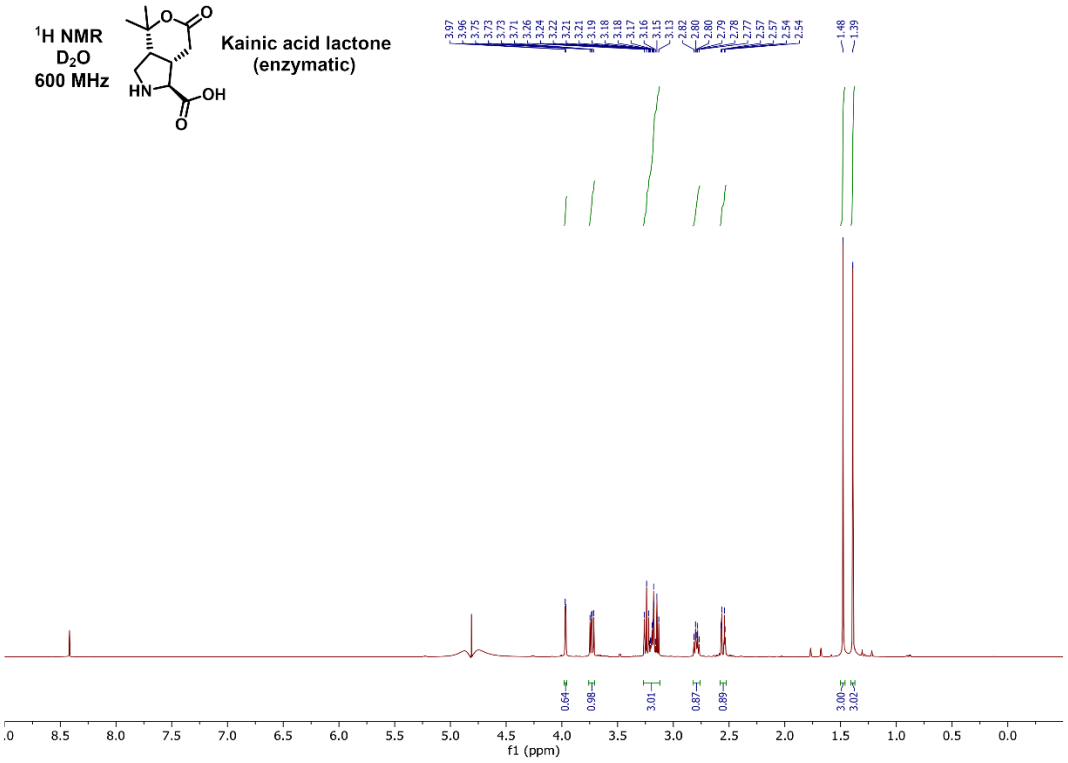

COSY
D2O/0.1% MeOH
600 MHz
Kainic acid
(enzymatic)
f2 (ppm)
f1 (ppm)

