ACS chemical biology

pubs.acs.org/acschemicalbiology

Articles

Genetic and Biochemical Reconstitution of Bromoform Biosynthesis in *Asparagopsis* Lends Insights into Seaweed Reactive Oxygen Species Enzymology

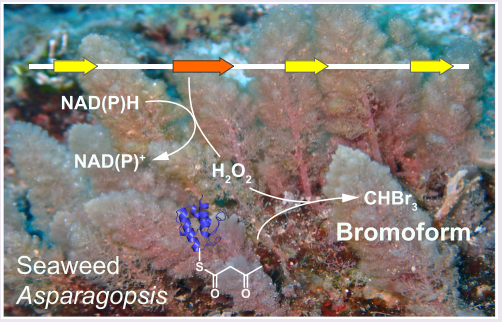
Hem R. Thapa, Zhenjian Lin, Dongqi Yi, Jennifer E. Smith, Eric W. Schmidt, and Vinayak Agarwal*

Cite This: *ACS Chem. Biol.* 2020, 15, 1662−1670

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Marine macroalgae, seaweeds, are exceptionally prolific producers of halogenated natural products. Biosynthesis of halogenated molecules in seaweeds is inextricably linked to reactive oxygen species (ROS) signaling as hydrogen peroxide serves as a substrate for haloperoxidase enzymes that participate in the construction these halogenated molecules. Here, using red macroalga *Asparagopsis taxiformis*, a prolific producer of the ozone depleting molecule bromoform, we provide the discovery and biochemical characterization of a ROS-producing NAD(P)H oxidase from seaweeds. This discovery was enabled by our sequencing of *Asparagopsis* genomes, in which we find the gene encoding the ROS-producing enzyme to be clustered with genes encoding bromoform-producing haloperoxidases. Biochemical reconstitution of haloperoxidase activities establishes that fatty acid biosynthesis can provide viable hydrocarbon substrates for bromoform production. The ROS production haloperoxidase enzymology that we describe here advances seaweed biology and biochemistry by providing the molecular basis for decades worth of physiological observations in ROS and halogenated natural product biosyntheses.



INTRODUCTION

Halogen-mediated ozone depletion has traditionally been ascribed to man-made halogenated pollutants such as chlorofluorocarbons. Photolytic degradation of these molecules produces reactive halogen species that degrade ozone.[1] However, the atmospheric halogen flux, particularly that for bromine, cannot be satisfied by anthropogenic pollutants alone.[2] Measurements and simulations posit that biogenic processes contribute a large fraction, up to 30%, of atmospheric bromine.[3,4] Of the biogenetically derived brominated pollutants, bromoform ($CHBr_3$), due to its high bromine content, short half-life, and photolytic lability, is the major contributor to ozone depletion. The tropical ocean is the primary source of atmospheric bromoform and oceanic bromoform emission is estimated to be on the order of several hundred thousand metric tons per year.[1,5]

Marine macroalgae, such as seaweeds, are the primary sources of atmospheric bromoform.[6,7] Bromoform production also extends to phytoplankton.[7−9] Work by Hager and others suggested that haloperoxidases, a class of halogenating enzymes widely detected in seaweeds, were involved in the production of bromoform.[10−14] Haloperoxidases catalyze the conversion of halide anions to hypohalous acid with hydrogen peroxide acting as the oxidant (eq 1). Hypohalous acid delivers electrophilic halenium ion, which halogenates hydrocarbon substrates to furnish natural products such as bromoform (eq 2). The requirement for hydrogen peroxide in this scheme links halogenated natural product biosynthesis in seaweeds to the enzymatic production of reactive oxygen species (ROS).

$$Br^- + H_2O_2 \xrightarrow[\text{peroxidase}]{} HOBr + OH^- \quad (1)$$

$$3HOBr + [CH] \xrightarrow[\text{peroxidase}]{} CHBr_3 + 3OH^- \quad (2)$$

Seaweeds are extraordinarily prolific in their production of haloperoxidase-mediated halogenated natural products, and these specialized molecules play formative roles in shaping biotic interactions and in marine chemical ecology. Fundamental gaps in haloperoxidase biochemistry remain. Principal among these is the source of hydrogen peroxide (eq 1). To date, a hydrogen peroxide-producing enzyme has not been identified from seaweeds.[15] This lack of insight also begets knowledge gaps in basic seaweed biology. In seaweeds, just like other eukaryotic systems, the production of reactive oxygen species (ROS), such as hydrogen peroxide is the primary response to biotic and abiotic stress.[16,17] Lack of identification of hydrogen peroxide forming enzymes, in turn, precludes

Received: April 17, 2020
Accepted: May 26, 2020
Published: May 26, 2020

https://dx.doi.org/10.1021/acschembio.0c00299
ACS Chem. Biol. 2020, 15, 1662−1670

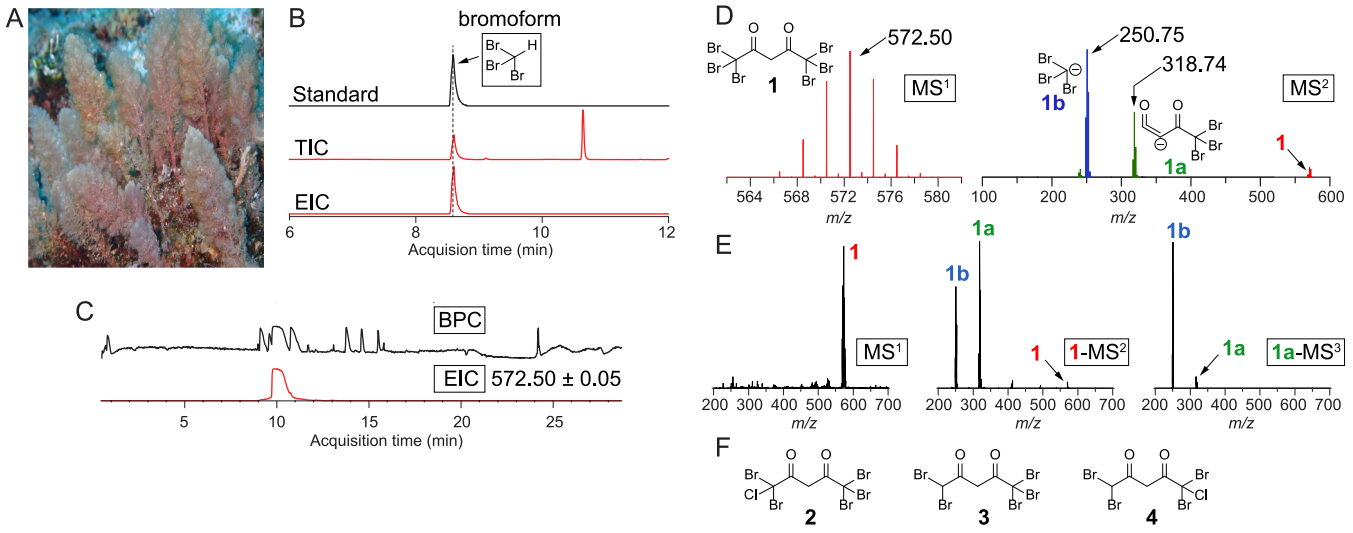


Figure 1. Marine macroalga *Asparagopsis taxiformis*. (A) Morphology of macroalgae *A. taxiformis* (Guam, 2015). (B) GC-MS chromatograms, from top to bottom, bromoform standard, *A. taxiformis* extract total ion chromatogram (TIC), and extracted ion chromatogram (EIC) corresponding to bromoform. (C) Organic extract of *A. taxiformis* analyzed by LC-MS. Base peak chromatogram (BPC) and EIC for m/z 572.49 ± 0.05 Da corresponding to the most abundant $[M - H]^{1-}$ isotopic ion for **1**. (D) MS^1 spectra for **1** showing the characteristic hexabrominated isotopic distribution corresponding to the molecular formula $[C_5H_2Br_6O_2]^{1-}$. MS^2 fragmentation spectra for **1** demonstrating product ions with inferred chemical structures **1a** and **1b**, as shown. The m/z values for the most abundant isotopic ions for **1a** and **1b** are labeled. (E) Ion-trap mass spectrometry-based MS^n characterization of **1**. (left) MS^1 spectra obtained by direct injection of algal *A. taxiformis* extract to an ion-trap mass spectrometer followed by isolation of cluster of ions corresponding to **1**. (middle) MS^2 spectra for **1** showing fragments identical to those observed in panel D. (right) The MS^2 ion corresponding to **1a** was isolated for MS^3 fragmentation leading to the production of the **1b** product ion, demonstrating that for **1a**, the three Br atoms each are localized to a single carbon atom only ($-CBr_3$). (F) Chemical structures of asymmetrical polyhalogenated 2,4-diones **2**–**4**.

characterization of the molecular circuitry underlying stress response and other ROS-mediated signaling in seaweeds.

The second principal gap is the complete lack of molecular-level description of the genetic localization and transcriptional status of genes encoding halogenated natural product biosynthesis within seaweed genomes. Without knowledge of their identities, hydrogen peroxide-producing enzymes are obviously bereft of such characterization. For haloperoxidases, the principal confounding factor is that seaweed genomes bear multiple haloperoxidase genes. In the absence of gene expression data, it is often unclear which specific enzyme is principally responsible for observed biosynthetic activity. In the event that a recombinant haloperoxidase is employed in an *in vitro* biochemical assay to reconstitute halogenation activity, it potentially strips away the molecular context in which the enzyme is produced, particularly as it relates to gene expression profile and proteomic abundance in the native organism. Only a handful of seaweed genome sequences are available, each of which is on the order of 100 Mb and contains extensive repeating and transposable elements.[18−20] As such, a molecular biological characterization of halogenated natural product biosynthesis in seaweeds is undeveloped.

The third major gap, particularly as it pertains to the biosynthesis of bromoform, is the lack of insight into the nature of the hydrocarbon substrate (eq 2). Prior studies afforded little rationale for organic substrates chosen in these studies for reconstituting bromoform production.[10,11] Dissolved organic material in the oceanic water column serves as a poor haloperoxidase substrate.[21,22] Confounding this further is the substrate promiscuity of haloperoxidases. Without metabolomic or genetic rationale, all studies thus far were restricted to employing biomimetic substrates for bromoform biosynthesis.

Here, we provide a genetic and biochemical characterization of bromoform production by the most prolific natural source of this pollutant: the seaweed *Asparagopsis taxiformis*. Together with its sister lineage *Asparagopsis armata*, *A. taxiformis* is globally distributed across the (sub)tropical and temperate coastlines and is receiving increasing attention as a feed additive to reduce methane production by dairy cattle.[23] Among biological sources, bromoform production levels produced by *Asparagopsis* are orders of magnitude higher than any other seaweed or phytoplankton source, up to 5% of algal dry weight can be bromoform, implying a dedicated bromoform biosynthetic enzymology that is not reliant on promiscuous reactivity of haloperoxidase enzymes alone. Employing a combination of metabolomics, genomics, and transcriptomics and enzymology, we uncover the molecular identities of the enzymes participating in hydrogen peroxide and bromoform production and reconstitute their respective biochemical activities. Our findings now set the stage for the discovery and characterization of bromoform biogenetic pathways in other oceanic sources.

RESULTS AND DISCUSSION

***Asparagopsis* Metabolomics.** Seaweed *A. taxiformis* samples used in this study were collected from Guam, Hawaii, Fiji, and the Florida Keys (Figure 1A). Samples were also collected in southern California and cultivated in the laboratory in sterilized seawater. Genomic DNA was extracted from all samples and barcoded by Sanger sequencing of PCR amplicons corresponding to the v4−v5 region of the 18S rRNA gene and the mitochondrial *cox2* and *cox3* intergenic region.[24] The 18S rRNA gene sequences confirmed the assignment of all samples as *A. taxiformis*. The *cox2*−*cox3* intergenic spacer sequences were used for phylogenetic assignment to reveal that

https://dx.doi.org/10.1021/acschembio.0c00299
ACS Chem. Biol. 2020, 15, 1662−1670

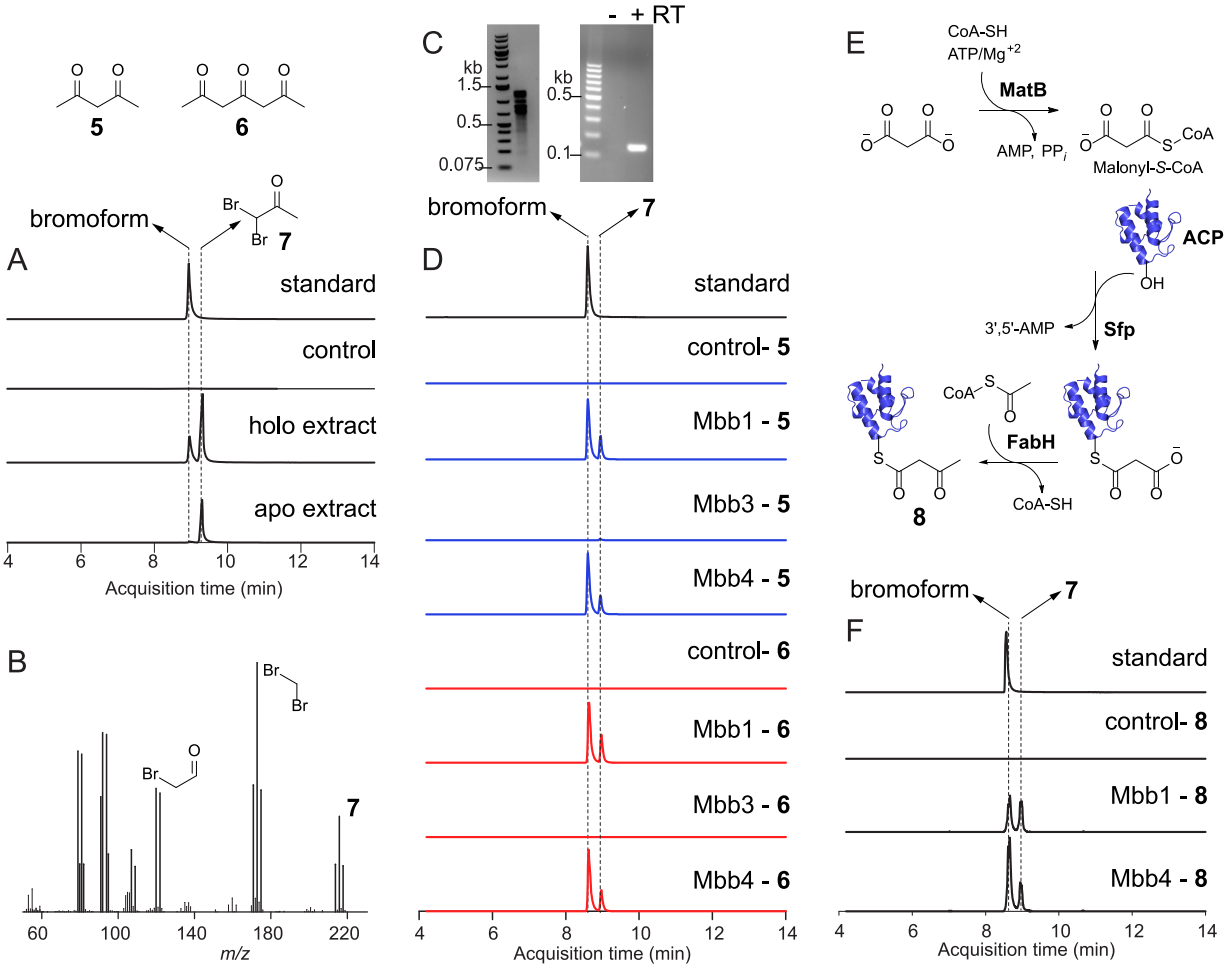


Figure 2. Enzymatic production of bromoform. (A) GC-MS chromatograms showing biosynthesis of bromoform using *A. taxiformis* total protein extract as the enzyme source and **5** as the substrate. Total protein extract dialyzed against buffer containing vanadate reconstitutes production of bromoform (assay with holo-extract), while extract dialyzed with vanadate-chelating EDTA does not lead to production of bromoform (assay with apo-extract). Control reaction omits the total protein extract. (B) Structural annotation of **7** based on GC-MS fragmentation spectra. (C) RNA isolation and verification of lack of DNA contamination by RT-PCR of RuBisCO-encoding chloroplast *rbcL* gene for field-collected Guam sample. (D) GC-MS chromatograms showing production of bromoform and **7** by Mbb1 and Mbb4 enzymes using **5** and **6** as substrates. (E) Scheme for enzymatic production of **8**. (F) GC-MS chromatograms showing production of bromoform and **7** by Mbb1 and Mbb4 enzymes using **8** as the substrate. Data shown are representative example of three independent experiments.

Guam, California, and Fiji samples formed distinct outgroups when integrated with previously reported sequences (Figure S1).[25] Samples used further in this study, from Guam and California, possess high concentrations of bromoform, ranging from 1% to 5% w/w (dry weight basis), as judged by gas chromatography−mass spectrometry (GC-MS) analyses of underivatized organic extracts of dried biomass. Metabolite identity was verified by comparison to a commercial bromoform standard (Figures 1B and S2).

Inventorying *Asparagopsis* metabolomes has traditionally relied on GC-MS, which, in addition to polybromomethanes, has revealed a rich inventory of polybrominated ketones, lactones, acrylates, and acetates.[26,27] Here, we employed untargeted liquid chromatography−mass spectrometry (LC-MS) to interrogate the *A. taxiformis* metabolomes. Using LC-MS, we could detect an exceptionally abundant hexabrominated metabolite in the negative ionization mode that could not be dereplicated across natural product repositories, mass spectrometric databases, or a manual curation of previously described *Asparagopsis* metabolites (Figure 1C). High-resolution mass determination and the isotopic distribution characteristic of polybrominated molecules led to the assignment of the molecular formula as $C_5H_2Br_6O_2$, which implies two degrees of unsaturation (**1**, inferred structure shown in Figure 1D; $[M - H]^{1-}$ calcd 566.5082; found 566.5085). The molecule was labile and could not be isolated for spectroscopic structure elucidation. Hence, we relied on mass spectrometry for structural assignment. High-resolution time-of-flight (ToF) spectrometry revealed two major fragment ions, **1a** and **1b** (Figure 1D). MS^2 ion **1b** demonstrates that three of the six bromine atoms of **1** are located on a single carbon atom. To gain insight into the structure of the other MS^2 product ion, **1a**, we employed ion trap-MS^n fragmentation. MS^1 ions corresponding to **1** were isolated in a low-resolution ion trap mass spectrometer and fragmented to reveal **1a** and **1b** product ions, and the MS^2 fragment ion **1a** was isolated and further fragmented to reveal MS^3 product ion identical to **1b** (Figure 1E). Thus, for ion MS^2 fragment **1b**, the three bromine atoms are also located on a single carbon atom only (**1a** and **1b** inferred structures shown in Figure 1D). Progressing from the requirement of two degrees of unsaturation, we posit the structure of **1** to be a symmetrical 1,1,1,5,5,5-hexabromo-2,4-

https://dx.doi.org/10.1021/acschembio.0c00299
ACS Chem. Biol. 2020, 15, 1662−1670

dione. Compound **1** was not observed by GC-MS. In addition to **1**, other 1,5-polyhalogenated 2,4-dione natural products, such as **2**−**4** were also detected (Figure 1F). High resolution ToF fragmentation spectra for **2**−**4** lead to assignment of asymmetric MS^2 product ions (Figure S3). In addition to bromoform, the hydrolytic cleavage of molecules **1**−**4** rationalizes the production of other *Asparagopsis*-derived polybromomethanes such as dibromomethane and chlorodibromomethane.[7]

Enzymatic Reconstitution of Bromoform Production. Given the abundance of polybrominated 2,4-dione natural products in the *A. taxiformis* metabolome, we rationalized that a physiological substrate for bromoform production would be different from a β-keto acid such as 3-oxooctanoic acid that was used by Hager et al.[10] and that has now been generally accepted as the physiological substrate for bromoform production. Instead, we posit that the physiological substrate for bromoform production in *A. taxiformis* should resemble β,δ-polyones. Thus, we employed small molecule substrates pentane-2,4-dione (**5**, commercial, Figure 2) and heptane-2,4,6-trione (**6**, synthesized, Figures S5 and S6) to reconstitute the enzymatic production of bromoform.

To reconstitute bromoform production, we generated a total protein extract from *A. taxiformis* collected from Guam. The protein extract was dialyzed against buffer containing sodium orthovanadate to yield "holo-extract" and against buffer containing vanadate-chelating EDTA to yield "apo-extract". Protein extracts thus prepared were used as catalysts for *in vitro* biochemical assays. Control reactions omitted the catalyst. Organic extracts of reactions were analyzed by GC-MS. In the presence of bromide and the hydrogen peroxide, only the holo-extract led to production of bromoform (Figure 2A). In addition to bromoform, the production of 1,1-dibromoacetone (**7**), a previously reported *Asparagopsis* metabolite, was also observed. Structural assignment for **7** is based on mass spectral fragmentation annotation (Figure 2B). The apo-extract led to production of **7** but not bromoform. The requirement for vanadate lends support that vanadium-dependent haloperoxidases (VHPOs), halogenating enzymes widely detected in other seaweeds, are involved in bromoform biosynthesis.[6,13,14]

To uncover molecular identities of *Asparagopsis* VHPOs, we undertook a transcriptome sequencing and mining approach. First, we extracted total RNA from the Guam and California *A. taxiformis* samples. We verified that the RNA preparations were free of DNA contamination by reverse transcription-PCR (RT-PCR) of a 150 bp fragment of the RuBisCO large subunit-encoding chloroplast *rbcL* gene.[28] The *rbcL* amplicon was observed only in the presence of the reverse transcriptase in the RT-PCR reaction (Figures 2C and S4). Subsequent subtraction of rRNA and poly-dT based mRNA enrichment allowed access to cDNA. The cDNA of the Guam *A. taxiformis* sample was sequenced on an Illumina HiSeq 2000 sequencer with a ∼350 bp inserts and 125 bp paired end reads. The cDNA of California sample was packaged into libraries that were sequenced using 150 bp paired end reads on an Illumina NovaSeq sequencer with a ∼450 bp insert size. After adaptor trimming and quality control, 98,748,611 and 44,450,663 reads were recovered for *A. taxiformis* transcriptomes from Guam and California samples, respectively, that were assembled to reveal 151,422 and 20,789 contigs (Table S1). We mined the transcriptomes using marine macroalgal VHPO sequences as queries to reveal three VHPO ORFs that were highly similar between the Guam and California transcriptomes (Table S2). Sequence homology to other algal VHPOs identified all residues that are involved in binding the vanadate cofactor (Figure S7).[15]

The three VHPOs, named Mbb1, Mbb3, and Mbb4 (nomenclature defined below), identified in the Guam transcriptome were cloned from cDNA into plasmid vectors for expression in *Escherichia coli* (Table S3). Purified recombinant enzymes were assayed using monochlorodimedone as a substrate, which offers a continuous spectroscopic readout for the halogenating activity. All three VHPOs were specific for the utilization of bromide; no chlorinating activity was observed (Figure S8).

We next assayed activities of Mbb1, -3, and -4 enzymes using **5** and **6** as substrates. For both substrates, Mbb1 and Mbb4, but not Mbb3, demonstrated production of bromoform and **7** (Figure 2D). Consumption of substrate **5** could be observed for Mbb1 and Mbb4, but not for Mbb3 (Figure S9). Extended reaction times with Mbb3 also did not lead to bromoform production. Reaction products observed using purified enzyme catalysts are comparable to the product profile observed with the total protein holo-extract prepared from the algal biomass (Figure 2A). All three enzymes demonstrated comparable production of monobromoacetone, a previously described *Asparagopsis* natural product (Figure S10).

Bromoform Can Be Derived from Fatty Acid Biosynthesis. Next, we intended to identify a possible biogenetic route that could furnish a VHPO substrate for the production of bromoform. Compounds **5** and **6** are reminiscent of fatty acid biosynthetic intermediates. Fatty acid biosynthesis in algae is localized to the plastid. Gene sequences for fatty acid biosynthetic enzymes were available from the previously reported *A. taxiformis* plastid genome.[28] To query whether the fatty acid biosynthetic intermediates could serve as substrates for bromoform production, first, we purified recombinant *A. taxiformis* plastid acyl carrier protein (ACP) and ketosynthase, FabH. Malonyl-coenzyme A (malonyl-CoA) was synthesized by the ATP-dependent condensation of malonate with CoA-SH by the enzyme MatB and used *in situ* to modify the purified *A. taxiformis* ACP to malonyl-*S*-ACP by the phosphopantetheinyl transferase Sfp (Figure 2E).[29,30] Stoichiometric production of malonyl-*S*-ACP was verified by the mass spectrometric phosphopantetheine ejection assay.[31] Malonyl-*S*-ACP thus synthesized was incubated with acetyl-CoA and *A. taxiformis* FabH to synthesize acetoacetyl-*S*-ACP (**8**). Formation of **8** was verified by mass spectrometry, and the product was purified chromatographically (Figures S11−13). Incubation of **8** with Mbb1 and Mbb4 led to production of bromoform and **7** (Figure 2F).

In light of the distinguishingly high amount of bromoform produced by *Asparagopsis* as compared to other seaweeds, it is likely that the primary hydrocarbon substrates employed by *Asparagopsis* for bromoform production are endogenously produced, as opposed to the utilization of dissolved organic matter in the water column, which serves as a poor substrate pool for bromoform production.[21,22] Contrary to the conventional biomimetic experiments that employ β-keto acid substrates for bromoform biosynthesis, the presence of diones **1**−**4** in the *Asparagopsis* metabolome lead us to posit that haloform-like reaction using methyl ketone substrates that are likely derived from fatty acid biosynthesis underlies bromoform biosynthesis. Thioesterified methyl ketones, analogous to **8**, are generated by type III polyketide ketosynthases, three of which are found in the *Asparagopsis* nuclear genome. Thus, several

https://dx.doi.org/10.1021/acschembio.0c00299
ACS Chem. Biol. 2020, 15, 1662−1670

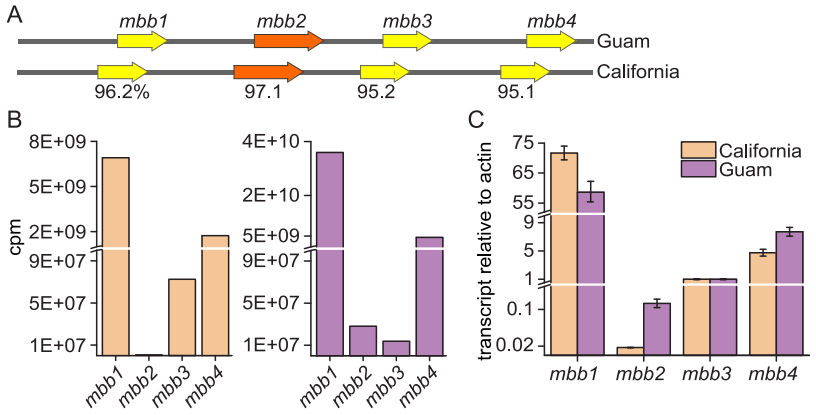


Figure 3. Bromoform biosynthetic genes. (A) The *A. taxiformis mbb* loci from Guam and California samples. Nucleotide identity between Guam and California sequences is listed for each gene. (B) Computationally determined counts per million (cpm) values determined for each *mbb* gene in the laboratory cultivated (California, left) and field collected (Guam, right) transcriptomes. (C) Experimentally determined transcript levels for *mbb* genes relative to actin in respective samples. Average values and standard deviation error from three independent experiments are plotted.

different hydrocarbon substrates, all containing terminal methyl ketones, might be involved in bromoform production in algae in line with the broad substrate tolerance of bromoperoxidases Mbb1 and Mbb4. The bromoform biosynthesis scheme that we describe here also recovers the production of bromoacetones, characteristic *Asparagopsis* natural products that represent intermediates or shunt metabolites generated along the bromoform biosynthetic route.

Algal Genome Sequencing. Having verified that Mbb1 and Mbb4 haloperoxidases biosynthesize bromoform, we proceeded to query their relative positioning in the algal genome. To this effect, we sequenced and assembled metagenomes of *A. taxiformis* from Guam and California using 150 bp paired end Illumina NovaSeq sequencing platform (Table S1). The draft genome was retrieved by removing the prokaryotic contigs using Autometa to yield 46.5 Mb and 83.6 Mb data, respectively.[32] The draft genomes reveal that all three VHPOs, Mbb1, Mbb3, and Mbb4, are present in genetic neighborhood of each other (Figure 3A). We have named this genetic neighborhood as the marine bromoform biosynthesis locus (*mbb* locus). The *mbb* locus is present in the nuclear genome, and all *mbb* genes are devoid of introns, an observation consistent with other red macroalgal genomes.[18,19]

With draft *Asparagopsis* genomes in hand, together with transcriptomic data that was derived from the same samples, we could query the relative transcript abundances for the *mbb* genes. Transcriptomic read recruitment to the *mbb* locus revealed that in the California sample, gene *mbb1* is transcribed at a much higher level as compared, in decreasing order, to *mbb4*, *mbb3*, and *mbb2*. This surprising relative transcript abundance trend was conserved in the field collected Guam sample as well; however, gene *mbb2* was expressed at a higher level (Figure 3B). To verify that the differential gene expression in the *mbb* loci was not due to cDNA sequencing and assembly artifacts, we queried gene expression using quantitative real time-PCR (qRT-PCR). The qRT-PCR data were normalized against actin, a constitutively expressed housekeeping gene, thus allowing for a between-sample comparison (Table S4).[33] Relative to actin, gene *mbb1* is transcribed 58−71-fold higher in *Asparagopsis*, while the gene encoding VHPO that does not produce bromoform, *mbb3*, is expressed at a much lower level. The experimental data were in agreement with the computational findings, in that the *mbb1* gene is transcribed at a much higher level than *mbb2*−*4* and that the expression of the *mbb2* gene differs between the California (laboratory cultivated) and Guam (field collected) *A. taxiformis* samples (Figure 3C).

ROS Production by Mbb2. Next, we interrogated the biochemical activity of Mbb2. Sequence similarity led to identification of motifs for binding NAD(P)H, flavin, and heme in the Mbb2 transmembrane peptide (Figures S14 and S15), hallmarks of an NAD(P)H-oxidase (NOX). NOXs are membrane bound enzymes that catalyze the transfer of electrons from the electron donor, NAD(P)H, via flavin and heme cofactors, to molecular oxygen to generate ROS such as hydrogen peroxide and superoxide anion (dismutated to hydrogen peroxide) (Figure 4A). However, the inferred architecture of Mbb2 was distinct from all other NOXs known thus far, in that the relative positioning of the membrane-spanning heme-binding/oxygen reducing domain and the flavin-binding/NAD(P)H-oxidizing cytoplasmic domain is uniquely different from previously characterized mammalian, yeast, and cyanobacterial NOXs (Figure S14).

Our attempts to express *mbb2* in bacteria and yeast were unsuccessful, despite employing synthetic DNA clones optimized for expression in the respective heterologous hosts. We thus searched for Mbb2 homologues in other algal genomes. The genome of the bromoform producing marine alga *Chondrus crispus*, colloquially referred to as "Irish moss", has been sequenced.[7,18] A Mbb2 homologue, annotated as a hypothetical protein with 63% similarity between protein primary sequences and identical domain organization (Figures S14 and S15), could be detected in the *C. crispus* genome. Synthetic cDNA optimized for expression in yeast for the *C. crispus* Mbb2 homologue, henceforth referred to as *Cc*Mbb2, was packaged into a plasmid vector for heterologous expression in *Saccharomyces cerevisiae*. In the *S. cerevisiae* expression strain that we employed here, the yeast ROS-producing NOX encoding gene, *yno1*, had been deleted.[34] In the Δ*yno1* genetic background, reconstituting the expression of *yno1* using an identical plasmid vector system as *Ccmbb2* serves as a positive control to evaluate the biochemical activity of *Cc*Mbb2, while an empty plasmid vector serves as the negative control. In this experiment, the dye dihydroethidium (DHE) captures ROS to deliver a fluorescent spectroscopic readout. Upon induction of

https://dx.doi.org/10.1021/acschembio.0c00299
ACS Chem. Biol. 2020, 15, 1662−1670

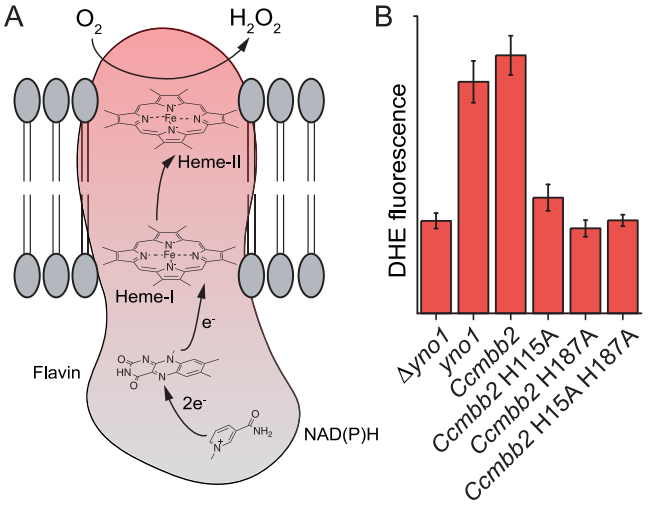


Figure 4. Hydrogen peroxide formation by a macroalgal NOX. (A) Cartoon representation of the architecture of NOXs. Electrons are delivered by NAD(P)H that are then transferred to the flavin cofactor and then via two heme cofactors, referred here as heme-I and heme-II, to dioxygen. (B) DHE fluorescence for Δ*yno1 S. cerevisiae* strain transformed with, from left to right, empty plasmid vector, vector carrying the *yno1* gene, vector carrying the *Ccmbb2* gene, vector carrying the *Ccmbb2* gene with H115A mutation, vector carrying the *Ccmbb2* gene with H187A mutation, and vector carrying the *Ccmbb2* gene with H115A, H187A double mutation. The H115 and H187 side chain imidazole heterocycles coordinate the heme-I iron atom. Average values and standard deviation from three independent measurements are plotted.

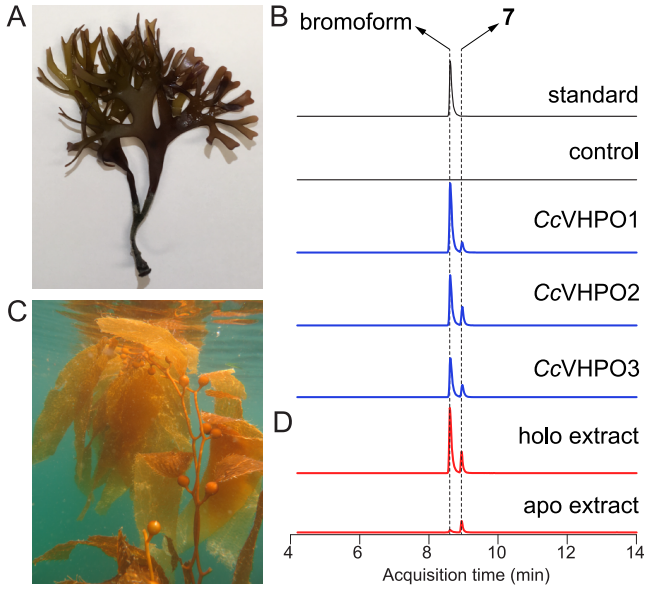


Figure 5. Bromoform production by seaweeds. (A) Morphology of red macroalga *C. crispus*. (B) GC-MS chromatograms showing *in vitro* production of bromoform and **7** starting from **5** as the substrate by purified *C. crispus*-derived enzymes *Cc*VHPO1−3 (in blue). Control reaction omits enzyme in the assay. (C) Morphology of brown macroalga *M. pyrifera*. (D) GC-MS chromatograms showing production of bromoform and **7** by *M. pyrifera* total protein extracts (in red) using **5** as the substrate. Data shown are representative of three independent experiments.

protein expression via the inducible galactose promoter, both Yno1 and *Cc*Mbb2 demonstrated an increase in the production of ROS (Figure 4B). It should be noted that deletion of *yno1* does not abolish ROS production in *S. cerevisiae*, as has been observed previously, due to the presence of additional as yet uncharacterized NOX encoding genes in the yeast genome.[34]

Next, based on sequence homology, we identified two histidine residues that coordinate the iron atom in the heme-I cofactor involved in conducting electrons from the flavin cofactor to the heme-II cofactor, which is the site for dioxygen reduction (Figure 4A, S15). Mutation of either or both of these histidine residues causes disruption of the electron transport chain leading to loss of NOX activity, as monitored by DHE fluorescence retreating to levels observed for the negative control (Figure 4B).

Conservation of Bromoform Enzymology. While *A. taxiformis* is exceptionally prolific in its production of bromoform, it is not the only macroalgal source of bromoform. As described in the previous section, the seaweed *C. crispus* is a known bromoform producer (Figure 5A).[7] The *C. crispus* genome harbors five VHPO-encoding genes.[18] Of these five, optimized for expression in *E. coli*, synthetic DNA sequences corresponding to three *Cc*VHPOs with highest sequence homology to Mbb1, -3, and -4 were packaged into plasmid vectors. All three *Cc*VHPOs were specific for bromide incorporation and demonstrated *in vitro* production of bromoform and **7** using **5** as a substrate (Figures 5B and S16). Together with the experimental demonstration of ROS production by *Cc*Mbb2, the bromoform biosynthetic enzymology in *C. crispus* is expected to resemble that from *A. taxiformis*.

We next turned our attention to the brown macroalga *Macrocystis pyrifera*, the globally distributed oceanic giant kelp and a known producer of bromoform (Figure 5C).[35] A California coastal collection of *M. pyrifera* demonstrated the presence of bromoform, albeit at a much-reduced level compared with that for *A. taxiformis*; 1% to 5% of algal dry weight in *A. taxiformis* compared to 0.0003% in *M. pyrifera*, values in agreement with literature.[35] We prepared holo and apo total protein extracts from *M. pyrifera* by dialyzing against buffers with and without vanadate, as we had done previously for *A. taxiformis*. Only the holo-extract catalyzed the *in vitro* production of bromoform using **5** as the substrate (Figure 5D). While the molecular identities of the catalysts remain unknown in the absence of the as yet unavailable genomic data, bromoform production in *M. pyrifera* is also dependent on the activity of VHPO(s), in addition to previously reported halogenating activity of organic substrates.[36]

Seaweed ROS Enzymology and Stress Response. Bromoform production by seaweeds has been extensively documented.[7,14] Likewise, at the organismal level, ROS production by seaweeds in response to stress is known.[16] However, a genetic and molecular connection between these activities, as is necessitated by haloperoxidase biochemistry, had not been realized. Specifically, while hydrogen peroxide has traditionally been used to reconstitute the *in vitro* activity of haloperoxidases, endogenous enzymatic production of hydrogen peroxide in seaweeds had not been characterized. Although we could not reconstitute the *A. taxiformis* Mbb2 activity, we characterized a homologous NOX from a bromoform producing alga and demonstrated the production of hydrogen peroxide. Our discovery of the NOX activity of Mbb2 was guided by the clustering of the *mbb2* gene with haloperoxidase genes *mbb1* and *mbb3−4* in the *Asparagopsis* genome (Figure 3A). Homologues of Mbb2 from previously sequenced algal genomes were hiding in plain sight with no

https://dx.doi.org/10.1021/acschembio.0c00299
ACS Chem. Biol. 2020, 15, 1662−1670

bioinformatic or biochemical rationale to interrogate their NOX activities.[18,19] Similar clustered organization of genes, though it is the norm in prokaryotic natural product biosynthesis, has been observed in seaweeds only once before.[37] The mechanistic coupling of NOXs and haloperoxidases is akin to the dependence of flavin-dependent halogenases on flavin reductases[15] and of cytochrome P450s on ferredoxin/ferredoxin reductases.[38] Unlike bacteria that have formed the basis of much of our understanding of halogenation biochemistry, this coupling of haloperoxidases with NOXs has important consequences for seaweed biology, marine chemical ecology, and, as exemplified by bromoform, environmental health.

Our experimental RT-qPCR and computational transcriptome sequencing data converge to reveal that the gene *mbb2* is expressed at a lower level as compared to *mbb1* (Figure 3B,C). This might be due to highly efficient Mbb2 activity, toxicity associated with high expression of this membrane-localizing NOX as was observed for the yeast NOX Yno1,[34] or possibly compensation by other NOX enzymes in the algal genome. Indeed, *mbb2* is not the only NOX encoding gene transcribed in *A. taxiformis*. In addition to *mbb2*, we detect the expression of at least four other genes encoding NOXs (Figure S17). Each of these proteins is predicted to contain the transmembrane architecture with cytoplasmic NAD(P)H and flavin cofactor binding domains characteristic of NOXs (Figure S18). Why does the *Asparagopsis* genome encode multiple NOXs? Participation of different NOXs in different signaling cascades to respond to stress and regulate growth in terrestrial plants is well established and it seems that seaweeds too conserve this biological logic.[17] In a similar vein, it was known that seaweed genomes encode multiple haloperoxidases. It is likely that the presence of different haloperoxidases allows them to access different hydrocarbon substrate pools or that they are trafficked to different subcellular locations for bromoform production. However, clustering of multiple haloperoxidases together in a single genetic locus with varying levels of gene expression was surprising. While the rationale for the genetic clustering of haloperoxidase genes is not immediately clear, overall, the haloperoxidase gene expression levels correlate with the observed physiological activities, in that the expression levels of bromoform producing Mbb1 and Mbb4 haloperoxidases is significantly higher than that of the Mbb3, which does not produce bromoform (Figure 3B,C).

Asparagopsis accumulates an extraordinarily high tissue concentration of bromoform (1% to 5% of algal dry weight). Bromoform protects *Asparagopsis* against epiphytic bacterial colonization, and other haloperoxidase-mediated halogenated natural products have been postulated to serve similar defensive roles.[16,39] Our findings allow us to posit that bromoform and other halogenated molecules in marine algae are produced as a part of a ROS production and manipulation cascade, which converts hydrogen peroxide to defensive halogenated chemicals with longer half-lives. Interestingly, expression of the *mbb2* homologue in the bromoform producing macroalga *C. crispus*, *Ccmbb2*, was found to be upregulated in response to pathogen attack and bacterial immunogens.[40] A similar upregulation of NOX expression in response to bacterial challenge was observed for macroalga *Laurencia dendroidea*, a prolific source of VHPO-mediated halogenated natural product chemistry.[41] A NOX-mediated oxidative burst resulting in increased hydrogen peroxide production has been elicited in other bromoform producing macroalgae by bacterial immunogens as well.[16,42] Hydrogen peroxide can also be transduced to hypohalous acids by heme-dependent peroxidases. We detect multiple highly transcribed heme-dependent peroxidases in the *Asparagopsis* transcriptomes (Table S5). Our characterization of the first hydrogen peroxide-producing NOX serves as an initiation point to determine, at the molecular level, the ROS enzymology in seaweeds and the response of these enzymes to biotic and abiotic stressors.[16] Already, with a change in culture conditions, that is, between the field collected Guam sample and laboratory cultivated California sample, we can observe a rewiring of the NOX-encoding gene transcription levels (Figure S17).

Bromoform production is not limited to marine macroalgae. While macroalgae are indeed the dominant sources of bromoform in coastal environments, phytoplankton in the oceanic water column also produce bromoform with estimated equal contributions to the atmospheric bromoform flux as coastal macroalgae.[7,8] Other sources of bromoform may indeed exist in marine ecosystems, but our inventory of them is limited by access to biomass and analytical procedures. The discovery and experimental validation of the Mbb enzymes now provide for an alternative mechanism to mine for the presence of the bromoform biosynthetic potential in marine metagenomes. The contemporary marine environment is rapidly changing. Increasing eutrophication of the oceans and rising CO_2 levels are altering marine ecosystems, for example, by favoring massive phototrophic algal blooms. How will these changes affect the production of bromoform, and can we inform future activities that can mitigate the environmental exposure to this pollutant? We can begin to address these important questions only when the biological sources have been definitively identified and the biosynthetic processes fully understood.

MATERIALS AND METHODS

Materials and methods used in this study are described in the Supporting Information.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.0c00299.

Description of experimental methods and supplementary figures and tables (PDF)

Accession Codes

Algal 18S rRNA and intergenic *cox2*−*cox3* sequences have been deposited to GenBank (with accession numbers MN547332-6 and MN563722-6, respectively. The *mbb* loci from California and Guam samples have been deposited to GenBank with accession numbers MN966723 and MN893468, respectively. GenBank accession numbers for *nox1*−*4* from Guam sample are MN911451−MN911454 and for *nox1*−*4* from California sample are MN966724−MN966727. Whole genome shotgun projects have been deposited at GenBank under accession number PRJNA596236.