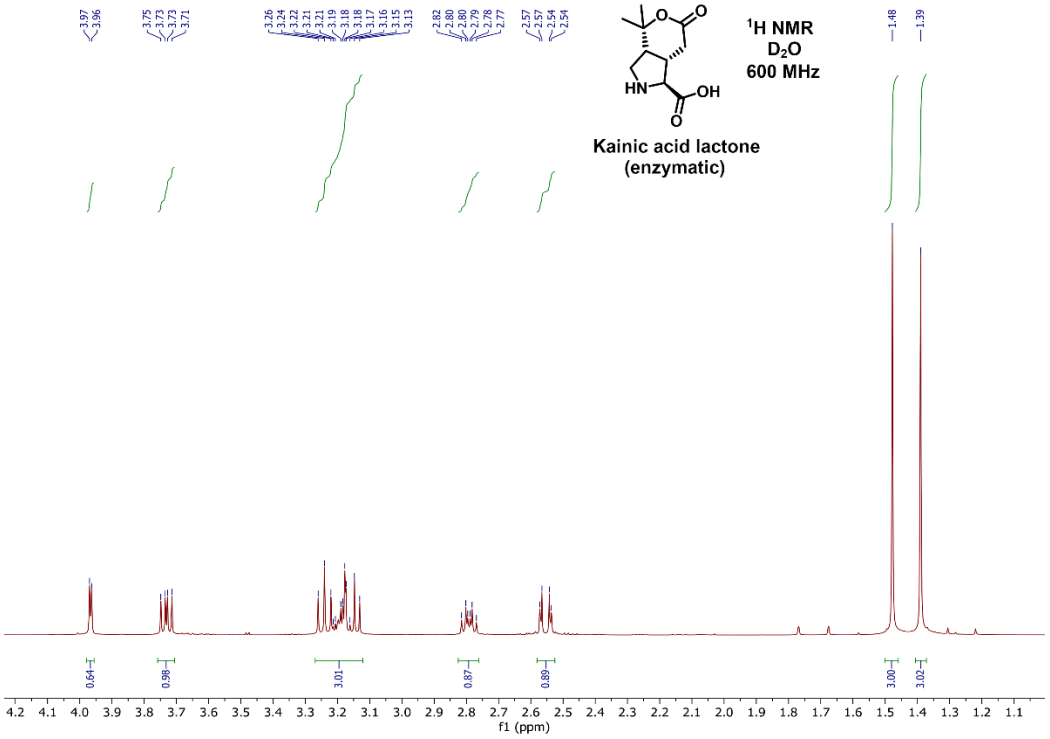

HSQC
D2O/0.1% MeOH
600 MHz
Kainic acid
(enzymatic)
f2 (ppm)
f1 (ppm)

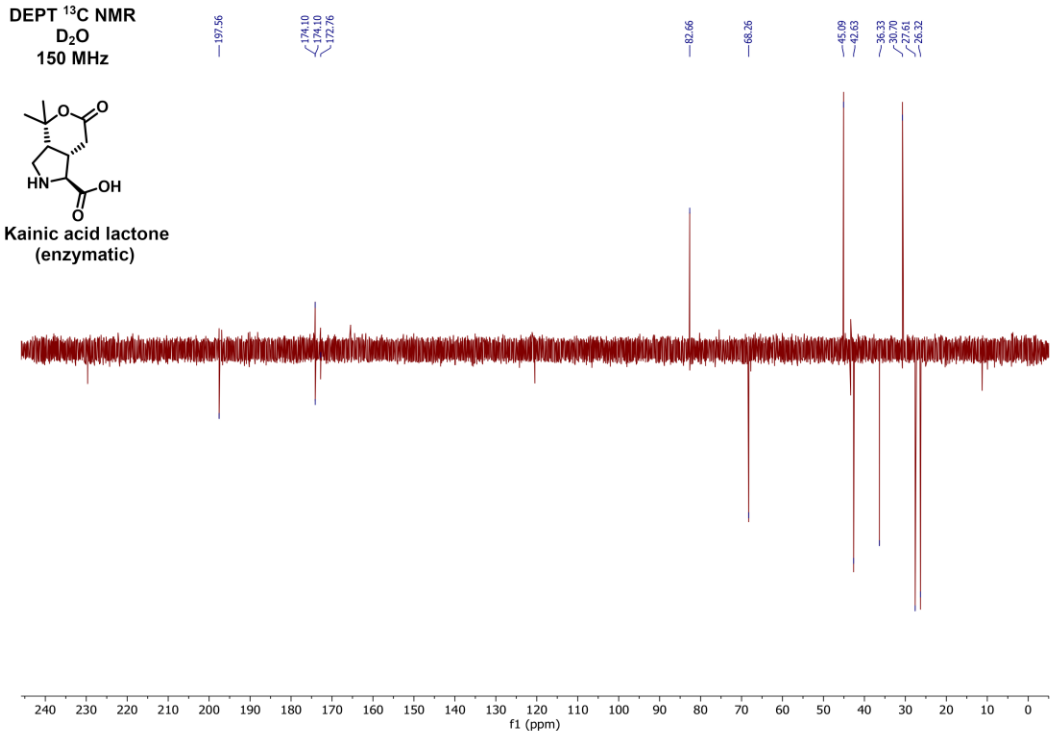

HMBC
D2O
600 MHz
Kainic acid
(enzymatic)
f2 (ppm)
f1 (ppm)