AUTHOR INFORMATION

Corresponding Author

Vinayak Agarwal − *School of Chemistry and Biochemistry and School of Biological Sciences, Georgia Institute of Technology,*

https://dx.doi.org/10.1021/acschembio.0c00299
ACS Chem. Biol. 2020, 15, 1662−1670

Atlanta, Georgia 30332, United States; orcid.org/0000-0002-2517-589X; Phone: 404-385-3798; Email: vagarwal@gatech.edu

Authors

Hem R. Thapa − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States;* orcid.org/0000-0003-3864-6141

Zhenjian Lin − *Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States*

Dongqi Yi − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States*

Jennifer E. Smith − *Center for Marine Biodiversity and Conservation, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, United States*

Eric W. Schmidt − *Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States;* orcid.org/0000-0001-5839-694X

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.0c00299

Author Contributions

H.R.T. and V.A. designed the research; H.R.T. and Z.L. performed the research; D.Y., J.E.S., and E.W.S. contributed new reagents and analytical tools; H.R.T., Z.L., and V.A. analyzed the data; H.R.T. and V.A. wrote the paper with contributions from all authors.

Notes

The authors declare no competing financial interest.

High-resolution LC-MS data set has been deposited to MassIVE repository (ID no. MSV000084456).

■ ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (ES026620 to V.A.; GM122521 to E.W.S.) and an Alfred P. Sloan Foundation fellowship (to V.A.). The authors thank P. Jensen (University of California, San Diego), J. Kubanek, (Georgia Institute of Technology), V. Paul (Smithsonian Marine Station at Ft. Pierce, Florida), and J. Biggs (University of Guam Marine Station) for algae samples from California, Fiji, Florida Keys, and Guam, respectively. The authors also thank A. Reddi and N. Garg (Georgia Institute of Technology) for yeast expression strains and helpful discussions and for the use of Bruker ImpactII UHR-QqToF mass spectrometer. Clone for MatB gene was thankfully provided by A. Keatinge-Clay (The University of Texas at Austin).

■ REFERENCES

(1) Simpson, W. R., Brown, S. S., Saiz-Lopez, A., Thornton, J. A., and von Glasow, R. (2015) Tropospheric halogen chemistry: sources, cycling, and impacts. *Chem. Rev. 115* (10), 4035−4062.

(2) Salawitch, R. J. (2006) Biogenic bromine. *Nature 439* (7074), 275−277.

(3) Navarro, M. A., Atlas, E. L., Saiz-Lopez, A., Rodriguez-Lloveras, X., Kinnison, D. E., Lamarque, J.-F., Tilmes, S., Filus, M., Harris, N. R. P., Meneguz, E., Ashfold, M. J., Manning, A. J., Cuevas, C. A., Schauffler, S. M., and Donets, V. (2015) Airborne measurements of organic bromine compounds in the Pacific tropical tropopause layer. *Proc. Natl. Acad. Sci. U. S. A. 112* (45), 13789.

(4) Tegtmeier, S., Ziska, F., Pisso, I., Quack, B., Velders, G. J. M., Yang, X., and Krüger, K. (2015) Oceanic bromoform emissions weighted by their ozone depletion potential. *Atmos. Chem. Phys. 15* (23), 13647−13663.

(5) Warwick, N. J., Pyle, J. A., Carver, G. D., Yang, X., Savage, N. H., O'Connor, F. M., and Cox, R. A. (2006) Global modeling of biogenic bromocarbons. *J. Geophys. Res. 111* (D24), D24305.

(6) Paul, C., and Pohnert, G. (2011) Production and role of volatile halogenated compounds from marine algae. *Nat. Prod. Rep. 28* (2), 186−195.

(7) Carpenter, L. J., and Liss, P. S. (2000) On temperate sources of bromoform and other reactive organic bromine gases. *J. Geophys. Res. 105* (D16), 20539−20547.

(8) Tokarczyk, R., and Moore, R. M. (1994) Production of volatile organohalogens by phytoplankton cultures. *Geophys. Res. Lett. 21* (4), 285−288.

(9) Moore, R. M., Webb, M., Tokarczyk, R., and Wever, R. (1996) Bromoperoxidase and iodoperoxidase enzymes and production of halogenated methanes in marine diatom cultures. *J. Geophys. Res. 101* (C9), 20899−20908.

(10) Theiler, R., Cook, J. C., Hager, L. P., and Siuda, J. F. (1978) Halohydrocarbon synthesis by bromoperoxidase. *Science 202* (4372), 1094−6.

(11) Beissner, R. S., Guilford, W. J., Coates, R. M., and Hager, L. P. (1981) Synthesis of brominated heptanones and bromoform by a bromoperoxidase of marine origin. *Biochemistry 20* (13), 3724−31.

(12) Itoh, N., and Shinya, M. (1994) Seasonal evolution of bromomethanes from Coralline algae (Corallinaceae) and its effect on atmospheric ozone. *Mar. Chem. 45* (1−2), 95−103.

(13) Butler, A., and Carter-Franklin, J. N. (2004) The role of vanadium bromoperoxidase in the biosynthesis of halogenated marine natural products. *Nat. Prod. Rep. 21* (1), 180−8.

(14) Wever, R., Krenn, B. E., and Renirie, R. (2018) Marine vanadium-dependent haloperoxidases, their isolation, characterization, and application. in *Methods in Enzymology* (Moore, B. S., Ed.), Vol. *605*, Chapter 6, pp 141−201, Academic Press.

(15) Agarwal, V., Miles, Z. D., Winter, J. M., Eustaquio, A. S., El Gamal, A. A., and Moore, B. S. (2017) Enzymatic halogenation and dehalogenation reactions: pervasive and mechanistically diverse. *Chem. Rev. 117* (8), 5619−5674.

(16) Potin, P. (2008) Oxidative burst and related responses in biotic interactions of algae, in *Algal Chemical Ecology* (Amsler, C. D., Ed.), pp 245−271, Springer, Berlin, Heidelberg.

(17) Petrov, V. D., and Van Breusegem, F. (2012) Hydrogen peroxide—a central hub for information flow in plant cells. *AoB Plants 2012*, No. pls014, DOI: 10.1093/aobpla/pls014.

(18) Collén, J., et al. (2013) Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proc. Natl. Acad. Sci. U. S. A. 110* (13), 5247−5252.

(19) Brawley, S. H., et al. (2017) Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). *Proc. Natl. Acad. Sci. U. S. A. 114* (31), E6361.

(20) Lee, J., et al. (2018) Analysis of the draft genome of the red seaweed *Gracilariopsis chorda* provides insights into genome size evolution in rhodophyta. *Mol. Biol. Evol. 35* (8), 1869−1886.

(21) Liu, Y., Thornton, D. C. O., Bianchi, T. S., Arnold, W. A., Shields, M. R., Chen, J., and Yvon-Lewis, S. A. (2015) Dissolved organic matter composition drives the marine production of brominated very short-lived substances. *Environ. Sci. Technol. 49* (6), 3366−3374.

(22) Lin, C. Y., and Manley, S. L. (2012) Bromoform production from seawater treated with bromoperoxidase. *Limnol. Oceanogr. 57* (6), 1857−1866.

(23) Dijoux, L., Viard, F., and Payri, C. (2014) The more ee search, the more we find: discovery of a new lineage and a new species complex in the genus *Asparagopsis*. *PLoS One 9* (7), No. e103826.

(24) Zuccarello, G. C., Burger, G., West, J. A., and King, R. J. (1999) A mitochondrial marker for red algal intraspecific relationships. *Mol. Ecol. 8* (9), 1443−1447.

(25) Andreakis, N., Procaccini, G., Maggs, C., and Kooistra, W. H. C. F. (2007) Phylogeography of the invasive seaweed *Asparagopsis*

(Bonnemaisoniales, Rhodophyta) reveals cryptic diversity. *Mol. Ecol.* *16* (11), 2285−2299.
(26) McConnell, O., and Fenical, W. (1977) Halogen chemistry of the red alga *Asparagopsis*. *Phytochemistry 16* (3), 367−374.
(27) Greff, S., Zubia, M., Genta-Jouve, G., Massi, L., Perez, T., and Thomas, O. P. (2014) Mahorones, highly brominated cyclopentenones from the red alga *Asparagopsis taxiformis*. *J. Nat. Prod.* *77* (5), 1150−1155.
(28) Lee, J., Cho, C. H., Park, S. I., Choi, J. W., Song, H. S., West, J. A., Bhattacharya, D., and Yoon, H. S. (2016) Parallel evolution of highly conserved plastid genome architecture in red seaweeds and seed plants. *BMC Biol. 14* (1), 75.
(29) Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry 37* (6), 1585−95.
(30) Hughes, A. J., and Keatinge-Clay, A. (2011) Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol. 18* (2), 165−76.
(31) Dorrestein, P. C., Bumpus, S. B., Calderone, C. T., Garneau-Tsodikova, S., Aron, Z. D., Straight, P. D., Kolter, R., Walsh, C. T., and Kelleher, N. L. (2006) Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry 45* (42), 12756−66.
(32) Miller, I. J., Rees, E. R., Ross, J., Miller, I., Baxa, J., Lopera, J., Kerby, R. L., Rey, F. E., and Kwan, J. C. (2019) Autometa: automated extraction of microbial genomes from individual shotgun metagenomes. *Nucleic Acids Res. 47* (10), No. e57.
(33) Livak, K. J., and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods 25* (4), 402−408.
(34) Rinnerthaler, M., et al. (2012)et al Yno1p/Aim14p, a NADPH-oxidase ortholog, controls extramitochondrial reactive oxygen species generation, apoptosis, and actin cable formation in yeast. *Proc. Natl. Acad. Sci. U. S. A. 109* (22), 8658.
(35) Goodwin, K. D., North, W. J., and Lidstrom, M. E. (1997) Production of bromoform and dibromomethane by giant kelp: factors affecting release and comparison to anthropogenic bromine sources. *Limnol. Oceanogr. 42* (8), 1725−1734.
(36) Soedjak, H. S., and Butler, A. (1990) Characterization of vanadium bromoperoxidase from *Macrocystis* and *Fucus*: reactivity of vanadium bromoperoxidase toward acyl and alkyl peroxides and bromination of amines. *Biochemistry 29* (34), 7974−7981.
(37) Chekan, J. R., McKinnie, S. M. K., Moore, M. L., Poplawski, S. G., Michael, T. P., and Moore, B. S. (2019) Scalable biosynthesis of the seaweed neurochemical, kainic acid. *Angew. Chem., Int. Ed. 58* (25), 8454−8457.
(38) Denisov, I. G., Makris, T. M., Sligar, S. G., and Schlichting, I. (2005) Structure and chemistry of cytochrome P450. *Chem. Rev. 105* (6), 2253−77.
(39) Paul, N. A., de Nys, R., and Steinberg, P. D. (2006) Chemical defence against bacteria in the red alga *Asparagopsis armata*: linking structure with function. *Mar. Ecol.: Prog. Ser. 306*, 87−101.
(40) Hervé, C., Tonon, T., Collén, J., Corre, E., and Boyen, C. (2006) NADPH oxidases in eukaryotes: red algae provide new hints! *Curr. Curr. Genet. 49* (3), 190−204.
(41) de Oliveira, L. S., Tschoeke, D. A., Magalhães Lopes, A. C. R., Sudatti, D. B., Meirelles, P. M., Thompson, C. C., Pereira, R. C., and Thompson, F. L. (2017) Molecular mechanisms for microbe recognition and defense by the red seaweed *Laurencia dendroidea*. *mSphere 2* (6), e00094-17.
(42) Küpper, F. C., Gaquerel, E., Boneberg, E.-M., Morath, S., Salaün, J.-P., and Potin, P. (2006) Early events in the perception of lipopolysaccharides in the brown alga *Laminaria digitata* include an oxidative burst and activation of fatty acid oxidation cascades. *J. Exp. Bot. 57* (9), 1991−1999.

https://dx.doi.org/10.1021/acschembio.0c00299
ACS Chem. Biol. 2020, 15, 1662−1670

SUPPLEMENTARY INFORMATION FOR:

Genetic and biochemical reconstitution of bromoform biosynthesis in *Asparagopsis* lends insights into seaweed ROS enzymology

Hem R. Thapa,[1] Zhenjian Lin,[2] Dongqi Yi,[1] Jennifer E. Smith,[3] Eric W. Schmidt,[2] Vinayak Agarwal[1,4,*]

[1]School of Chemistry and Biochemistry, Georgia Institute of Technology

[2]Department of Medicinal Chemistry, University of Utah

[3]Center for Marine Biodiversity and Conservation, Scripps Institution of Oceanography, University of California, San Diego

[4]School of Biological Sciences, Georgia Institute of Technology

*Correspondence: vagarwal@gatech.edu; Ph: 404-385-3798

MATERIALS AND METHODS

Sample collection

Samples of the filamentous, sporophyte phase of the red macroalga *Asparagopsis taxiformis* were collected from five different locations: Guam, California, Hawaii, Florida Keys, and Fiji. Samples of *Macrocystis pyrifera* were collected from the southern California coast (San Diego). Algal biomass for nucleic acid preparation was stored in RNA*later* solution (3.54 M ammonium sulfate, 16.7 mM sodium citrate, 13.3 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0). Biomass for chemical analysis was stored without any treatment. Both types of biomass samples were kept in a cooler during field collection and samples once brought to the lab were stored at -80 °C until further use.

Culturing of *Asparagopsis taxiformis*

Upon collection of the filamentous phase of *A. taxiformis* in California, samples were cultured in the lab at Scripps Institution of Oceanography to allow for further experimentation. Specifically, samples were sorted and cleaned under stereomicroscope to isolate *A. taxiformis* from any other taxa growing within the algal matrix. Roughly 1 g of wet weight was placed in each of several 2-liter Erlenmeyer flasks filled with autoclaved and UV sterilized seawater from the Scripps Institution of Oceanography's running seawater system. Flasks were sealed with stoppers and all samples were gently aerated with ambient air to allow for a gentle tumble. Seawater was replaced weekly and sample biomass to volume ratios were kept below 5g/L. All flasks were cultured on a 12hr day/light cycle at 50 mE/cm^2/s at ambient temperature (roughly 22 °C).

GC-MS analysis of chemical extracts

Algal biomass collected either from the field or laboratory were lyophilized for 16 h. Freeze-dried biomass was used for extraction with MeOH. Biomass was soaked in MeOH for several hours before extraction with vigorous agitation on vortex mixer, centrifuged at 16,000×g for 30 min to remove debris, and an aliquot of supernatant was analyzed by GC-MS (1260G with 7890a MS; Agilent Technologies) in electron ionization (70 eV) mode using a DF-5ms ultra inert GC column (30 m length, 0.25 mm width and 0.5 µM film thickness). Bromoform production quantitation for *A. taxiformis* and *M. pyrifera* was quantified based on calibration curves generated from a bromoform standard. The column temperature conditions were as follows: 40 °C for 3 min, increased to 200 °C at 10 °C/min and held for 1 min with total run time of 20 min. Injection port, interface and ion source were kept at 250 °C, 300 °C and 230 °C, respectively. Helium was used as carrier gas at a flow rate of 0.9 mL/min.

LC-MS analysis

Chemical extracts of *A. taxiformis* prepared above were analyzed using 1290 Infinity II UHPLC system (Agilent Technologies) coupled to a high-resolution Impact II Q-TOF mass spectrometer (Bruker Daltonics). Mass spectrometry data were collected in the negative ionization mode in the mass range *m/z* 100-2000 Da, and the data collected were deposited in the MassIVE repository (ID# MSV000084456). Samples were analyzed using Kinetex™ 1.7 µm C18 reversed phase UHPLC column (50 × 2.1 mm) at a flow rate of 0.5 mL/min and the chromatographic separation was achieved using two solvents; solvent A (H_2O, 0.1% formic acid) and solvent B (MeCN, 0.1% formic acid). The chromatography elution profile was as follows: 5% solvent B from 0-5 min, linear gradient to 100% solvent B from 5-20 min, 100% solvent B from 20-25 min, linear gradient to 5% solvent B from 25-26 min, 5% solvent B from 26-28 min, linear gradient to 100% solvent B from 28-29 min, 100% solvent B from 29-33 min, linear gradient to 5% solvent B from 33-34 min, and 5% solvent B from 34-35 min.

MS^2 and MS^3 characterization of molecule **1** (1,1,1,5,5,5-hexabromo-2,4-dione) was achieved by direct infusion into the Bruker amaZon SL ion-trap mass spectrometer in a negative ionization mode. Algal extract was prepared and enriched for **1** prior to mass spectrometry-based characterization. Freeze-dried algal biomass (55 g) was extracted twice with 100 mL MeOH by stirring at room temperature for 16 h, extracts pooled, filtered, and concentrated using rotary evaporator. The concentrated extract was resuspended in small volume of hexane:EtOAc (40:1) before application to a silica gel chromatography column equilibrated in the same solvent. Elution was done with a gradient of hexane:EtOAc (40:1→10:1→8:1→4:1→2:1). The hexane:EtOAc (2:1) fraction containing molecule **1** was concentrated and then infused in to the mass spectrometer at a flow rate of 4 µL/min using a metered peristaltic pump. Signal was allowed to stabilize before isolation of MS^1 ion of 572.49 and then fragmented to obtain MS^2 ions *m/z* 318.74 and *m/z* 250.75. MS^2 ion *m/z* 318.74 was isolated and fragmented further to obtain MS^3 spectra.

Genomic DNA extraction

DNA from algal samples was extracted following the protocol described in previous study.[1] Algal biomass samples stored in RNAlater were thawed at 4 °C and washed twice with TE buffer (10 mm Tris-HCl (pH 7.5), 1 mM Na-EDTA) to remove excess salts. Biomass was frozen with liquid nitrogen and freeze-dried overnight in a lyophilizer. Dried biomass was ground to fine powder with mortar and pestle using liquid nitrogen. 500 mg of crushed powder was resuspended in 10 mL of CTAB buffer (3% w/v cetyltrimethylammonium bromide (CTAB), 1.4 M NaCl, 20 mM Na-EDTA, 100 mM Tris-HCl (pH 8.0),

0.2% polyvinylpolypyrrolidone, 0.2% β-mercaptoethanol, and 0.2 mg/mL proteinase K). Samples were incubated at 55 °C for 2 h with gentle mixing every 15 min and then centrifuged at 14,000×g for 15 min at room temperature to remove cellular debris. To the supernatant was added 2 mL of 5 M Na-acetate (pH 8.0), cooled on ice for 30 min, and centrifuged at 14,000×g for 15 min at 4 °C. The aqueous layer (supernatant) was extracted with equal volume of phenol:chloroform:isoamyl alcohol ((25:24:1) saturated with 10 mM Tris-HCl (pH 8.0), 1 mM Na-EDTA) and the mixture was centrifuged at 12,000×g for 5 min at 4 °C. The supernatant thus obtained was extracted with equal volume of chloroform and then centrifuged at 12,000×g for 5 min at 4 °C. To the supernatant thus obtained was added equal volume of ice-chilled isopropanol, mixed by inverting the tube, and centrifuged at 12,000×g for 15 min at 4 °C. The supernatant was removed and the pellet containing nucleic acids was washed with 75% (v/v) ethanol, centrifuged as before, supernatant removed, and the pellet was dried in a speedvac to remove any residual ethanol. The dried pellet was resuspended in 0.5 mL of 10 mM Tris-HCl (pH 8.0) containing 0.6 mg/mL of RNase A, incubated at room temperature for 1 h and then incubated overnight at 4 °C to digest any contaminating RNA. The RNase treated sample was extracted with equal volume of phenol:chloroform:isoamyl alcohol followed by chloroform extraction as described before. Na-acetate (pH 5.2) was added to the supernatant to a final concentration of 0.3 M followed by addition of three volumes of ice-chilled ethanol, and then incubated at -20 °C for 2 h to precipitate DNA. The sample was then centrifuged at 16,000×g for 20 min at 4 °C, the pellet washed with 75% (v/v) ethanol, centrifuged again, and the pellet was dried in a speedvac. The DNA pellet was resuspended in desired volume of 10 mM Tris-HCl (pH 8.0) and the quality of genomic DNA was analyzed by gel electrophoresis and nanodrop.

Phylogenetic analysis

Phylogenetic assignment of *Asparagopsis* samples was achieved by PCR amplification and amplicon sequencing of the *18S rRNA* gene. Universal primers, as described in literature, (forward primer 566:5'-CAGCAGCCGCGGTAATTCC-3' and reverse primer 1200:5'-CCCGTGTTGAGTCAAATTAAGC-3') were used for PCR amplification, and covered variable region v4-v5 of *18S rRNA* gene.[2] Intraspecific relationship of *A. taxiformis* samples were inferred by amplification of mitochondrial marker *cox2-cox3* spacer that spans the intergenic region between two conserved gene pair cytochrome oxidase subunit 2 (*cox2*) and cytochrome oxidase subunit 3 (*cox3*). Degenerate primers, as described before, (forward primer cox2:5'-GTACCWTCTTTDRGRRKDAAATGTGATGC-3' and reverse primer cox3:5'-GGATCTACWAGATGRAAWGGATGTC-3') was used for amplification of *cox2-cox3* spacer.[3] PCR reactions were done in total volume of 25 µL and contained 20 ng of genomic DNA, 0.4 µM primers, 2.5 mM dNTPs, 1× Phusion buffer and 0.5 unit of Phusion-high fidelity DNA polymerase. PCR reactions were

carried out in a thermocycler using the following program: initial denaturation at 98 °C for 2 min, 33 cycles of 98 °C for 30 sec, 55 °C for 30 sec, 72 °C for 60 sec for *18S rRNA* gene and 72 °C for 30 sec for *cox2-cox3* spacer, and final extension at 72 °C for 10 min. The PCR amplicons were purified using DNA Clean and Concentrator Kit (Zymo Research). An 'A' nucleotide was added to blunt-ends of PCR product using GoTaq polymerase (Promega). The amplicons were then cloned into pGEM-T Easy vector as per manufacturer's instructions followed by transformation into *Escherichia coli* (DH5α) under appropriate antibiotic selection. Individual colonies were selected for growth and plasmid DNA extracted using standard miniprep protocols followed by Sanger sequencing of the inserts. Sequences were deposited in the GenBank database.