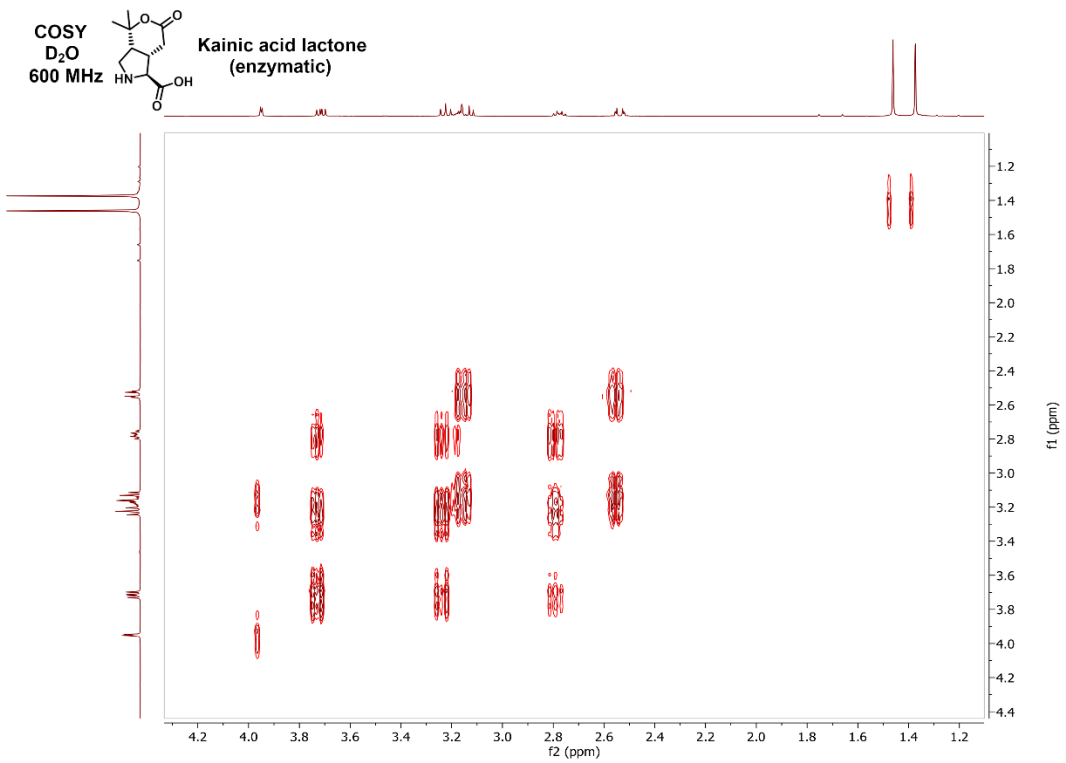

NOESY
D2O/0.1% MeOH
600 MHz
Kainic acid
(enzymatic)
f2 (ppm)
f1 (ppm)

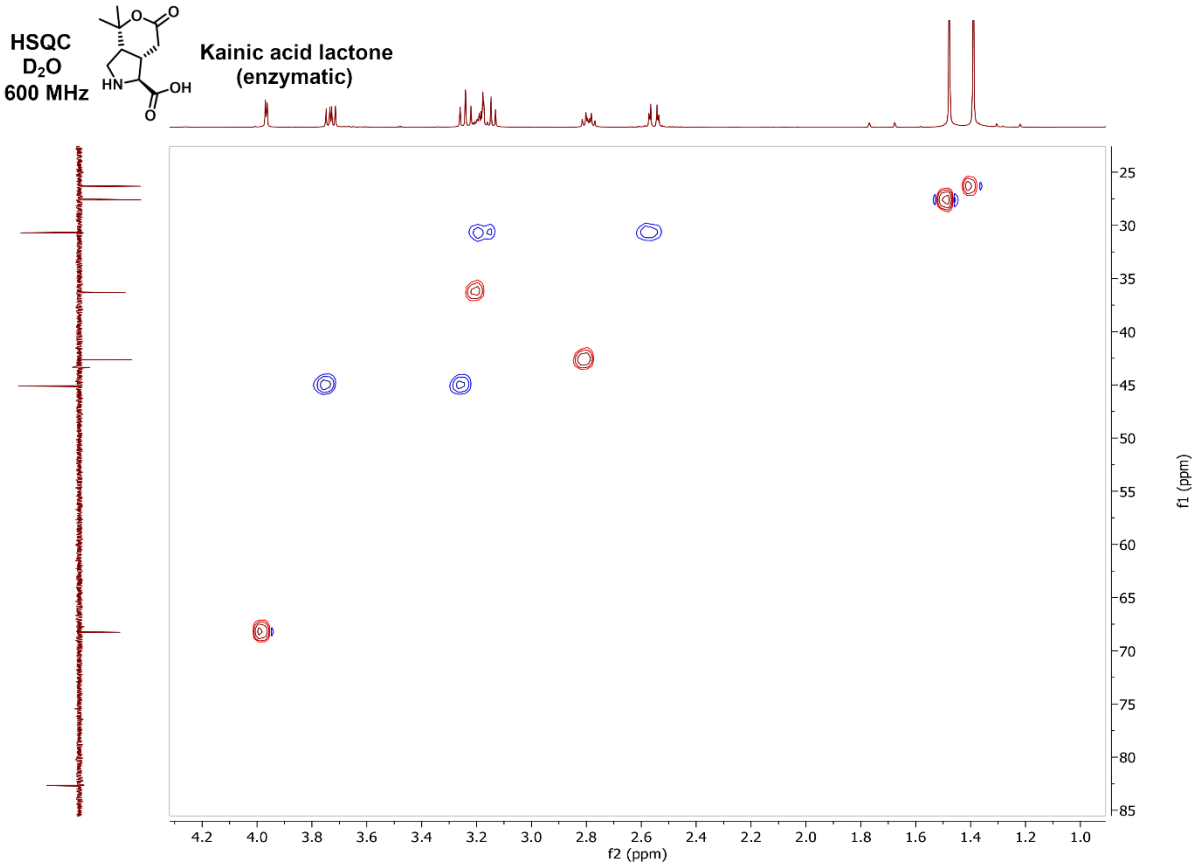
Kainic acid lactone (GfKabC product)