GenBank accession numbers for algal *18S rRNA* region are as follows:

Guam (MN547333)

California (MN547332)

Hawaii (MN547334)

Florida (MN547336)

Fiji (MN547335)

GenBank accession numbers for algal *cox2-cox3* spacer are as follows:

Guam (MN563722)

California (MN563723)

Hawaii (MN563724)

Florida (MN563725)

Fiji (MN563726)

Phylogenetic tree was constructed using the *cox2-cox3* spacer sequence information of five *A. taxiformis* samples from our study and sequences of other *A. taxiformis* samples available in the GenBank database. Tree was constructed in MEGA7 using default parameters of Maximum Likelihood method based on Tamura-Nei model.[3, 4]

Preparation of algal protein extract

Algal biomass for protein extract preparation were harvested, snap frozen in liquid nitrogen and stored at -80 °C until further use. 100 mg of frozen biomass was added to 2 mL safe-lock tube containing 0.8 mL of homogenization buffer (100 mM HEPES-Na (pH 7.6), 10% glycerol) and 1.1 g of 1.4 mm ceramic beads. Samples were homogenized at 4 °C using a Bullet Blender Storm 24 (Next Advance) for 10 min at a power of '10'. The lysate thus obtained was centrifuged at 9,000×g for 10 min to obtain the total protein extract as the supernatant. Typical protein preparation utilized 400-600 mg of total algal biomass which resulted in 2-3 mL of supernatant. The 9,000×g supernatants were pooled and divided equally into two aliquots and then dialyzed separately in dialysis buffer A (100 mM HEPES-Na (pH 7.6), 10% glycerol, 50 mM KCl and 10 µM sodium orthovanadate) to yield holo-enzyme extract and in buffer B (100 mM HEPES-Na (pH 7.6), 10% glycerol, 50 mM KCl and 1 mM Na-EDTA) to yield apo-enzyme extract. Dialysis with Na-EDTA will result in apo-enzyme extract as EDTA is known to form stable complex with vanadate.[5] Dialysis was done using Slide-A-Lyzer dialysis cassette (3.5K MWCO, 3 mL) overnight at 4 °C in 1 L of dialysis buffer A or B. Dialysis was repeated in fresh buffer for another 4 h before using the protein extract for enzyme assays.

Enzyme assays for bromoform production

Enzyme assays were typically done in 1 mL total volume at 30 °C and contained 100 mM HEPES-Na (pH 7.6), 50 mM KBr, 10 µM sodium orthovanadate, 1 mM substrate (**5** or **6**) and 50 µg of protein extract or 1 µM recombinant protein. Reactions were initiated by addition of 1 µL of 1 M H_2O_2, and same volume of H_2O_2 was added every 10 min for a total reaction time of 90 min. Assay with apo-enzyme extract was done in absence of sodium orthovanadate. Reactions were quenched by addition of 1 mL of GC-grade diethyl ether and 100 µL brine. Reaction products were then extracted by vigorous agitation on a vortex mixer, centrifuged at 1,500×g to obtain phase separation, and the top organic layer removed. Extraction was repeated once again with 1 mL of diethyl ether, organic extracts pooled, concentrated using rotary evaporator, and aliquots of extracts were analyzed by GC-MS using conditions described above.

RNA Isolation

Total RNAs from algal sample were isolated as reported before.[6] Algal biomass stored in RNAlater was washed with TE buffer as described earlier and then frozen in liquid nitrogen. 200 mg of frozen tissue was pulverized in liquid nitrogen using mortar and pestle. Frozen powder was added to 1 mL of TRIzol (Invitrogen) and mixed briefly using vortex to obtain homogenous solution. Sample incubated at room

temperature for 5 min, mixed again, and centrifuged at 10,000 x g for 5 min at room temperature to pellet cell debris. Centrifugation repeated when necessary to obtain the clear supernatant. The supernatant mixed with 0.2 mL of chloroform, mixed using vortex, incubated at room temperature for 5 min, and then centrifuged at 12,000×g for 15 min at 4 ºC to obtain phase separation. The supernatant (aqueous phase) containing RNA was carefully removed to a new tube, 0.5 mL of isopropanol added, and mixed by invert mixing. Sample incubated at room temperature for 10 min followed by centrifugation at 12,000×g for 15 min at 4 ºC to pellet RNA. The supernatant was removed, RNA pellet washed with 1 mL of 75% ethanol by inverting mixing and centrifuged at 12,000×g for 15 min at 4 ºC. The supernatant was removed, RNA pellet was partially dried using speed vacuum, and further processed to ensure removal of contaminating polysaccharides. This was achieved by resuspension of the RNA pellet in 2 M LiCl at room temperature followed by centrifugation at 12,000×g for 15 min at 4 ºC. This process was repeated until the size of the RNA pellet did not reduce further. The RNA pellet was then dissolved in 0.5 mL TE buffer, extracted with equal volume of phenol/chloroform/isoamyl alcohol ((25:24:1) saturated with 10 mM Tris-HCl (pH 8.0), 1 mM Na-EDTA), and centrifuged for phase separation. The aqueous phase was further extracted with equal volume of chloroform and centrifuged at 12,000×g for 15 min at 4 ºC. The supernatant containing RNA was treated with 0.1 volume of 3 M Na-acetate (pH 5.2) and 2.5 volumes of ethanol and incubated overnight at -80 ºC to precipitate RNA. The sample was centrifuged at 12,000×g for 15 min at 4 ºC, the supernatant removed, and the RNA pellet was washed with 75% (v/v) ethanol. The pellet was dried in a speedvac followed by resuspension in desired volume of nuclease free water. The isolated total RNA was treated with DNA-free™ DNA removal kit (Invitrogen) to remove any contaminating genomic DNA. Several micrograms of good quality RNA were obtained with A_{260}/A_{280} and A_{260}/A_{230} values greater than 2.0. Analysis of total RNA by agarose gel did not show any genomic DNA contamination and was further confirmed by reverse transcriptase-PCR (RT-PCR) analysis.

Transcriptome sequencing, assembly and analysis

Short-gun sequencing was performed using an Illumina HiSeq 2000 sequencer with a ~350 bp inserts and 125 bp paired-end runs, or an Illumina NovaSeq sequencer with a ~450 bp inserts and 150 bp paired-end runs at the Huntsman Cancer Institute's High Throughput Genomics Center at the University of Utah. Raw reads were trimmed and adaptors removed by Trimmomatic-0.39[7] with parameters (PE -phred33 ILLUMINACLIP TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:(80 or 150)). The trimmed reads were merged using BBMerge[8] and then assembled using rnaSPAdes (mink=21 maxk=121 step=10)[9] with standard parameters in the Center for High Performance Computing at the University of Utah.

The eukaryotic genes predicted by AUGUSTUS were used as reference genes, trimmed reads of each sample were multi-mapped to the reference using Salmon6 with parameters (salmon index -t -i index -k 31; salmon quant --index --validateMappings --libType A --dumpEq -r). The identical predicted genes from both samples were hierarchically clustered into clusters, and then cluster count was summarized using Corset 1.047 according to shared reads information with parameters (-f true -g -n -i salmon_eq_classes). Transcript abundance of each gene was estimated using EdgeR8 by normalized counts per million (cpm).

RT-PCR and cloning

All genes from *A. taxiformis* that are characterized in this study were cloned using cDNA as template. GeneBank accession numbers for *A. taxiformis* ACP and FabH are AOM65890.1 and AOM66007.1, respectively. The cDNA was prepared using SuperScript IV first-strand synthesis system (Thermo Fisher). The first step of cDNA synthesis reaction was set up in 13 µL reaction volume containing 600 ng of total RNA, 1 µL of 10 mM dNTP mix, and 1 µL of 50 µM oligo dT_{20} primer. The reaction was incubated at 65 °C for 5 min, incubated in ice for 2 min, and then added to a tube containing 4 µL of 5x superscript buffer, 1 µL of 100 mM DTT, 1 µL of ribonuclease inhibitor, and 200 unit of Superscript IV enzyme. The reaction mix was incubated at 55 °C for 10 min followed by heat inactivation at 80 °C for 10 min. The reaction mix was further treated with 2 unit of RNase H by incubating at 37 °C for 20 min. The product (3 µL) from cDNA synthesis reaction was used as template for PCR with high fidelity Phusion DNA polymerase and gene specific primers. Negative control reaction for RT-PCR was identical to other reactions except omission of the reverse transcriptase enzyme. RT-PCR products were cleaned with DNA clean and concentrator kit, nucleotide A added using GoTaq DNA polymerase, and then cloned into pGEM-T Easy vector as described above followed by transformation into *E. coli* DH5α, colony selection and miniprep. The pGEM-T constructs were used as template for cloning into desired expression vectors.

Cloning and protein purification

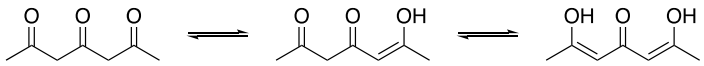
For characterization of *C. crispus* VHPOs[10], three of the five VHPOs with highest sequence identity to *A. taxiformis* Mbb4 protein were chosen, gene fragments optimized for expression in *E.coli* were synthesized (Twist Biosciences), and cloned into pGEM-T Easy vector. The *Cc*VHPOs were arbitrarily numbered. The GenBank accession numbers of *Cc*VHPO1, *Cc*VHPO2 and *Cc*VHPO3 are XP_005714237.1, XP_005719435.1 and XP_005710894.1, respectively. The pGEM-T gene construct for each gene was used as a template for PCR with high fidelity Phusion DNA polymerase. The MatB gene construct as described in previous study[11] was a gift from Dr. Keatinge-Clay (UT Austin) and was used as a template for cloning

into pET28MBP vector. The PCR amplicons for each gene were cloned into expression vector using NEBuilder HiFi DNA Assembly master mix following the manufacturer's instructions. The *E. coli* expression constructs used in this study are listed in Supplementary Table 3. The expression constructs for algal genes were transformed into *E. coli* (Rosetta) cells, whereas MatB and Sfp constructs were transformed into *E. coli* (BL21DE3) cells. In a typical protein expression experiment, 10-mL overnight culture was used to inoculate two 1-liter terrific broth media (1-liter for MatB and Sfp) containing appropriate antibiotic (kanamycin at 50 mg/L, ampicillin at 100 mg/L and chloramphenicol at 34 mg/L). *E. coli* cultures were grown at 30 ºC with shaking at 180 rpm until OD_{600} of 0.8, cultures cooled down to 18 ºC, incubated for another 1 h at 18 ºC, and then induced with 0.3 mM IPTG. All subsequent steps for purification were done on ice or at 4 ºC using the standard protocol unless mentioned otherwise. Cultures were harvested by centrifugation, pellet resuspended in lysis buffer (20 mM Tris pH 8.0, 500 mM NaCl, 1 mM EDTA), cells lysed by sonication, centrifuged at 30,000×g for 40 min, and the supernatant loaded into 5 mL His-Trap Ni-NTA column using AKTA Prime FPLC system. Column was washed extensively with wash buffer (20 mM Tris pH 8.0, 500 mM NaCl, 30 mM imidazole) and protein eluted using linear gradient to 100% elution buffer (20 mM Tris pH 8.0, 500 mM NaCl, 250 mM imidazole) for a total volume of 40 mL. Purity of eluant fractions were checked using SDS-PAGE, and fractions containing protein of interest were pooled and concentrated to 2.5 mL using Amicon centrifugal filters. Protein samples were desalted using PD-10 column and eluted in storage buffer (20 mM HEPES pH 7.5, 50 mM KCl, 10% glycerol). Purified proteins were stored as small aliquots at -80 °C and fresh aliquots of each protein were used for enzyme assays. None of the proteins characterized in this study showed loss of activity during storage at -80 °C for several months.

Monocholorodimedone (MCD) Assay

Halogenation activity of Mbb enzymes and *Cc*VHPOs were determined using MCD assay.[5] Reactions were performed in a quartz cuvette at 25 ºC, and the decrease in absorbance of MCD substrate were monitored spectrophotometrically at 290 nm every 30 sec for 15 min using NanoDrop OneC UV-Vis spectrophotometer. Reactions were initiated by adding the recombinant enzymes in a 0.75 mL total reaction volume containing 100 mM Na-citrate (pH 6.5), 100 mM KBr or KCl, 100 µM MCD, 10 µM sodium orthovanadate, and 2 mM hydrogen peroxide. Assays were done in triplicate for total of six enzymes at a varied concentration, Mbb1-250 nM, Mbb3-1 µM, Mbb4-100 nM, *Cc*VHPO1-50 nM, *Cc*VHPO2-50 nM, and *Cc*VHPO3-250 nM. Enzyme concentrations used in the assays were empirically determined due to the difference in the activity of individual enzymes.

Synthesis of heptane-2,4,6-trione (6)



6a **6** **6b**

6 was prepared as described before.[12] Commercial 2,6-dimethyl-γ-pyrone (500 mg, 4.03 mmol) was dissolved in 2.5 mL ethanol and treated with 0.5 mL of 16 M NaOH. The reaction mixture was stirred at 60 ºC for 5 h, and then heated at 100 ºC for another 1 h. Precipitate formed during the reaction was filtered and washed with diethyl ether. The crystals were dissolved in water followed by the addition of 5 mL of 3 M HCl solution. The aqueous phase was extracted thrice with diethyl ether. The organic layer was dried with anhydrous Na-sulfate and concentrated under vacuum which afforded the target compound as a yellow solid. Tautomeric equilibrium of product in chloroform-*d* was 68.8% **6**, 24.6 % **6b** and 6.6 % **6a**. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ 15.21 (s, 1H), 14.18 (s, 1H), 5.56 (s, 1H), 5.14 (s, 1H), δ 3.70 (s, 4H), δ 3.40 (s, 2H), 2.25 (s, 6H), 2.08 (s, 3H), 1.98 (s, 3H). **^{13}C-NMR** (101 MHz, $CDCl_3$) δ 202.07, 193.88, 191.23, 187.02, 178.65, 101.21, 98.76, 57.85, 54.12, 30.52, 24.76, 21.99.

Synthesis of malonyl-CoA

Previously described enzyme MatB was used to synthesize malonyl-CoA.[11] MatB enzyme assay was performed in a 100 µL reaction volume containing 100 mM HEPES-Na (pH 7.6), 20 mM $MgCl_2$, 20 mM Na-malonate, 20 mM ATP, 5 mM TCEP, 5 mM CoA-SH, 15% (v/v) glycerol and 5 µM MatB. Reaction was incubated at 30 ºC for 23 h. Control reaction omitted MatB. An aliquot of reaction was quenched with equal volume of solvent (MeOH + 1% trifluoroacetic acid (TFA) for HPLC analysis and MeCN + 0.2% formic acid for LC-MS/MS analysis), precipitate removed by centrifugation, and supernatant analyzed by HPLC system and LC-MS/MS using Phenomenex Luna C8 column 5 µm (250 × 4.6 mm) column. LC-MS/MS analysis was done using Agilent 1290 Infinity II UHPLC system coupled to a high-resolution Bruker Impact II Q-TOF mass spectrometer. MS data were collected in the positive ionization mode from *m/z* 100-2000 Da. HPLC analysis used solvent A (100% H_2O, 0.1% TFA) and solvent B (100% MeOH, 0.1% TFA). LC-MS analysis used solvent A (100% H_2O, 0.1% formic acid) and solvent B (100% MeCN, 0.1% formic acid). Samples were analyzed at a flow rate of 0.5 mL/min, and elution profile for both HPLC and LC-MS/MS analysis were: 5% solvent B from 0-3 min, linear gradient to 25% solvent B from 3-8 min, linear gradient to 50% solvent B from 8-20 min, linear gradient to 100% solvent B from 20-23 min, 100% solvent B from 23-26 min, linear gradient to 5% solvent B from 26-27 min, 5% solvent B from 27-28 min, linear gradient to 100% solvent B from 28-29 min, 100% solvent B from 29-30 min, and linear gradient to 5% solvent B from 30-31 min.

Synthesis of malonyl-*S*-ACP

Malonyl-CoA synthesized above was used as the substrate to modify ACP using Sfp enzyme.[13] Assays were performed at 30 ºC for 3 h in a 100 µL total reaction volume containing 100 mM HEPES-Na (pH 7.6), 500 µM ACP, 50 µL reaction product from MatB assay, 10 mM $MgCl_2$, and 2 µM Sfp. Control reaction omitted the Sfp enzyme. An aliquot of reaction was quenched by adding equal volume of MeCN + 0.2% formic acid, precipitate observed was removed by centrifugation, and supernatant was analyzed by LC-MS/MS using Aeris 3.6 µm widepore XB-C18 (200 Å 250 × 4.6 mm) column at a flow rate of 0.5 mL/min. MS data were collected in the positive ionization mode from *m/z* 100-2000 Da, and data analysis for ACP proteins (with or without acyl-groups) were done following the method described in previous study.[14] The elution profile for LC-MS/MS analysis was as follows: 5% solvent B from 0-5 min, linear gradient to 70% solvent B from 5-30 min, linear gradient to 95% solvent B from 30-31 min, 95% solvent B from 31-35 min, linear gradient to 5% solvent B from 35-36 min, 5% solvent B from 36-38 min, linear gradient to 95% solvent B from 38-39 min, 95% solvent B from 39-42 min, linear gradient to 5% solvent B from 42-43 min, and 5% solvent B from 43-46 min.

Synthesis of acetoacetyl-*S*-ACP (8)

Synthesis of **8** was performed at 30 ºC for 3 h in a 100 µL total reaction volume containing 100 mM HEPES-Na (pH 7.6), 50 µL reaction product from Sfp loading assay, 1 mM acetyl-CoA, 1 mM TCEP and 2 µM FabH. After 3 h incubation, an aliquot of reaction was quenched with equal volume of MeCN + 0.2% formic acid, precipitate removed by centrifugation and analyzed by LC-MS/MS using the procedure described above for malonyl-*S*-ACP.

Preparative scale synthesis of 8 and enzyme assay

2×2.5 mL assays for synthesis of malonyl-CoA were set up as described above. After incubation at 30 ºC for 23 h, both samples were pooled, and reaction product was used to set up 4×2.5 mL assays for synthesis of malonyl-*S*-ACP using the reaction conditions described above. After incubation of Sfp loading reaction at 30 ºC for 3 h, precipitate was removed by centrifugation at 1500×g for 10 min, and the supernatant was used to set up 8×2.5 mL assays with the FabH enzyme. Reaction conditions for synthesis of acetoacetyl-*S*-ACP were same as described above. After incubation at 30 ºC for 3 h, samples were pooled, centrifuged, and the supernatant was concentrated to 2.5 mL using Amicon centrifugal filters (3 kDa MWCO). Concentrated sample was loaded onto a desalting PD-10 column, and acetoacetyl-*S*-ACP was eluted using the buffer 20 mM HEPES-Na (pH 7.5), 50 mM KCl, 10% glycerol. Aliquots of acetoacetyl-*S*-ACP were

stored at -80 °C. Production of malonyl-CoA, malonyl-*S*-ACP and acetoacetyl-*S*-ACP in each reaction step was verified using LC-MS/MS. Enzyme assays for bromoform synthesis were performed at 30 ºC for 1.5 h in a 1 mL reaction volume containing 100 mM HEPES-Na (pH 7.6), 50 mM KBr, 10 µM sodium orthovanadate, 0.67 mM acetoacetyl-*S*-ACP and 1 µM Mbb1 or Mbb4 enzymes. Reactions were initiated by addition of 1 µL of 1 M H_2O_2. H_2O_2 was replenished by adding 1 µL of 1 M H_2O_2 after every 10 min through the course of reaction. After 90 min, reactions were quenched by addition of equal volume of diethyl ether and 100 µL brine. Reactions were extracted by vigorous agitation, centrifuged at 1,500×g for 10 min and the top organic layer removed. Extraction was repeated thrice. Organics extracts were pooled, concentrated, and 10 µL of concentrated extract was analyzed by GC-MS using manual injection using the conditions described above.

Metagenome sequencing, assembly, binning and annotation.

Short-gun sequencing was performed using an Illumina NovaSeq sequencer with a ~450 bp inserts and 150 bp paired-end runs at the Huntsman Cancer Institute's High Throughput Genomics Center at the University of Utah. Raw reads were trimmed and adaptors removed by Trimmomatic-0.39[7] with parameters (PE -phred33 ILLUMINACLIP TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:150). The trimmed reads were merged using BBMerge[8] and then assembled using metaSPADES (mink=21 maxk=121 step=10)[9] with standard parameters in the Center for High Performance Computing at the University of Utah. Bacterial contigs were binned out of metagenome assembly using Autometa.[15] The eukaryotic contigs were annotated using AUGUSTUS[16] with transcriptome assembly as training data.

Quantitative RT-PCR

The total RNA from each sample were isolated and cDNA synthesized as above. The product from cDNA synthesis reaction was diluted five-fold in nuclease free water and 3 µL was used as template in a 20 µL PCR reaction containing 10 µL of PowerUp SYBR green master mix (ThermoFisher) and 0.5 µM primers. The qRT-PCR reactions were carried out on a StepOnePlus Real-Time PCR system (Applied Biosystems) using following program: initial denaturation at 95 °C for 2 min, 40 cycles of (95 °C for 15 sec, 55 °C for 15 sec and 72 °C for 60 sec. Expression level for each gene were calculated relative to actin gene using the $2^{-\Delta\Delta CT}$ method.[17] Primers used for qRT-PCR are listed on supplementary Table 4.

Cloning of *yno1*, *mbb2* and *CcNOX (Ccmbb2)* in yeast expression plasmids

The *yno1*/pYES2 plasmid and BY4741 yeast strain (MATa, *his3*Δ1, *leu2*Δ0, *met15*Δ0, *ura3*Δ0, *yno1Δ*), which has a deletion of *yno1* gene, was a gift from Reddi laboratory at Georgia Tech. Synthetic clones for *mbb2* and *CcNOX* with sequences optimized for expression in yeast were obtained commercially. All genes, *yno1*, *mbb2*, and *CcNOX (Ccmbb2)*, were amplified using high fidelity Phusion DNA polymerase and PCR amplicons were cloned into pESC-URA expression vector using NotI and SpeI restriction sites. Gene constructs were transformed into *yno1Δ* yeast strain using standard lithium acetate protocol.[18] Transformants were selected and patches of positive colonies were maintained on synthetic dropout medium (SC-URA) supplemented with 2% glucose.