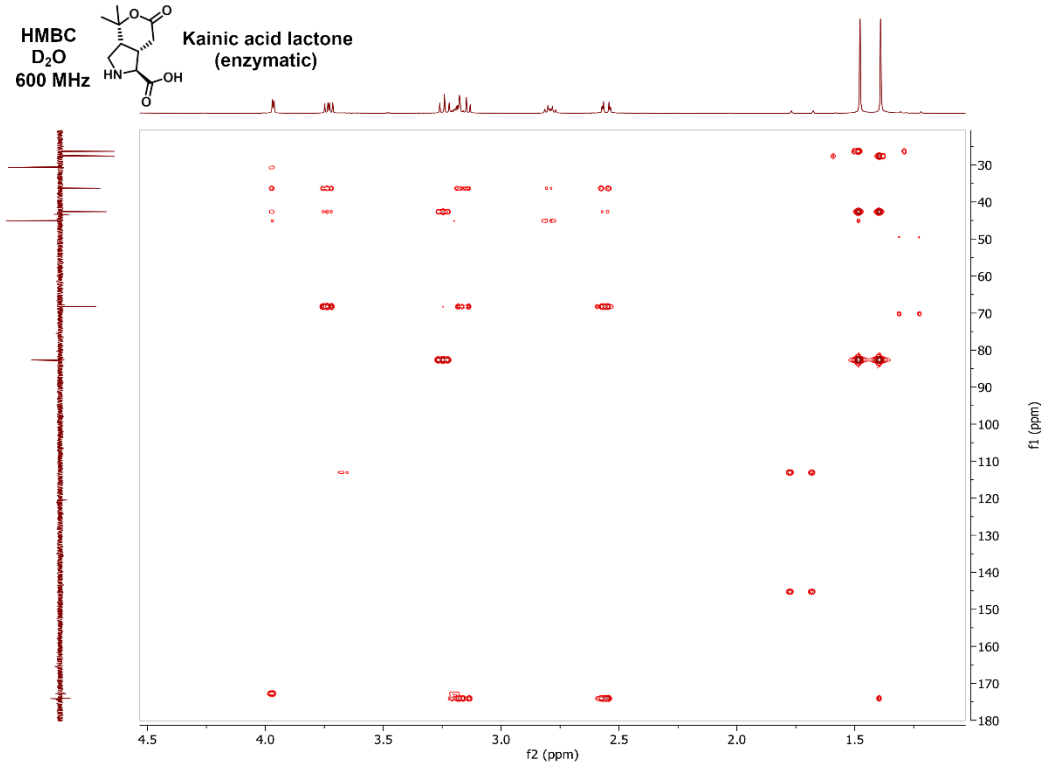


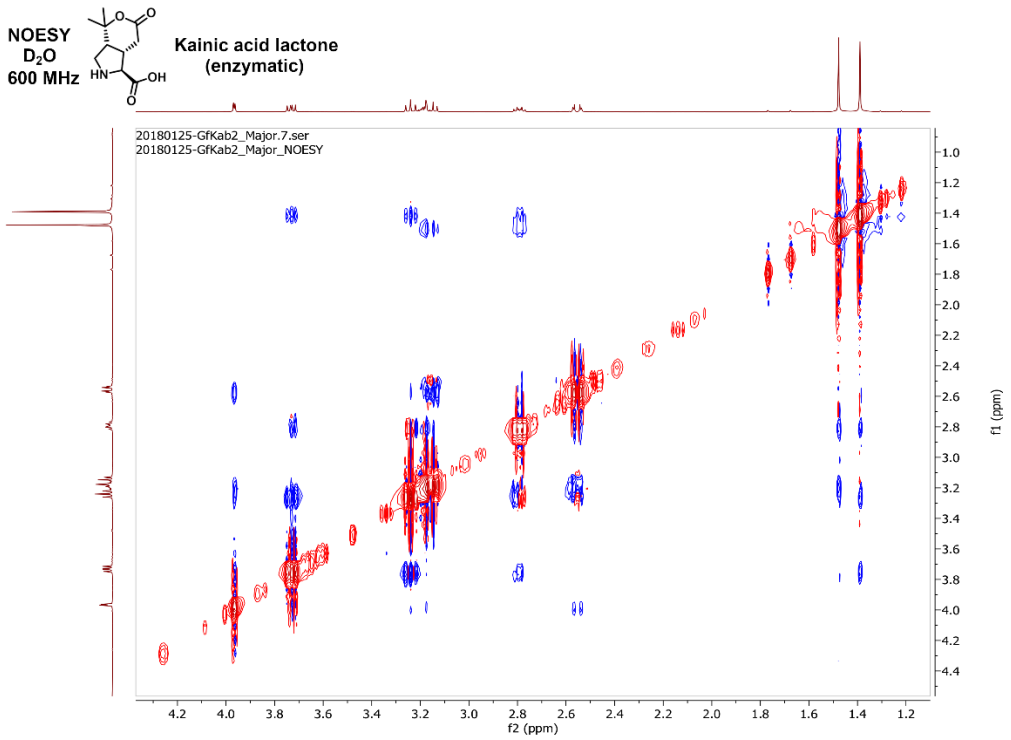



DEPT ^{13}C NMR
D_2O
150 MHz
197.56
174.10
174.10
172.76
82.66
68.26
45.09
42.63
36.33
30.70
27.61
26.32
Kainic acid lactone
(enzymatic)
f1 (ppm)