Dihydroethidium (DHE) Assay

Reactive oxygen species production by yeast cells were measured using DHE as a fluorescent probe and following the methodology as described before.[19] Yeast cultures were maintained in SC-URA supplemented with 2% raffinose. Triplicate cultures for each gene constructs were used for DHE assay. Several time points and galactose concentration were screened to determine optimal signal for DHE assay. In brief, 25 µL of overnight cultures was used to inoculate 5 mL of media containing 10% galactose. The cultures were grown at 30 ºC by shaking at 200 rpm for 43 h, harvested by centrifugation, washed, and resuspended in distilled water. Cells corresponding to 6.25 OD_{600} units were harvested by centrifugation at 6,000 x g and washed three times with PBS buffer. Cells were resuspended in 300 µL PBS buffer, DHE added to final concentration of 10 µM, and then incubated in dark at 30 ºC for 10 min with invert mixing every two min. After incubation, cells were harvested by centrifugation, washed at least three times with PBS buffer, pellet resuspended in 300 µL PBS buffer, 100 µL of cell suspension aliquoted to black microwell plate, and fluorescence measured using Biotek Synergy Mx microplate reader with an excitation wavelength of 518 nm and an emission wavelength of 605 nm.

SUPPLEMENTARY FIGURES

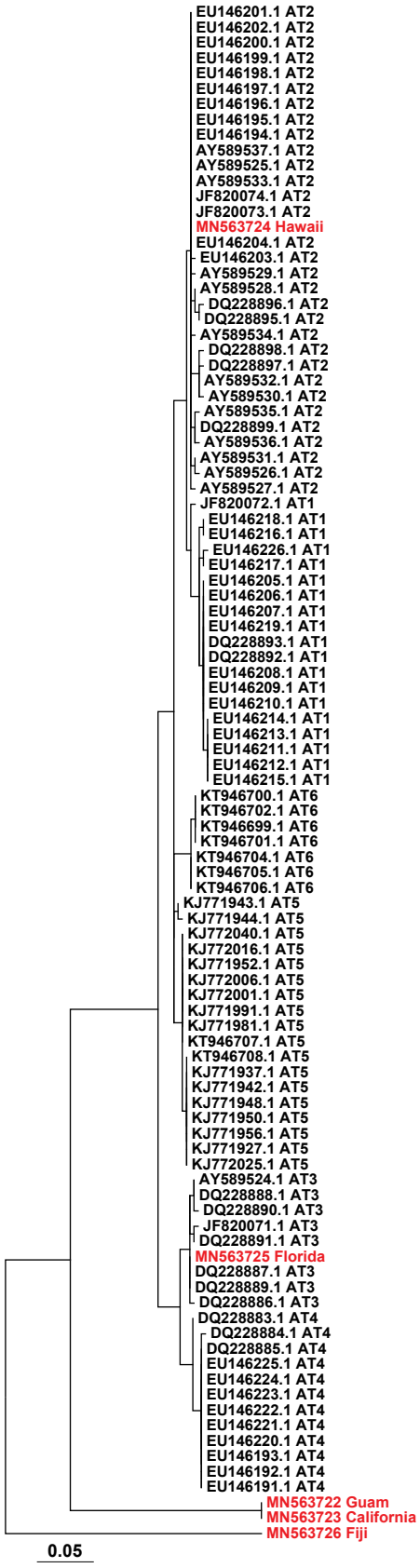


Figure S1: Maximum Likelihood phylogenetic tree of *Asparagopsis taxiformis* using *cox2-cox3* spacer. GenBank accession number followed by lineage information is shown for all *cox2-cox3* spacer sequences. Samples from this study are highlighted in red. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model.[4] The tree with the highest log likelihood (-1151.26) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood[10] approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 100 nucleotide sequences. All positions containing gaps and missing data were eliminated. There was a total of 267 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.[20]

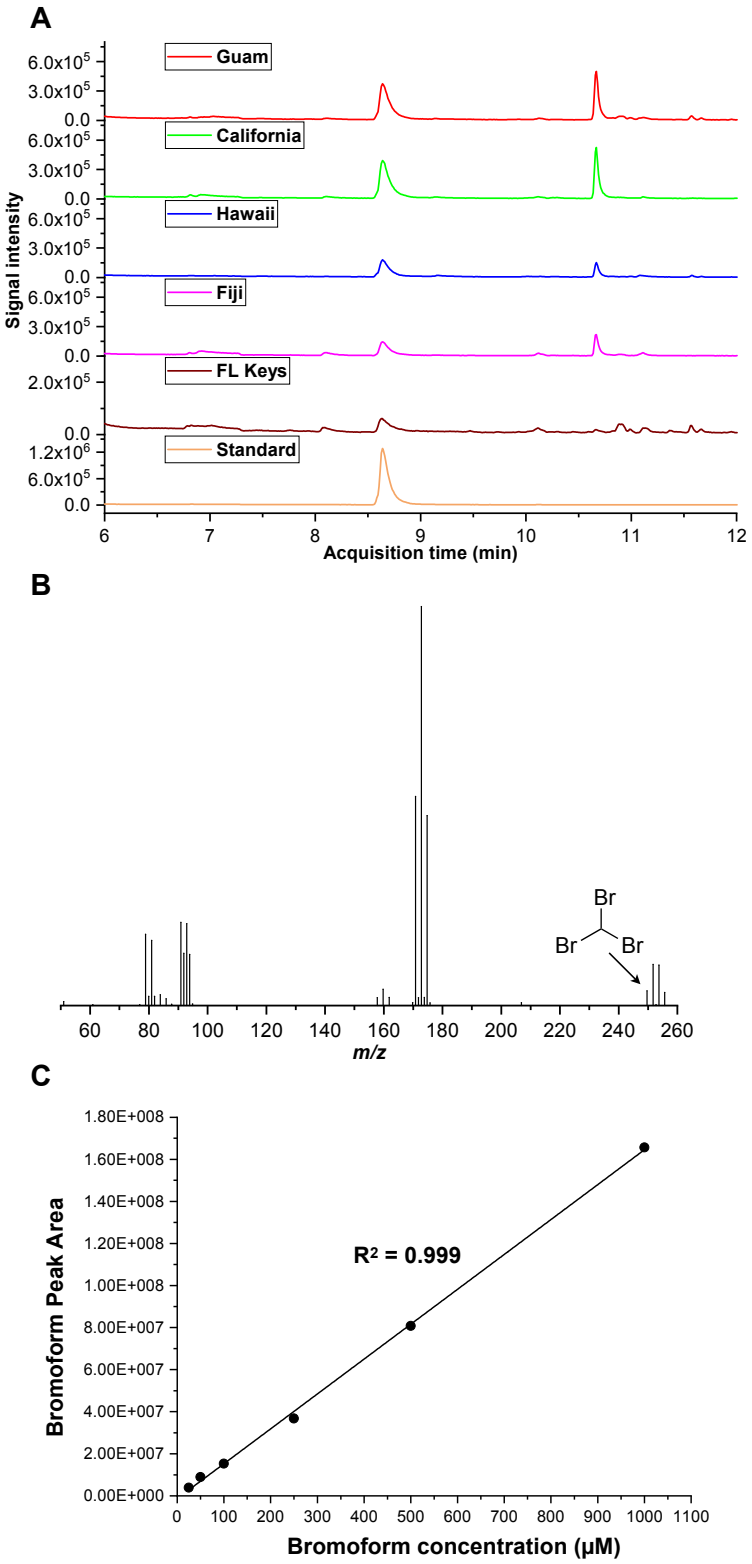


Figure S2: Bromoform production in red macroalga *Asparagopsis taxiformis*. (**A**) GC-MS chromatograms for MeOH extracts demonstrating bromoform production by *A. taxiformis* in geographically disperse samples. (**B**) MS spectrum of bromoform. Arrow corresponds to monoisotopic ion. (**C**) Calibration curve generated based on GC-MS analysis of different concentrations of bromoform standard. 25 mg freeze dried algal biomass for each sample was soaked in 625 µL MeOH for several hours, metabolites extracted using vortex mixer, and then centrifuged. The supernatant (MeOH extracts) was diluted 25-fold in DCM and 1 µL was analyzed by GC-MS. Bromoform peak area for each sample was determined and then used for quantification using bromoform standard calibration curve. Representative calculation: peak area 46,655,308 for California sample corresponds to 294.405 µM of bromoform. When adjusted for dilution, peak area is equivalent to 46.5 µg of bromoform in 1 mg of dry biomass (4.65% dry weight).

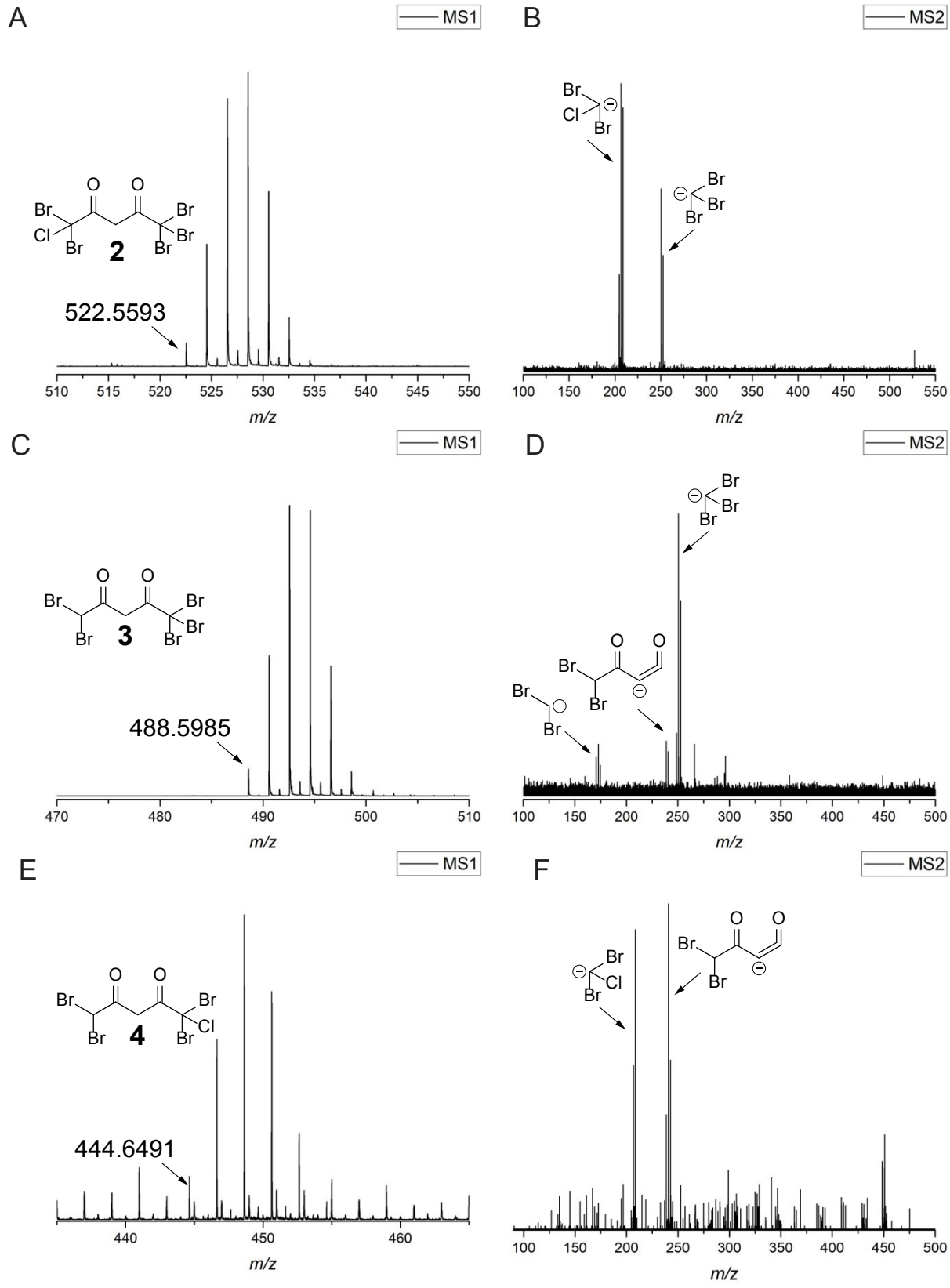


Figure S3: MS^1 and MS^2 spectra for **2**–**4**.

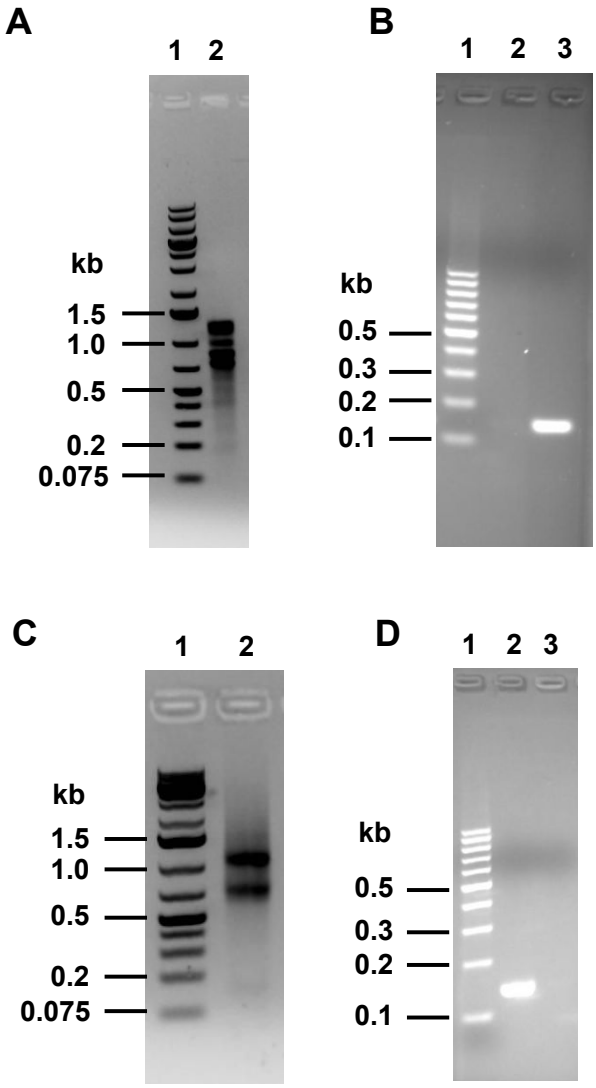


Figure S4: Total RNA and RT-PCR analysis by gel electrophoresis. Agarose gel pictures showing total RNA isolated from *A. taxiformis* collected from Guam (panel **A**, lane 2) and California (panel **C**, lane 2). RT-PCR reaction was done to amplify RuBisCo large subunit gene fragment using isolated RNA samples. Panel **B**; RT-PCR with Guam RNA in absence (lane 2) or presence (lane 3) of RT enzyme. Panel **D**; RT-PCR with California RNA in presence (lane 2) or absence (lane 3) of RT enzyme. Lane 1 in each panel corresponds to DNA ladder.

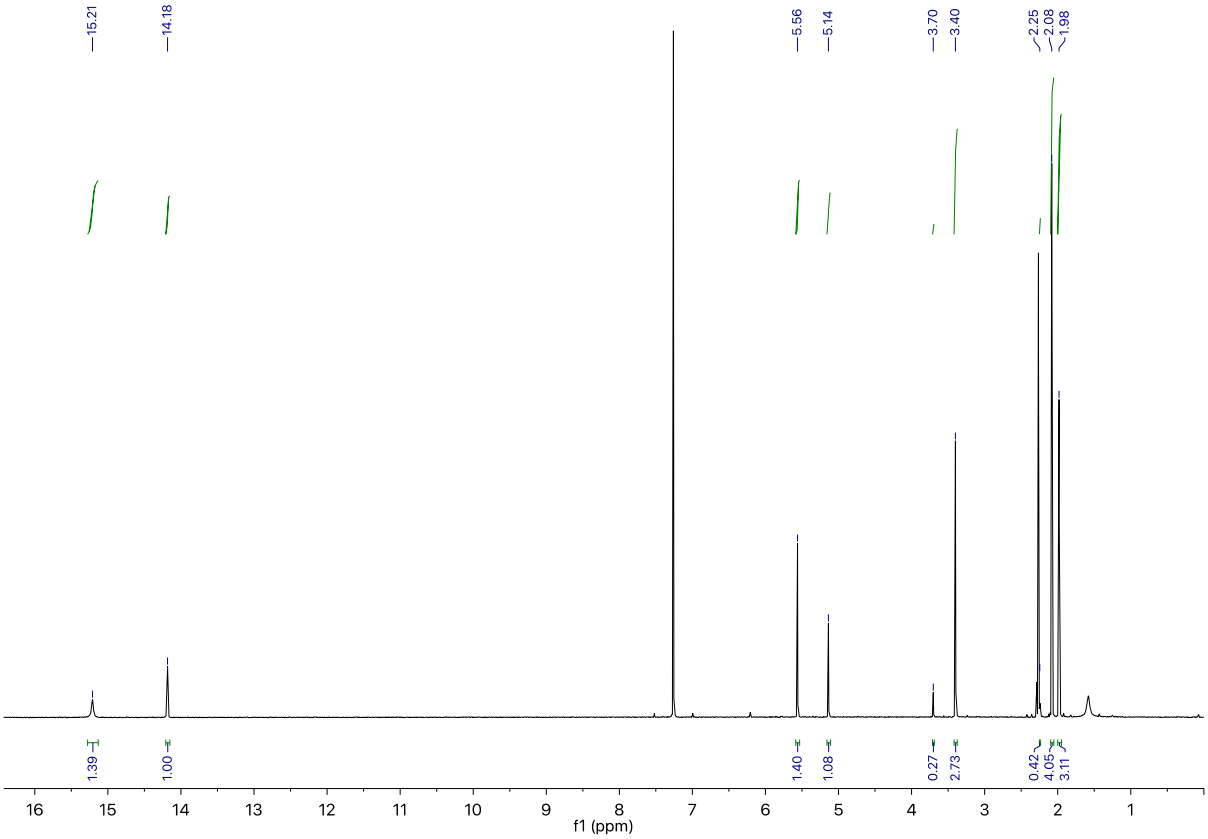


Figure S5: ^{1}H-NMR (400 MHz, $CDCl_3$) spectrum for **6**.

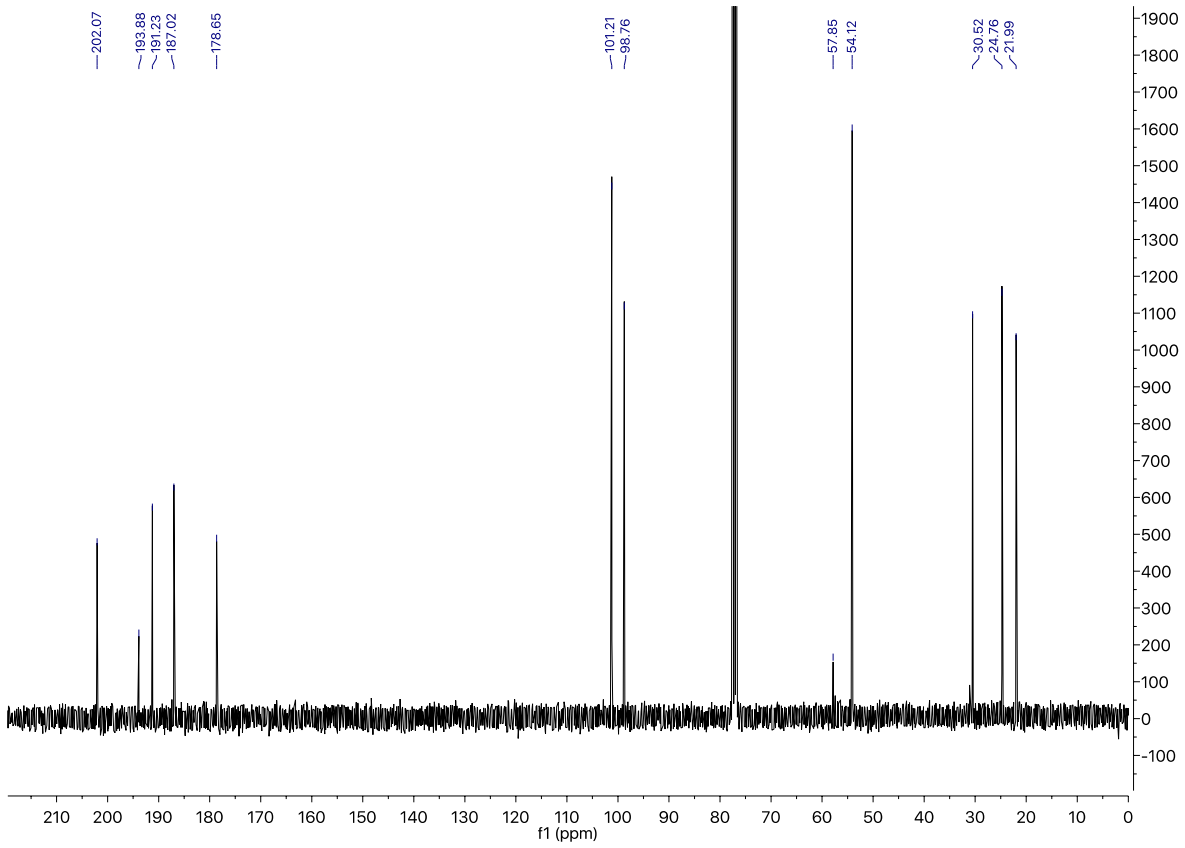


Figure S6: ^{13}C-NMR (101 MHz, $CDCl_3$) spectrum for **6**.

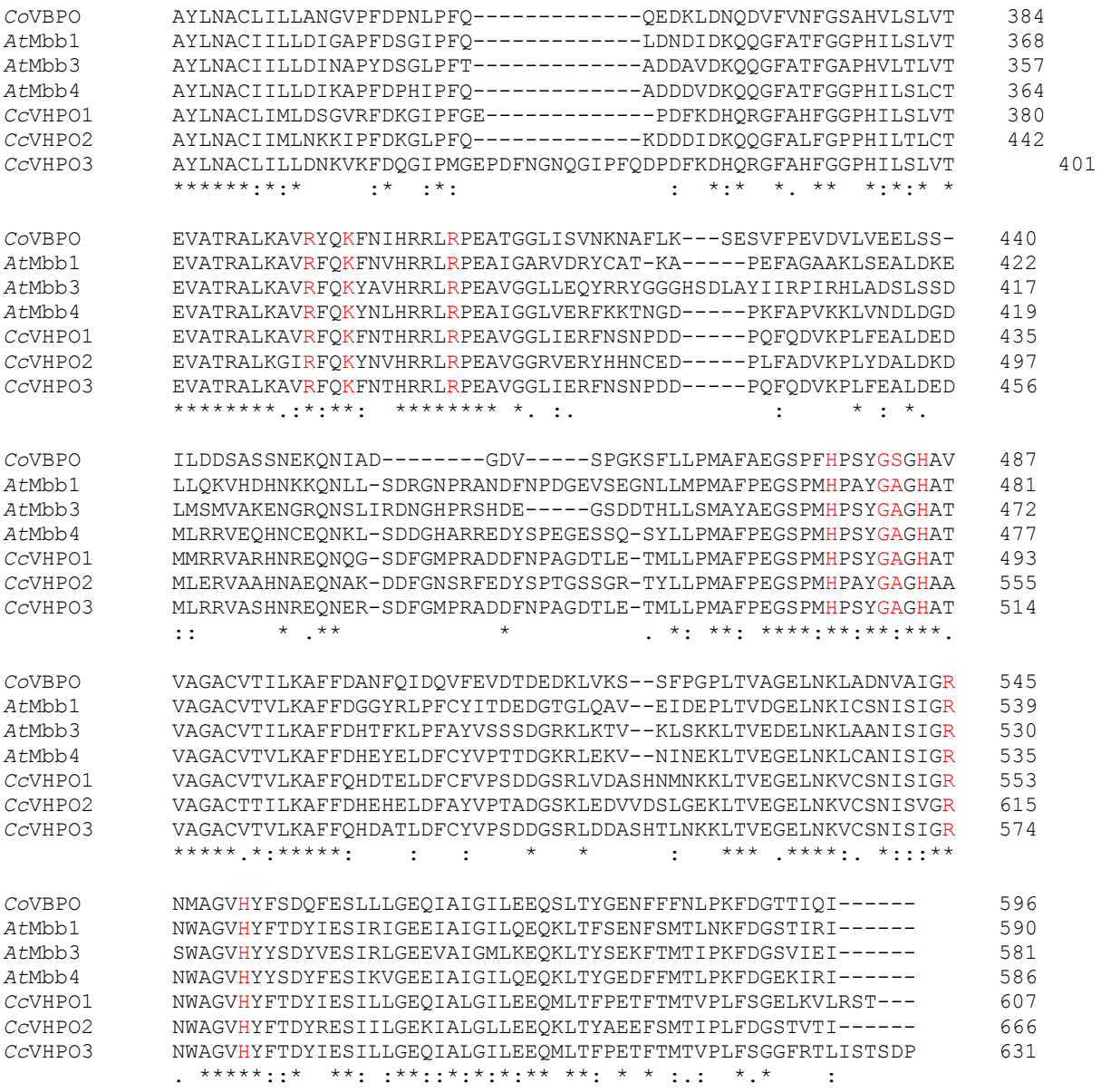


Figure S7: **Amino acid sequence alignment of VHPO and Mbb proteins from marine red macroalgal species.** Only partial amino acid sequences for each protein is shown and alignment was generate using Clustal Omega.[21] All vanadate binding residues except for Ser483 in *Co*VBPO is strictly conserved between *At*Mbb and *Cc*VHPO proteins. Accession numbers and percent identity of proteins with *Co*VBPO are shown in parenthesis: *Corallina officinalis Co*VBPO (PDB 1QHB_A; 100%), *At*Mbb1 (50.44%), *At*Mbb3 (47.53%), *At*Mbb4 (52.24%), *C. crispus Cc*VHPO1 (Genbank XP_005714237.1, 50.53%), *Cc*VHPO2 (GenBank XP_005719435.1, 47.83%), and *Cc*VHPO3 (GenBank XP_005710894.1, 50.26%).

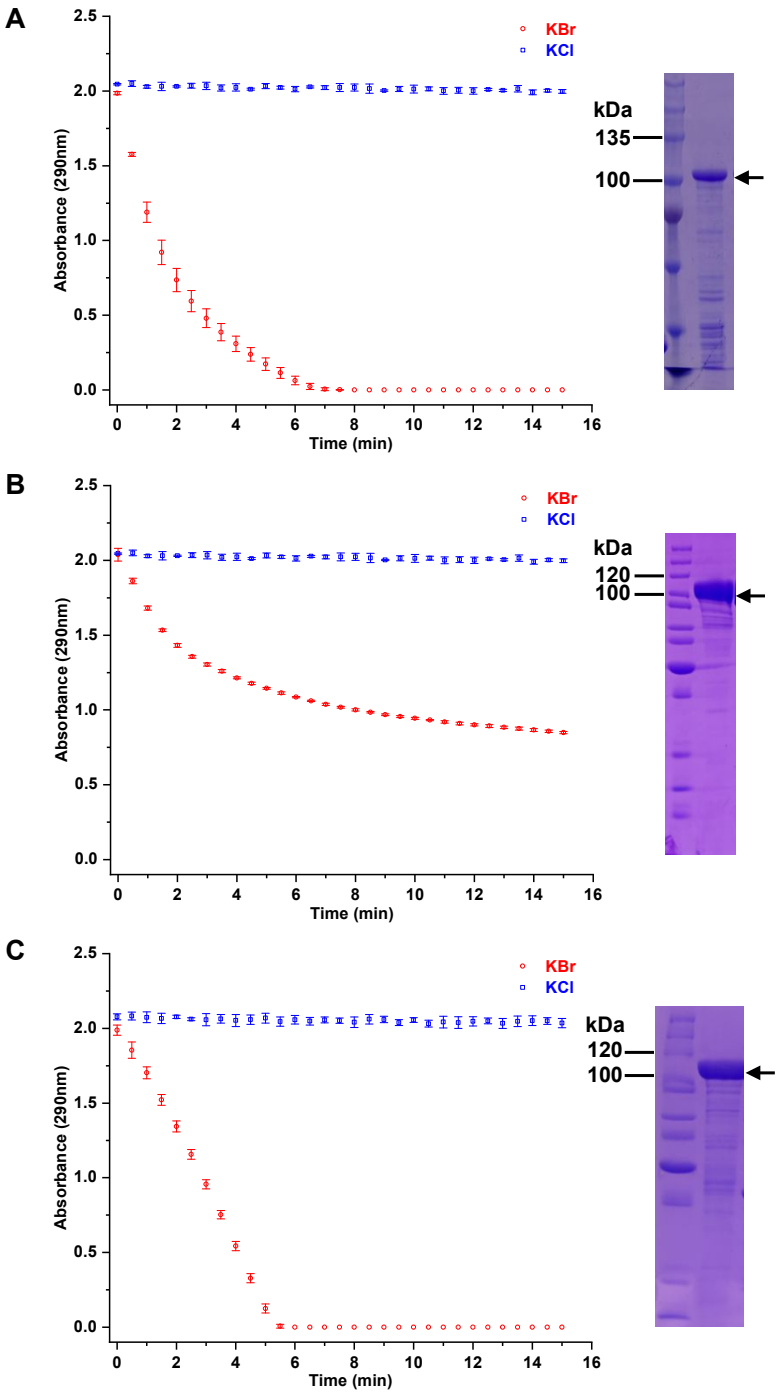


Figure S8: **Monochlorodimedone (MCD) assay with recombinant Mbb VHPOs using KBr and KCl as halide sources.** Change in absorbance of MCD substrate at 290nm over the time when incubated with Mbb enzymes in a MCD assay: (**A**) Mbb1 (250 nM), (**B**) Mbb3 (1 µM), and (**C**) Mbb4 (100 nM). Concentration for each enzyme used here was empirically determined due to difference in the activity of individual enzymes. SDS-PAGE gel on right panel of each graph demonstrates the purity of Mbb proteins: His-MBP-Mbb1 (110.6 kDa), His-MBP-Mbb3 (110.9 kDa) and His-MBP-Mbb4 (111.8 kDa).

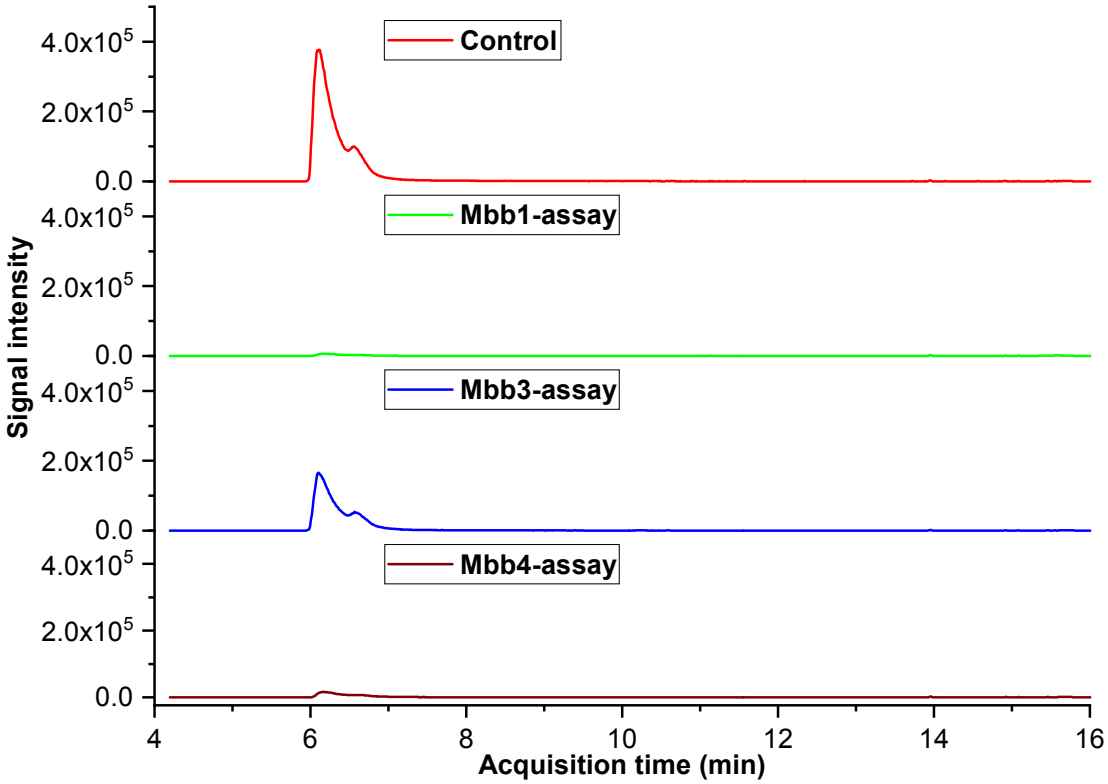


Figure S9: **GC-MS analysis of enzyme assay products showing turnover of substrate 5**. Assays were conducted with recombinant Mbb VHPOs. Extracted ion chromatograms for molecular ion *m/z* 100.05 for **5**.

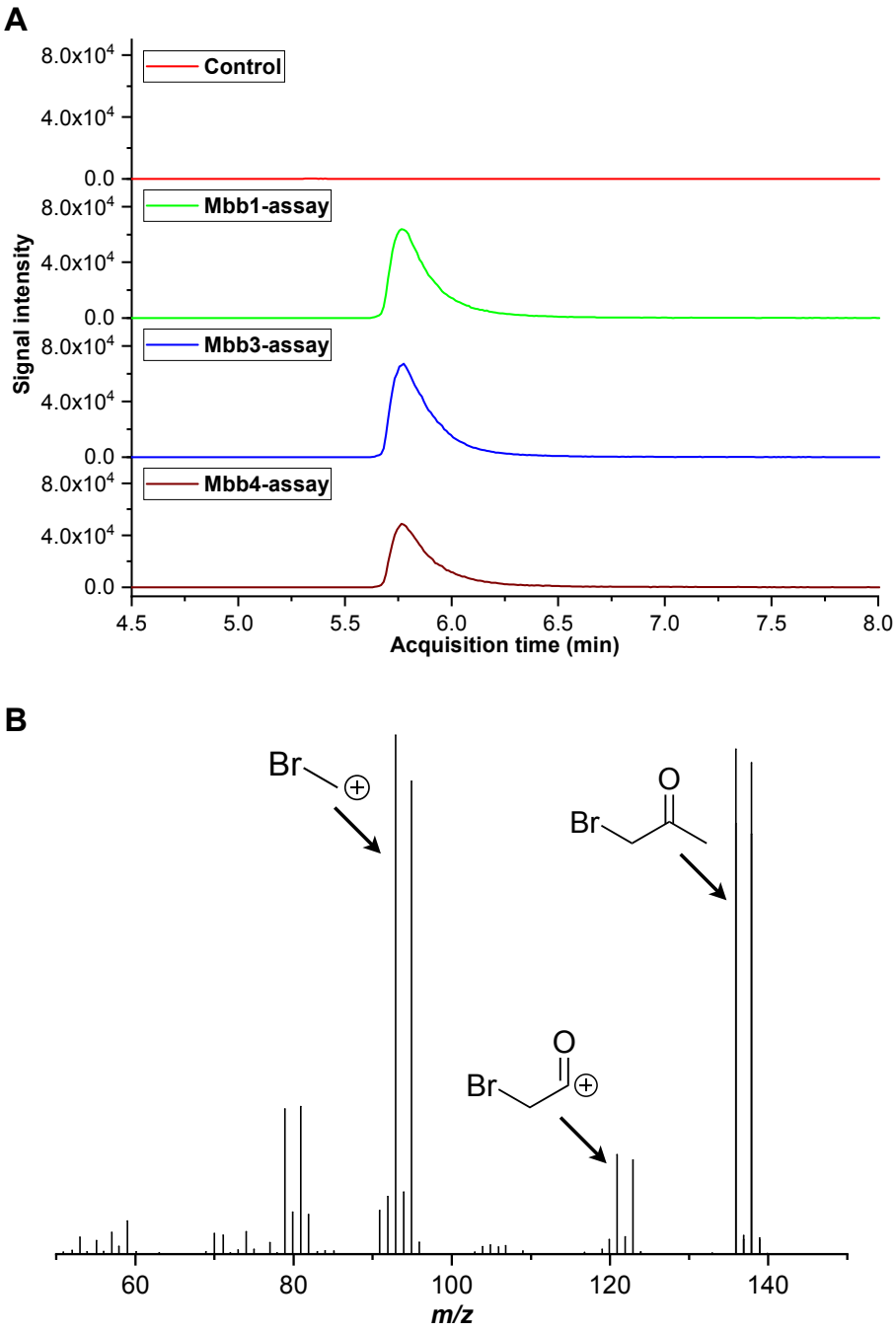


Figure S10: **GC-MS analysis of enzyme assay products.** Assays were conducted with recombinant Mbb VHPOs using **5** as substrate. (**A**) Extracted ion chromatograms for most abundant isotopic ion *m/z* 135.95 for monobromoacetone. (**B**) MS spectrum supports the structural assignment for monobromoacetone by annotation of fragment ions. Arrow corresponds to monoisotopic ion of each molecule.

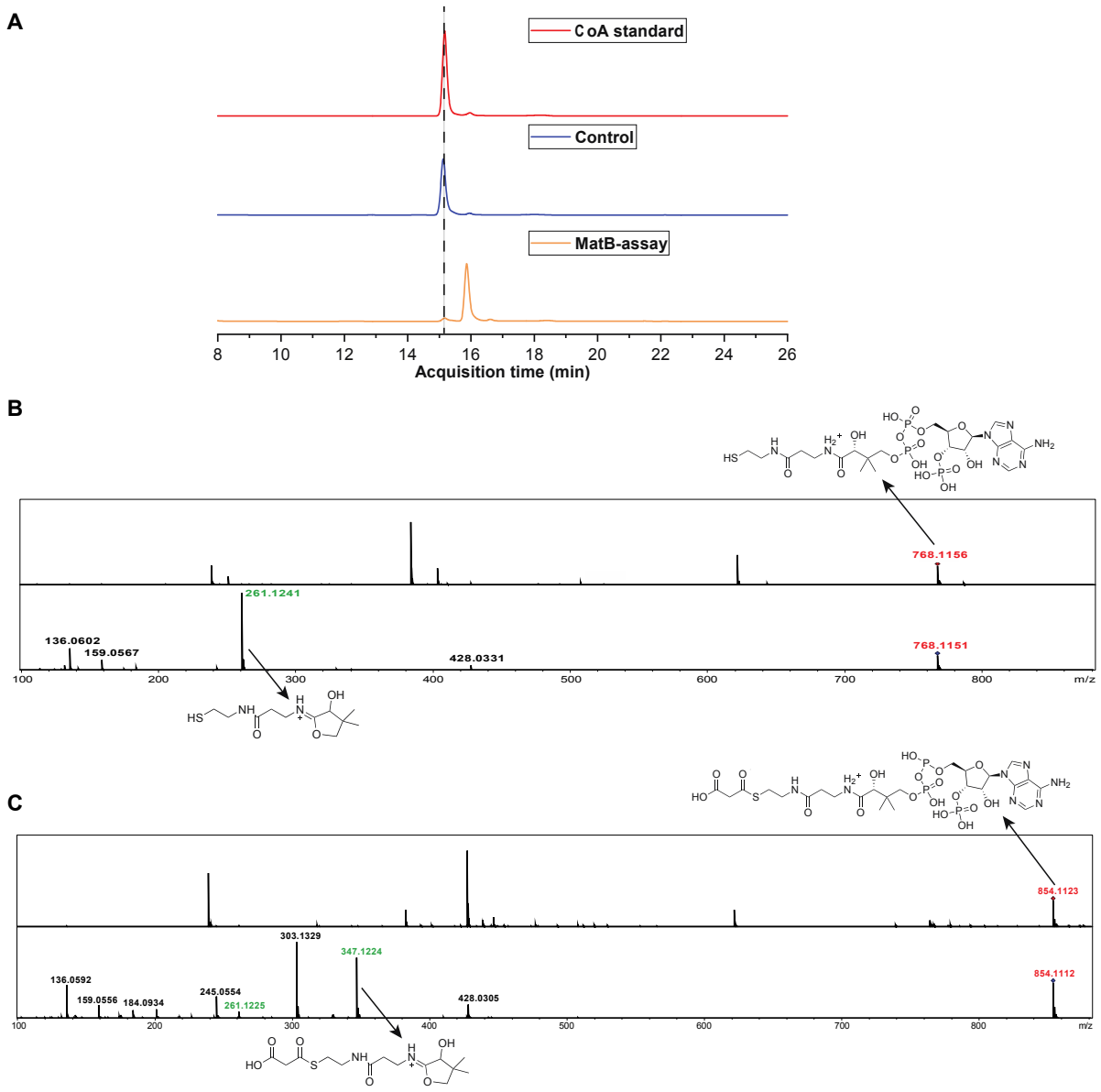


Figure S11: **Enzymatic synthesis of malonyl-CoA.** (**A**) HPLC chromatogram showing synthesis of malonyl-CoA when CoA-SH and malonic acid were used as substrates in an ATP/Mg^{+2} dependent assay with MatB enzyme.[11] Control reaction was done in the absence of MatB. (**B**) MS^1 (top) and MS^2 (bottom panel) spectra for CoA-SH. The characteristic (cyclo)pantetheine MS^2 product ion is observed in the MS^2 spectra.[22] (**C**) MS^1 (top) and MS^2 (bottom panel) spectra for malonyl-CoA. The characteristic malonyl-*S*-(cyclo)pantetheine MS^2 product ion is observed in the MS^2 spectra.

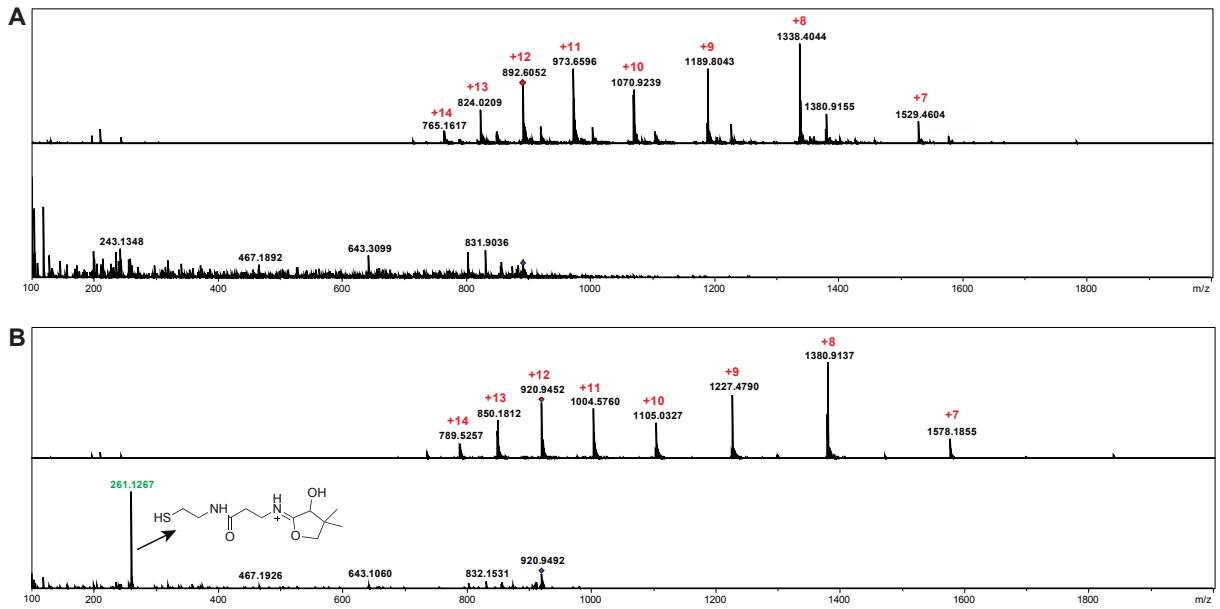


Figure S12: **Mass spectrometric characterization of apo-ACP and holo-ACP.** (**A**) MS^1 (top) and MS^2 (bottom panel) spectra corresponding to apo-ACP. Multiple charge states are observed in MS^1 spectrum. The (cyclo)pantetheine ejection ion is not observed for apo-ACP in the MS^2 spectrum. (**B**) MS^1 (top) and MS^2 (bottom panel) spectra corresponding to holo-ACP. Multiple charge states of peptides are observed in MS^1 spectrum and MS^2 spectrum of holo-ACP shows a characteristic (cyclo)pantetheine ejection ion at *m/z* 261.12.