COSY
D_2O
600 MHz
Kainic acid lactone
(enzymatic)
f1 (ppm)
f2 (ppm)


HSQC
D2O
600 MHz
Kainic acid lactone
(enzymatic)
f2 (ppm)
f1 (ppm)


HMBC
D2O
600 MHz
Kainic acid lactone
(enzymatic)
f2 (ppm)
f1 (ppm)


NOESY
D2O
600 MHz
Kainic acid lactone
(enzymatic)
20180125-GfKab2_Major.7.ser
20180125-GfKab2_Major_NOESY
f2 (ppm)
f1 (ppm)

Supplementary References

[1] T. P. Michael, F. Jupe, F. Bemm, S. T. Motley, J. P. Sandoval, C. Lanz, O. Loudet, D. Weigel, J. R. Ecker, *Nat. Commun.* **2018**, *9*, 541.

[2] H. Li, *Bioinformatics* **2018**, *34*, 3094–100.

[3] H. Li, *Bioinformatics* **2016**, *32*, 2103–10.

[4] R. R. Wick, M. B. Schultz, J. Zobel, K. E. Holt, *Bioinformatics* **2015**, *31*, 3350–2.

[5] R. Vaser, I. Sović, N. Nagarajan, M. Šikić, *Genome Res.* **2017**, *27*, 737–46.

[6] B. J. Walker, T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C. A. Cuomo, Q. Zeng, J. Wortman, S. K. Young, et al., *PLoS One* **2014**, *9*, e112963.

[7] K. Katoh, J. Rozewicki, K. D. Yamada, *Brief. Bioinform.* **2017**, 1–7.

[8] W. P. Maddison, D. R. Maddison, “Mesquite: a modular system for evolutionary analysis,” **2018**.

[9] D. Silvestro, I. Michalak, *Org. Divers. Evol.* **2012**, *12*, 335–7.

[10] J. K. Brunson, S. M. K. McKinnie, J. R. Chekan, J. P. McCrow, Z. D. Miles, E. M. Bertrand, V. A. Bielinski, H. Luhavaya, M. Oborník, G. J. Smith, et al., *Science* **2018**, *361*, 1356–8.

[11] A. B. Woodside, Z. Huang, C. D. Poulter, *Org. Synth.* **1988**, *66*, 211–6.

[12] G. W. Saunders, C. Jackson, E. D. Salomaki, *Mol. Phylogenet. Evol.* **2018**, *119*, 151–9.

[13] N. Matasci, L. Hung, Z. Yan, E. J. Carpenter, N. J. Wickett, S. Mirarab, N. Nguyen, T. Warnow, S. Ayyampalayam, M. Barker, et al., *Gigascience* **2014**, *3*, 17.

[14] J. Janouskovec, A. Horák, M. Oborník, J. Lukes, P. J. Keeling, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 10949–54.

[15] C. Bowler, A. E. Allen, J. H. Badger, J. Grimwood, K. Jabbari, A. Kuo, U. Maheswari, C. Martens, F. Maumus, R. P. Otillar, et al., *Nature* **2008**, *456*, 239–44.

[16] N. P. Dunham, W.-C. Chang, A. J. Mitchell, R. J. Martinie, B. Zhang, J. A. Bergman, L. J. Rajakovich, B. Wang, A. Silakov, C. Krebs, et al., *J. Am. Chem. Soc.* **2018**, *140*, 7116–26.

[17] R. Calaf, A. Barlatier, D. Garçon, G. Balansard, M. Pellegrini, J. Reynaud, *J. Appl. Phycol.* **1989**, *1*, 257–66.

[18] O. Goldberg, A. Luini, V. I. Teichberg, *J. Med. Chem.* **1983**, *26*, 39–42.