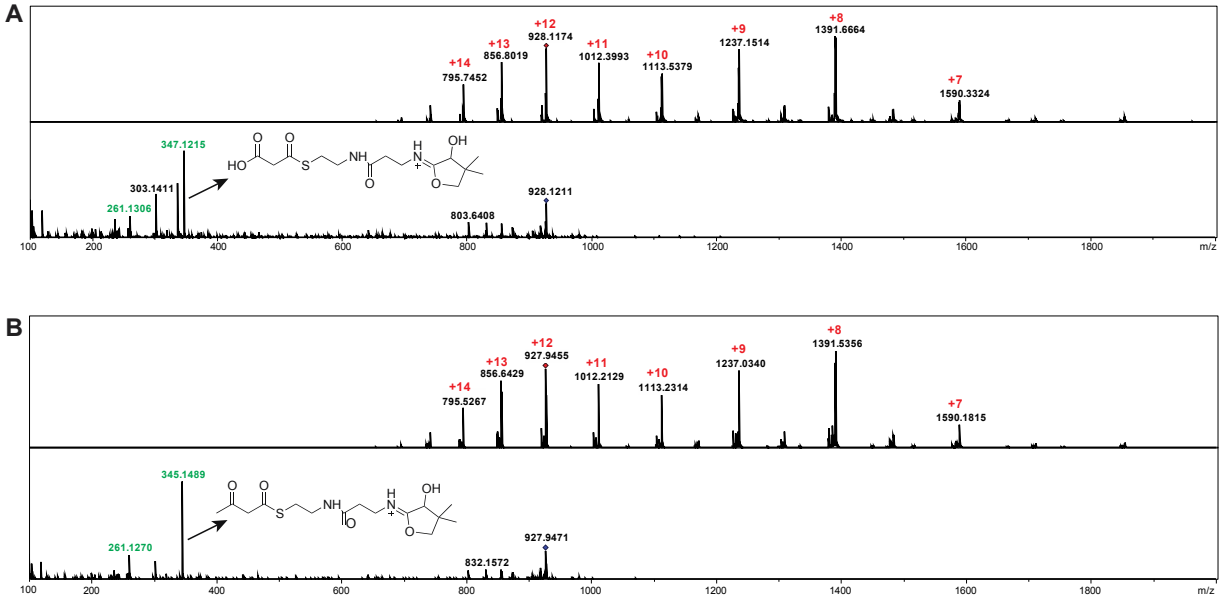


Figure S13: **Mass spectrometric characterization of malonyl-*S*-ACP and acetoacetyl-*S*-ACP (8).** (**A**) MS^1 (top) and MS^2 (bottom panel) spectra corresponding to malonyl-*S*-ACP. Multiple charge states of peptides are observed in MS^1 spectrum and MS^2 spectrum of malonyl-*S*-ACP shows a characteristic malonyl-*S*-(cyclo)pantetheine ejection ion at *m/z* 347.12. (**B**) MS^1 (top) and MS^2 (bottom panel) spectra corresponding to acetoacetyl-*S*-ACP. Multiple charge states of peptides are observed in MS^1 spectrum and MS^2 spectrum of acetoacetyl-*S*-ACP shows a characteristic acetoacetyl-*S*-(cyclo)pantetheine ejection ion at *m/z* 345.14.

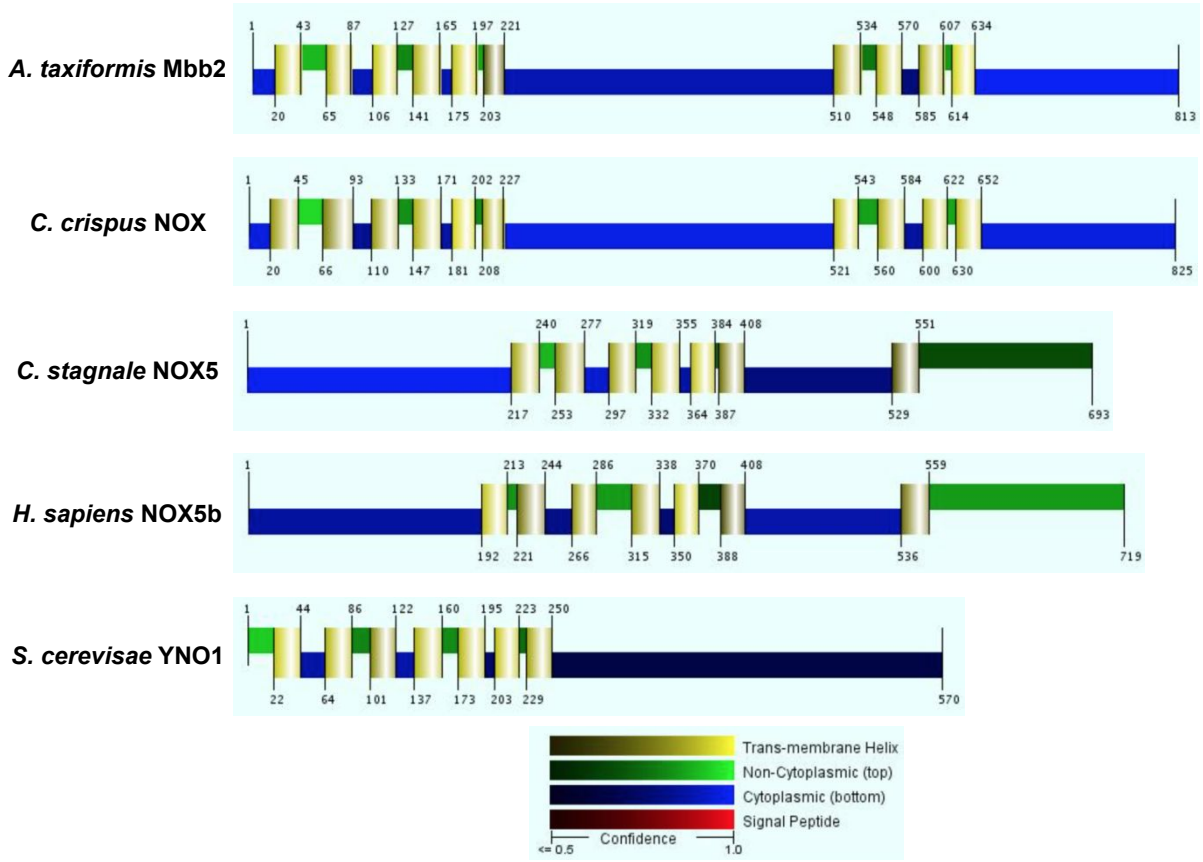
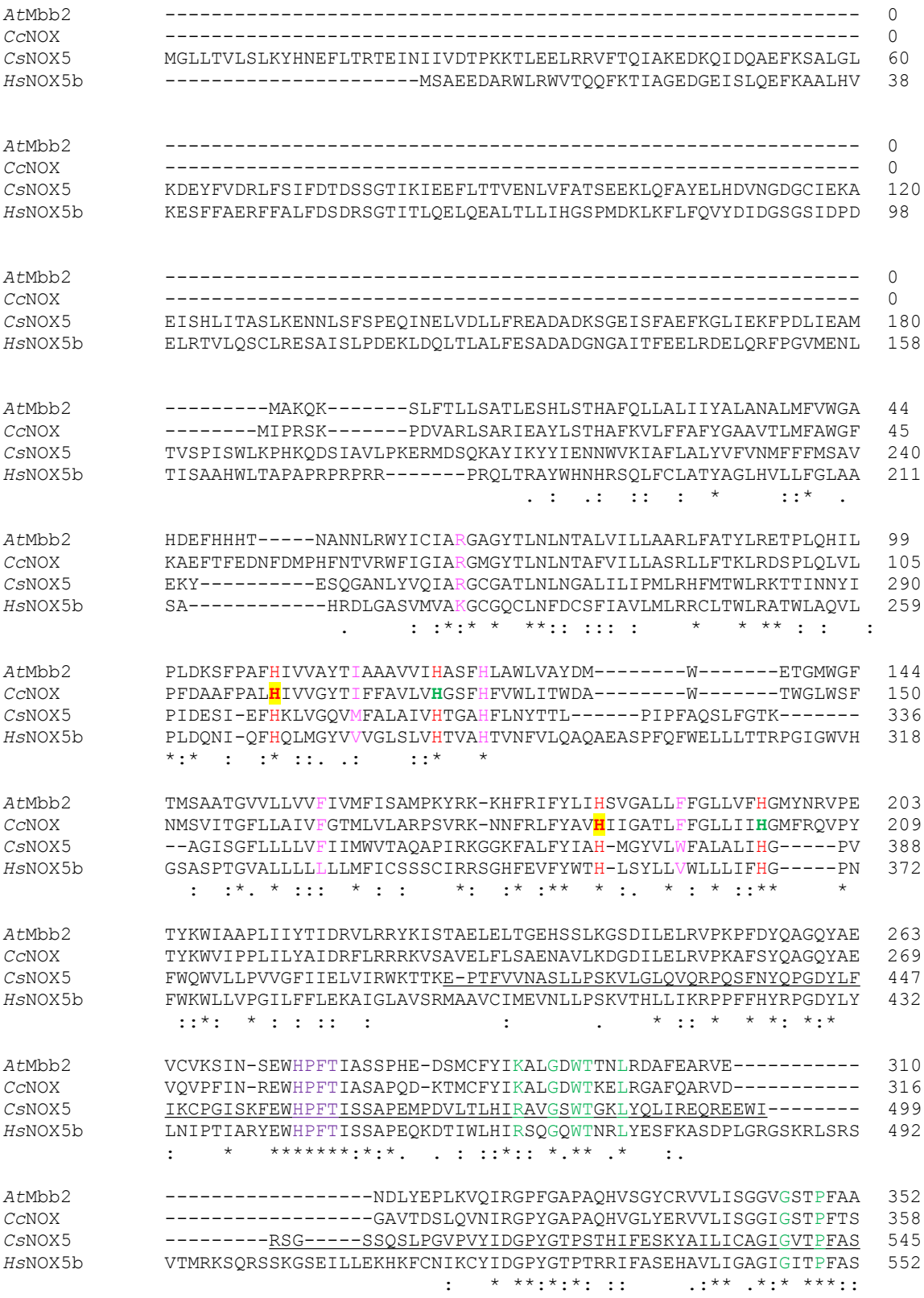


Figure S14: Transmembrane domain prediction for NOX protein sequences using the Phillius transmembrane prediction server.[23] Genbank accession numbers and percent identity of proteins with Mbb2 is shown in parenthesis: Red macroalga *Asparagopsis taxiformis* Mbb2 (100%), red macroalga *Chondrus crispus* NOX (Genbank XP_005719187; 48.81%), cyanobacterium *Cylindrospermum stagnale* (Genbank WP_015206836; 27.27%), human NOX5b (Genbank AAK57193; 30.51%), and yeast *Saccharomyces cerevisiae* YNO1 (Genbank NP_011355; 15.78 %). Human NOX5b is used for comparison as it showed highest identity to Mbb2 when compared to all seven human NOX proteins.[24] Cyanobacterial NOX5 is used for comparison as its crystal structure has been elucidated.[25] Although yeast YNO1 showed low identity to Mbb2 when compared to other NOX proteins, it is shown here to highlight the difference in organization of transmembrane domains between NOX protein sequences. Furthermore, yeast with deletion of *yno1* gene is used for *in vivo* characterization of *Atmbb2* and *CcNOX* genes.


AtMbb2 ---------- 0
CcNOX ---------- 0
CsNOX5 MGLLTVLSLKYHNEFLTRTEINIIVDTPKKTLEELRRVFTQIAKEDKQIDQAEFKSALGL 60
HsNOX5b ----------MSAEEDARWLRWVTQQFKTIAGEDGEISLQEFKAALHV 38
AtMbb2 ---------- 0
CcNOX ---------- 0
CsNOX5 KDEYFVDRLFSIFDTDSSGTIKIEEFLTTVENLVFATSEEKLQFAYELHDVNGDGCIEKA 120
HsNOX5b KESFFAERFFALFDSDRSGTITLQELQEALTLLIHGSPMDKLKFLFQVYDIDGSGSIDPD 98
AtMbb2 ---------- 0
CcNOX ---------- 0
CsNOX5 EISHLITASLKENNLSFSPEQINELVDLLFREADADKSGEISFAEFKGLIEKFPDLIEAM 180
HsNOX5b ELRTVLQSCLRESAISLPDEKLDQLTLALFESADADGNGAITFEELRDELQRFPGVMENL 158
AtMbb2 ---------MAKQK-------SLFTLLSATLESHLSTHAFQLLALIIYALANALMFVWGA 44
CcNOX --------MIPRSK-------PDVARLSARIEAYLSTHAFKVLFFAFYGAAVTLMFAWGF 45
CsNOX5 TVSPISWLKPHKQDSIAVLPKERMDSQKAYIKYYIENNWVKIAFLALYVFVNMFFFMSAV 240
HsNOX5b TISAAHWLTAPAPRPRPRR-------PRQLTRAYWHNHRSQLFCLATYAGLHVLLFGLAA 211
AtMbb2 HDEFHHHT-----NANNLRWYICIARGAGYTLNLNTALVILLAARLFATYLRETPLQHIL 99
CcNOX KAEFTFEDNFDMPHFNTVRWFIGIARGMGYTLNLNTAFVILLASRLLFTKLRDSPLQLVL 105
CsNOX5 EKY----------ESQGANLYVQIARGCGATLNLNGALILIPMLRHFMTWLRKTTINNYI 290
HsNOX5b SA----------HRDLGASVMVAKGCGQCLNFDCSFIAVLMLRRCLTWLRATWLAQVL 259
AtMbb2 PLDKSFPAFHIVVAYTIAAAVVIHASFHLAWLVAYDM--------W-------ETGMWGF 144
CcNOX PFDAAFPALHIVVGYTIFFAVLVHGSFHFVWLITWDA--------W-------TWGLWSF 150
CsNOX5 PIDESI-EFHKLVGQVMFALAIVHTGAHFLNYTTL------PIPFAQSLFGTK------- 336
HsNOX5b PLDQNI-QFHQLMGYVVVGLSLVHTVAHTVNFVLQAQAEASPFQFWELLLTTRPGIGWVH 318
AtMbb2 TMSAATGVVLLVVFIVMFISAMPKYRK-KHFRIFYLIHSVGALLFFGLLVFHGMYNRVPE 203
CcNOX NMSVITGFLLAIVFGTMLVLARPSVRK-NNFRLFYAVHIIGATLFFGLLIIHGMFRQVPY 209
CsNOX5 --AGISGFLLLLVFIIMWVTAQAPIRKGGKFALFYIAH-MGYVLWFALALIHG-----PV 388
HsNOX5b GSASPTGVALLLLLLLMFICSSSCIRRSGHFEVFYWTH-LSYLLVWLLLIFHG-----PN 372
AtMbb2 TYKWIAAPLIIYTIDRVLRRYKISTAELELTGEHSSLKGSDILELRVPKPFDYQAGQYAE 263
CcNOX TYKWVIPPLILYAIDRFLRRRKVSAVELFLSAENAVLKDGDILELRVPKAFSYQAGQYAE 269
CsNOX5 FWQWVLLPVVGFIIELVIRWKTTKE-PTFVVNASLLPSKVLGLQVQRPQSFNYQPGDYLF 447
HsNOX5b FWKWLLVPGILFFLEKAIGLAVSRMAAVCIMEVNLLPSKVTHLLIKRPPFFHYRPGDYLY 432
AtMbb2 VCVKSIN-SEWHPFTIASSPHE-DSMCFYIKALGDWTTNLRDAFEARVE---------- 310
CcNOX VQVPFIN-REWHPFTIASAPQD-KTMCFYIKALGDWTKELRGAFQARVD---------- 316
CsNOX5 IKCPGISKFEWHPFTISSAPEMPDVLTLHIRAVGSWTGKLYQLIREQREEWI-------- 499
HsNOX5b LNIPTIARYEWHPFTISSAPEQKDTIWLHIRSQGQWTNRLYESFKASDPLGRGSKRLSRS 492
AtMbb2 ------------------NDLYEPLKVQIRGPFGAPAQHVSGYCRVVLISGGVGSTPFAA 352
CcNOX ------------------GAVTDSLQVNIRGPYGAPAQHVGLYERVVLISGGIGSTPFTS 358
CsNOX5 ---------RSG-----SSQSLPGVPVYIDGPYGTPSTHIFESKYAILICAGIGVTPFAS 545
HsNOX5b VTMRKSQRSSKGSEILLEKHKFCNIKCYIDGPYGTPTRRIFASEHAVLIGAGIGITPFAS 552

```
AtMbb2        ICKHLHHLNKSENHSKALESHA-SSK-----RMS--QVQWRIRDAISILFDVSLDDSRND  404
CcNOX         ICKDLHHHKVKENATSATGFEPSTST-----LLK--RIESRVSTAISTLYGVDISNAKDI  411
CsNOX5        ILKSILHRNQQNP---------------AKMPLKKVHFYWLNREQKAFEWFVELLSKIEA  590
HsNOX5b       ILQSIMYRHQKRKHTCPSCQHSWIEGVQDNMKLHKVDFIWINRDQRSFEWFVSLLTKLEM  612
              * : : : : ..                    :    .       :  : *.:    :

AtMbb2        EAANQQRRQQLADMLNMSPKGDVHNVSFKRTQSKRSDDFDSVEDSKQMNF----------  454
CcNOX         NQEEEEKRVYLANMLNLTAPGSGSSSGETTE-----LEVEMVDASKQADESSSDSDRSTS  466
CsNOX5        EDT-----------------------------------------------------------  593
HsNOX5b       DQAEEA--------------------------------------------------------  618
              :

AtMbb2        -------SLCRHDSKSSFNLDDCSDAIRYYNRPSFQFFSLSTRHVINLYEYRTRLLAFLH  507
CcNOX         MESYNVKNMLRKEQDEEYILDDIKNSAN--------ARRGNRERLSHLYEGRSKVLEFLH  518
CsNOX5        -----------------------------------------------------------  593
HsNOX5b       -----------------------------------------------------------  618

AtMbb2        TTRFTFALLLTLVARIVILCIVSIFNLGHIGLYN---SHIDATWVIATNSILGLILGVAL  564
CcNOX         TSRVNLLLLFVLIARIFFICISSIIKADYIMINAEPHAIESGLWIVIVDTVLSIIFAIVL  578
CsNOX5        -----------------------------------------------------------  593
HsNOX5b       -----------------------------------------------------------  618

AtMbb2        LTTILLEISFMRMRFFYRVWRCVDFFIFLPITFLCNISNFAS-WNGHAQPKFLIFLDLII  623
CcNOX         PLTIFLELSYMGSRFFRTVGRTLDFFVFLPLTITSASLGIKALVTERTDEQIVLFLHYIV  638
CsNOX5        ----------------------------------------------------NNLFDLNL  601
HsNOX5b       ----------------------------------------------------QYGRFLELHM  628
                                                                   ::.  :

AtMbb2        VLPIMLLLLCHRMYRSVGSRNLLDDTAECRQGCKCNKTIPDVDFVWTTPRDSDDEWLRNE  683
CcNOX         FLPTLFVLLAVRMYRALGKRTLLTDAP---CHCSHRDIVPNVDFVWTVPHENDDEWLRSE  695
CsNOX5        ----------YLTGAQQKSDMKSSTL----------FVAMD-----------------  622
HsNOX5b       ----------YMTSALGKNDMKAIGL----------QMALD-----------------  649
                         :  :  .  :                 :*

AtMbb2        LYPLATGTELRLHRYVTRENMADLEDPERFITTANAGRPQWDEIFAQIAEQTPSHSKIGV  743
CcNOX         LEPLADGTELKLHRYVTRAKEVDMEAGSEFITSSNTGRPEWDEIFGKIAAEAPSNSVVGV  755
CsNOX5        --L----------MH--QETKVDLITGLKSRTKTGRPDWEEIFKDVAKQHA-PDNVEV  665
HsNOX5b       --L----------LA--NKEKKDSITGLQTRTQPGRPDWSKVFQKVAAEK--KGKVQV  691
                           :   :  . :    : : :: ***:*.::* .:* :    . : *

AtMbb2        FFCGPHPMGAAVQKSMRKVEVMSNLRGSYLRKTESAVLVDDLILRDEGEVKLLREYGCNI  803
CcNOX         FFCGPHKMGDSVQSAMRRAEINSNLRGAYLRSTKEKTLMKDLGLPQRGLIKMLMGTGCSV  815
CsNOX5        FFCGPTGLALQ-----------------------------------------LRHLCTKY  684
HsNOX5b       FFCGSPALAKV-----------------------------------------LKGHCEKF  710
              ****   :.                                             *     .

AtMbb2        RFVFREENFS 813
CcNOX         RFVFREENFG 825
CsNOX5        GFGYRKENF- 693
HsNOX5b       GFRFFQENF- 719
               * : :***
```

Figure S15: Amino acid sequence alignment of NOX protein sequences using Clustal Omega.[21] Four NOX protein sequences, *Cylindrospermum stagnale* NOX (*Cs*NOX), *Asparagopsis taxiformis* NOX (*At*Mbb2), *Chondrus crispus* NOX (*Cc*NOX) and Human NOX5b, were used for alignment. Yeast NOX was excluded in alignment study due to its low amino acid identity to other NOX proteins. Residues numbers described

here correspond to *Cs*NOX protein and their structure to function relationship is inferred from study describing the crystal structures of transmembrane domain and dehydrogenase domain of *Cs*NOX protein.[25] Four histidine residues (His313 and His385 ligated to phorphyrin of extracytoplasmic heme, highlighted in green, and His299 and His372 ligated to phorphyrin of cytoplasmic heme, highlighted in red) involved in heme binding in *Cs*NOX5 protein are conserved in *At*Mbb2, *Cc*NOX and *Hs*NOX5b proteins. The two His residues mutated to Ala (data for loss of NOX activity shown in Figure 4B) are highlighted in bold. Similarly, other hydrophobic residues (Met306, Phe348 and Trp378) that intercalates between two heme groups and dioxygen substrate binding residues (His317 and Arg256) in *Cs*NOX5 are also conserved among other NOX proteins, and are highlighted in magenta letters. Residues corresponding to dehydrogenase domain of *Cs*NOX is underlined. The NADPH and FAD binding residues are highlighted in green and purple letters, respectively.

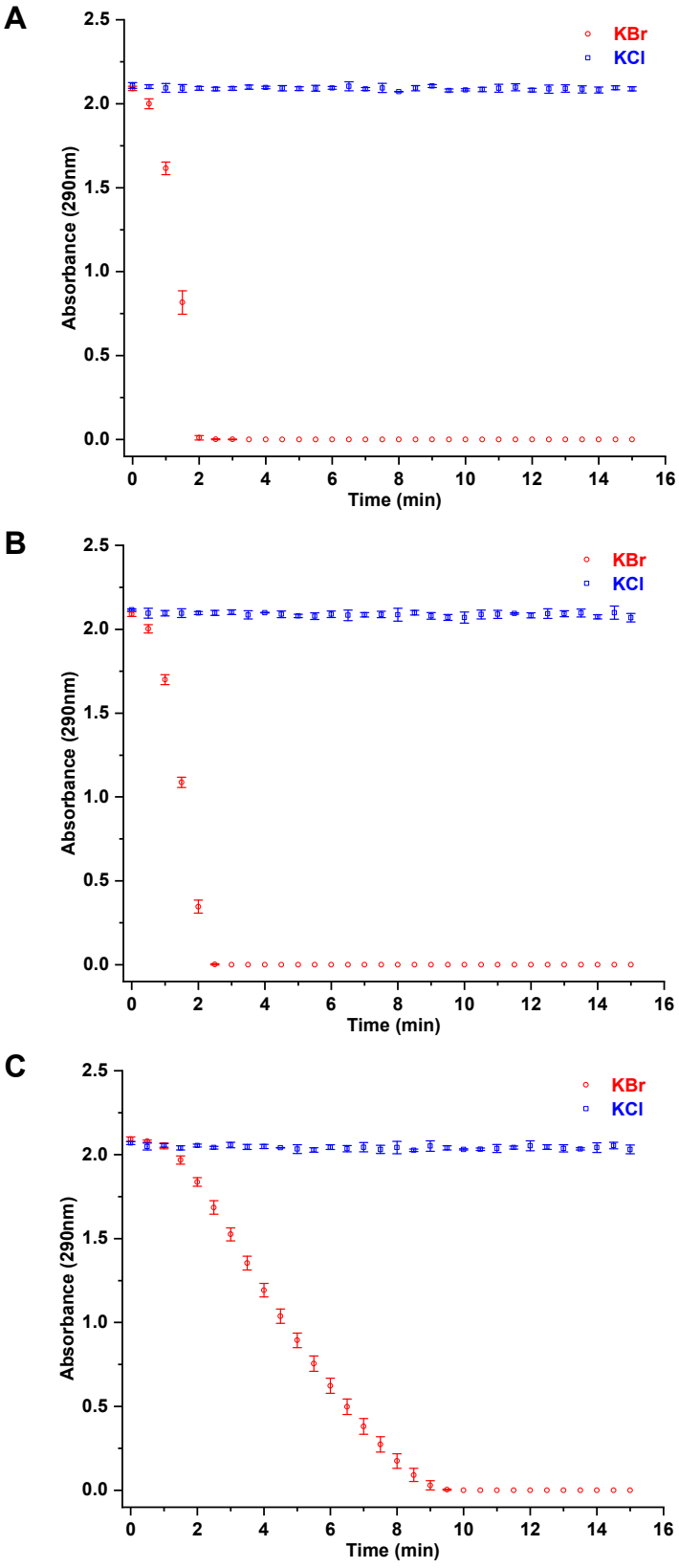


Figure S16: **Monochlorodimedone (MCD) assays with recombinant *Cc*VHPOs**. Change in absorbance of MCD at 290nm was recorded over the time in assays with (**A**) *Cc*VHPO1 (50 nM), (**B**) *Cc*VHPO2 (50 nM), and (**C**) *Cc*VHPO3 (250 nM). Two different halide sources, KBr and KCl, were tested in separate assays, and the enzyme concentrations used here were empirically determined.

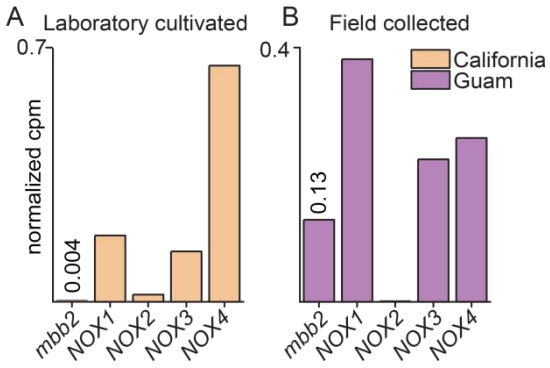
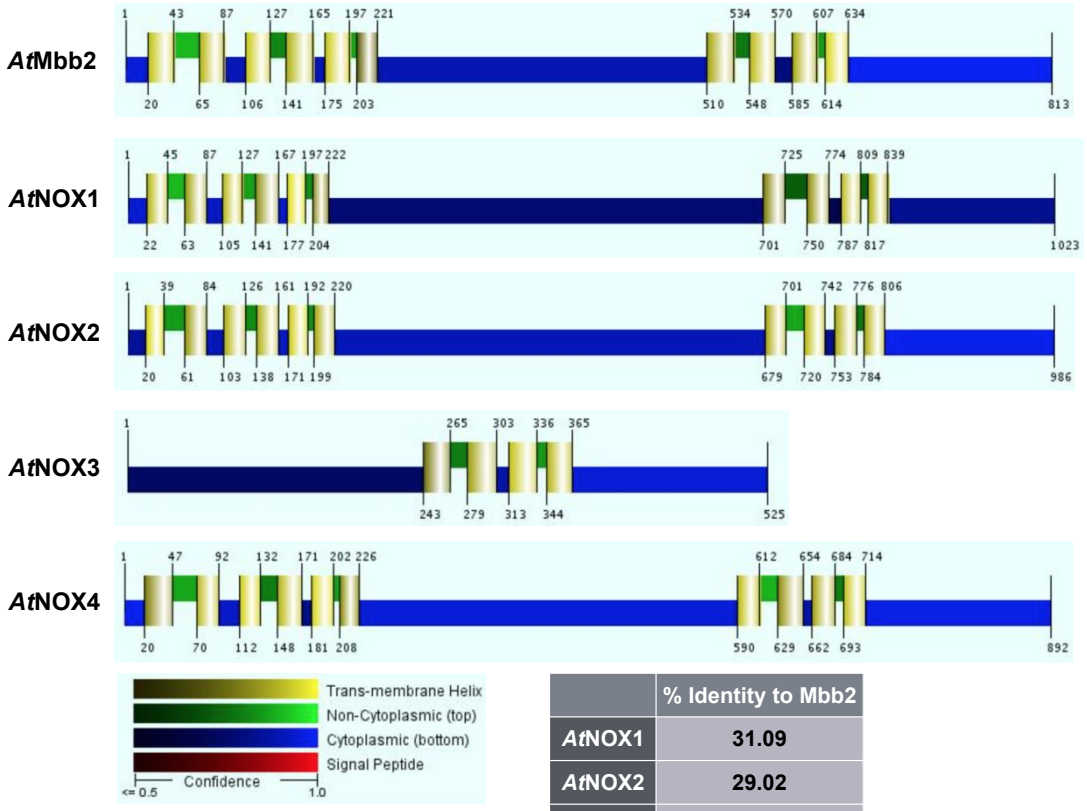


Figure S17. Individually normalized counts per million (cpm) transcript levels for NOX encoding genes in (**A**) laboratory cultivated (California) and (**B**) field collected (Guam) *A. taxiformis* samples.



	% Identity to Mbb2
*At*NOX1	31.09
*At*NOX2	29.02
*At*NOX3	25.32
*At*NOX4	58.37

Figure S18: Transmembrane domain prediction for *Asparagopsis taxiformis* NOX-like protein sequences using the Phillius transmembrane prediction server.[23] Percent identity of *At*NOX proteins to *At*Mbb2 is shown in table, and was determined based on protein sequence alignment generated using Clustal Omega.[21] GenBank accession numbers is shown in parenthesis: *At*NOX1 (GenBank MN911452), *At*NOX2 (GenBank MN911451), *At*Mbb2 (GenBank MN893468), *At*NOX3 (GenBank MN911453) and *At*NOX4 (GenBank MN911454).

SUPPLEMENTARY TABLES

Supplementary Table S1. Shortgun sequencing and assembly results.

	trimmed reads (pair)	contig #	n50 (bp)	longest contig (bp)	Total assembly size (bp)	Sequencing method	Library size	GC%	Read length (bp)
California gDNA	85,676,804	1104938	2261	2019246	1,517,284,223	Illumina NovaSeq	~450	51	151
Guam gDNA	105,547,008	980759	773	744171	765,833,334	Illumina NovaSeq	~450	46	151
California mRNA	44,450,663	20789	3018	20984	35,660,953	Illumina NovaSeq	~450	50	151
Guam mRNA	98,748,611	151422	783	38841	113,504,178	Illumina HiSeq 2000	~350	47	125

Supplementary Table 2: Amino acid and nucleotide sequence identity between Mbb proteins and *mbb* genes from Guam and California samples. Genes *mbb1*, *mbb3* and *mbb4* encodes for VHPO proteins.

	Amino acid similarity (%)	Nucleotide similarity (%)
Mbb1	96.949	96.221
Mbb2	98.770	97.052
Mbb3	96.558	95.246
Mbb4	96.587	95.060

Supplementary Table 3: List of *E. coli* expression constructs used in this study.

Gene	Plasmid	Fusion tag	Source organism	Reference
mbb1	pET-28MBP[26]	N-terminal 6xHis-MBP	*A. taxiformis* (Guam)	This study
mbb3	pET-28MBP	N-terminal 6xHis-MBP	*A. taxiformis* (Guam)	This study
mbb4	pET-28MBP	N-terminal 6xHis-MBP	*A. taxiformis* (Guam)	This study
acp	pET-28b	N-terminal 6xHis	*A. taxiformis* (Guam)	This study
fabH	pET-28b	N-terminal 6xHis	*A. taxiformis* (Guam)	This study
CcVBPO1	pET-28MBP	N-terminal 6xHis-MBP	*C. crispus*	This study
CcVBPO2	pET-28MBP	N-terminal 6xHis-MBP	*C. crispus*	This study
CcVBPO3	pET-28MBP	N-terminal 6xHis-MBP	*C. crispus*	This study
matB[11]	pET-28MBP[26]	N-terminal 6xHis-MBP	*Streptomyces coelicolor*	This study
sfp	pET-24b	C-terminal 6xHis	*Bacillus subtilis*	Agarwal et al. (2014)[27]; Quadri et al. (1998)[13]

Supplementary Table 4: **List of primers used for qRT-PCR**

Gene	Sample	qPCR-forward primer	qPCR-reverse primer
actin	Guam	GCATGACCCAGATCATGTTCGAG	CATAGATAGGGACGGTATGGGTGAC
mbb1	Guam	GAAGTTGGCGATGGATTCATCCAG	CTTCTTTGTACAATTCGCGACCGG
mbb2	Guam	GTGAGTGGCATCCGTTTACAATTG	CTTGAGCGGTTCGTACAGATCATTC
mbb3	Guam	GTCCAACAGATGCTGTTTGCCTC	GTGTACATCGGCGAGACCCTTG
mbb4	Guam	GAGATTACGGAGAGCTATTGGATGGC	CTTCATCGGTCAATGTTGACGGAC
actin	California	GCATGACCCAGATCATGTTCGAG	CATAGATAGGGACGGTATGGGTGAC
mbb1	California	GCAGAGTCCATCAATAACACCCAATG	CGACAAGGAGGAACTGTGACAGATAG
mbb2	California	GTGAGTGGCATCCGTTTACAATTG	GTGGTTCGTACAGATCATTCTCAACTCG
mbb3	California	CTGTGCACGACATGAATCATACCC	CTTTGTGCCGACAAGGAGAAACTG
mbb4	California	GAGATTACGGAGAGCTATTGGATGGC	CTTCATCGGTCAATGTTGACGGAC

All primers are listed from 5’ to 3’.

Supplementary Table 5: List of heme-peroxidase genes and their corresponding transcript level (counts per million) in California and Guam transcriptomes. Amino acid sequences for each heme-peroxidase gene is listed below.

	California cpm	Guam cpm
heme-peroxidase1	19,036,838	96,857,763
heme-peroxidase2	22,132,794	716,554,924
heme-peroxidase3	31,815,891	43,638,708
heme-peroxidase4	296,421	162,957,571
heme-peroxidase5	430,173,245	57,252,885

>Heme-peroxidase1
MSLIIHLLILSLLVVRHPAAAFPLCPARRTFSGACSSLFNPSWGAINTPHRRLIASPIPLPSDLPSPRLISNILCRHNDVKSTRYL
NELTTFFGQFVDHNIVLTPADGPLFPIPVPPSDPLFANFSNGLPFHRSRRVQAFTDHVFSERGPPDVFVPINVVSSVLDLSAVYGP
TISRVRALRGNGGKLKTSHNGRFLPFNHPHLDNEPTTGNIFFVAGDTRSNEHPVLTALHTLFVREHNDLVDELASKFSNYSDDWLF
QTARAINIAQFQKITLEHFYPAMTGRGLPPSSLFSDKYVDASIIDVFATAAFRVGHTMVNDVISPRDKHGPLASIKMSETFFQPGE
HMVHTDIDRFLRGAAWVRAQQVDLAVVDALRNHLFTSVRGEEGVDLVAMNIQRSRDNGLPSYNDIREAVGIPRASCFANISRKVSV
QTALSTAYGSVDRVEAWIGMVAEDHAPGAAMGETLIAVWELQFRKIRDGDRLYFRKPGLFPDVVTDQIARVRDLFYDSDTLRNIVL
RNTDIEDDELPRRMFFV

>Heme-peroxidase2
MISATILSLLLAASLTAALPSPAVHQDALTRVNCEISARFFDGTCTNPSNPAYGQATEAVFSYIPNLSSATFSSRGRPDARLISNA
VCDQDSPVFNDRNLNELLVFFGQFLDHDLVLSEAGDETVPIPVPADDPNTDVSSLPFRRNIRVSVPGEAPGTVRPENLLSQIIDLS
MVYGSDSGRANALRTFEEGRLKVSSDNLLPLNTEGLSNEPSTSENFFVAGDIRSNETPMLTVLHTIWVREHNNIIDELVQSRTFNR
LSDDGLYELARRINIAQFQKVVWEEFFPAIVGRSLPRYRGFNPNVNPTVSNIFATAAFRVGHTMVGDGVQVVQGSTRTVLPVTEMF
FREQSFFRSTGLDSFLLGAVTTRCQEIDNKVVNVLRNGLFSSIPGVSGFDLISLNIQRGRDHNIPSYNEVRELFGIERAQSFADIT
SDTDVQARLASVYDDIDDVDAFIGLISEEHVSGGSLGVTMRAIWTTEFTRLRDGDQFFYLRNGLFSSPLRREIRRVRMLRGRRGIF
RDIVLRNTGLTPSQVPRRPFFV

>Heme-peroxidase3
MTATPSNLLRPIFLVALLAAFATQAFSQAARGNGFGASVLQPIEPGVPEFRERKTDPCRDDYRTYDGSCNNKRMKLWGSAGLPHFS
YLPLRLSTKPKGRNLKSARQISNILSKQTTDIFNSRCLSEFFVFFGQFIDHTFAATPVENTKEFPIKIPADDPIFANFSGGVLPFE
RSRRGRVAGGLADLPINSVTSFLDLSSVYGSDDIRIQKLRTYKNGRMRTTKGNLLPLNTDSLRNAPTNGPMFFAAGDHRANEHPML
TSIHTLFVREHNSLADELRAKFPQWDDERLFQTARKIAIAEFQNIVFEGFFPAMTGRKLWEYRGYKRNVNPTLSDEFVTAAFRVGH
TMVGNEVKRAGPNNTPLSPISMKKMFFQPHTVMSRGVEQFLRGAIITPCQEVDVHVRNSLRDFLFTDIKEEKGFDLVALNLQRGRD
HGLPTYNELRVRFGRPPAFRFSEITRKRNLQSALANAYGNPNKVEAWIGLMCEDHIKGASIGKTLYRIWRREFRRMRAGDRFFYMV
PGLFEKEVRDKIQRVQDLFTDKDIMKGILLRNTKLTSEEIGESVWKADRCLANSN

>Heme-peroxidase4
MKFNLKFSFLSILALLSTSFALPTRLNTNVRQFNCQYQTRMLDGTCTNNIQPELGSTGRAVSSLLRNRSSKRPSSAKTTLPSARFI
SNTVSKQDGDILNDRGLNELVTFFGQFIDHTVVFTPVSETHMDIPIPEDDDIFANFTGGVLEFRRSERVPLAPGSRRERPVNELSA
ALDLASVYGVDQERNEELRTLVDGKLKTSPGNLLPLNTAGLPNAPSTGPNFFLAGDERANENPTLSALHTLFVLEHNNICDELKTN
FPSYTDEQLYETARKINIAEFQKITYEEFYPALTGRRIRRYRGFRATTDVGVLNEFATAGFRVGHTLVGNAIHRTGPGNSPRPDIP
FGEMFFRSAEVLQDNGIDEFIRGATQFEAQEVDLKVHDALRNGLFQGIPGEDAGFDLIALNLQRSRDHNLISYNQLRKRLGGGRAR
NFAQINRNVNIQNLLSTAYDGDVDKVEAWVGMMAERHERGGSFGPTLRKLWDKQFRMFRDGDRFFYLNDIFSTELRNAIPRIDALR
ADSETFRDIIIRNTDITDSELPRRIFFTQ

>heme-peroxidase5
MMSHIRSTALVAFVLLALSSFVVGMPPPQGRKYASRARIPTAQRQISCDIDTRSLDGQCTNPTDPSWSATDTAQFSYIEGHSSNIP
TGENLPSPRLISNTLCMQSEDLFNDREISEFFTFFGQLVDHDLYLTPTSEDEFPIPIGPEDEIFGMFPGDVLEFTRSERVPIVEGE
IAERPVNTVSGALDLSTVYGSSEERNAALREENSCKLLVSEGDLLPLNTMQIPNSPTTDPNMFVAGDTRSNEHPVLATMHTLFVRE
HNYICDQLAVLMPNTTAEEQYENARTINIAEFQDVVYDEFFPALIGRTLAPYEGFDPSVDPTPSNVYGAAGFRIGHTLVGNSLSRA
GPGNEPLEPLTMEEMFFRSTELLTDLGIEEFMRGSMQTTAQEVDIMIHDALRNFLFSEVEEEEGFDLIALNLQRGRDHAIPKFNEL
RVALNMEPLGSFAELTANEEVQAGMEEVYGTIDDVEAWIGMVAEDHMPDSSVGPTMGELWRTEYTRIRDGDMFFYQNSETFPAELA
ELPLITRLDEPGSVLRDIIIRNSEITEEEMNESPFFTS

SUPPLEMENTARY REFERENCES

1 Chekan, J. R. *et al.* (2019) Scalable biosynthesis of the seaweed neurochemical, kainic acid. *Angew. Chem.* 58, 8454-8457.

2 Hadziavdic, K. *et al.* (2014) Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS One* 9, e87624.

3 Zuccarello, G. C., Burger, G., West, J. A. & King, R. J. (1999) A mitochondrial marker for red algal intraspecific relationships. *Mol. Ecol.* 8, 1443-1447.

4 Tamura, K. & Nei, M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512-526.

5 Wever, R., Krenn, B. E. & Renirie, R. (2018) Marine vanadium-dependent haloperoxidases, their isolation, characterization, and application. *Method Enzymol.* 605, 141-201.

6 Thapa, H. R. *et al.*(2016) A squalene synthase-like enzyme initiates production of tetraterpenoid hydrocarbons in *Botryococcus braunii* Race L. *Nat. Commun.* 7, 11198.

7 Bolger, A. M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120.

8 Bushnell, B., Rood, J. & Singer, E. (2017) BBMerge - Accurate paired shotgun read merging via overlap. *PLoS One* 12, e0185056.

9 Bankevich, A. *et al.* (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455-477.

10 Collen, J. *et al.* (2013) Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proc. Natl. Acad. Sci. USA* 110, 5247-5252.

11 Hughes, A. J. & Keatinge-Clay, A. (2011) Enzymatic extender unit generation for in vitro polyketide synthase teactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol.* 18, 165-176.

12 Maier, T. *et al.* (2007) PYRAZOLOPYRIMIDONES. PCT/EP2007/055846.

13 Quadri, L. E. N. *et al.* (1998) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* 37, 1585-1595.

14 Thapa, H. R., Lail, A. J., Garg, N. & Agarwal, V. (2018) Chemoenzymatic synthesis of starting materials and characterization of halogenases requiring acyl carrier protein-tethered substrates. *Method Enzymol.* 604, 333-366.

15 Miller, I. J. *et al.* (2019) Autometa: automated extraction of microbial genomes from individual shotgun metagenomes. *Nucleic Acids Res.* 47, e57.

16 Stanke, M. & Morgenstern, B. (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 33 W465-7.

17 Livak, K. J. & Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402-408.

18 Gietz, R. D. & Schiestl, R. H. (1991) Applications of high-efficiency lithium-acetate transformation of Intact yeast-cells using single-stranded nucleic-acids as carrier. *Yeast* **7**, 253-263.

19 Montllor-Albalate, C. *et al.* (2019) Extra-mitochondrial Cu/Zn superoxide dismutase (Sod1) is dispensable for protection against oxidative stress but mediates peroxide signaling in *Saccharomyces cerevisiae*. *Redox Biol.* 21.

20 Kumar, S., Stecher, G. & Tamura, K. (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870-1874.

21 Sievers, F. & Higgins, D. G. (2018) Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* 27, 135-145.

22 Agarwal, V. *et al.* (2015) Chemoenzymatic synthesis of acyl Coenzyme A substrates enables in situ labeling of small molecules and proteins. *Org Lett* 17, 4452-4455.

23 Reynolds, S. M., Kall, L., Riffle, M. E., Bilmes, J. A. & Noble, W. S. (2008) Transmembrane topology and signal peptide prediction using dynamic bayesian networks. *Plos Comput. Biol.* 4, e1000213.

24 Panday, A., Sahoo, M. K., Osorio, D. & Batra, S. (2015) NADPH oxidases: an overview from structure to innate immunity-associated pathologies. *Cell Mol. Immunol.* 12.

25 Magnani, F. *et al.* (2017) Crystal structures and atomic model of NADPH oxidase. *Proc. Natl. Acad. Sci. USA* 114, 6764-6769.

26 Lee, J. *et al.* (2013) Structural and functional insight into an unexpectedly selective N-methyltransferase involved in plantazolicin biosynthesis. *Proc. Natl. Acad. Sci. USA* **110**, 12954-12959.

27 Agarwal, V. *et al.* (2014) Biosynthesis of polybrominated aromatic organic compounds by marine bacteria. *Nat. Chem. Biol.* 10.