Supporting Information

Probing the Mechanism of Cyanobacterial Aldehyde Decarbonylase using a Cyclopropyl Aldehyde

Bishwajit Paul,[†] Debasis Das,[†] Benjamin Ellington,[‖] and E. Neil G. Marsh[*,†,‖]

[†]Department of Chemistry and [‖]Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

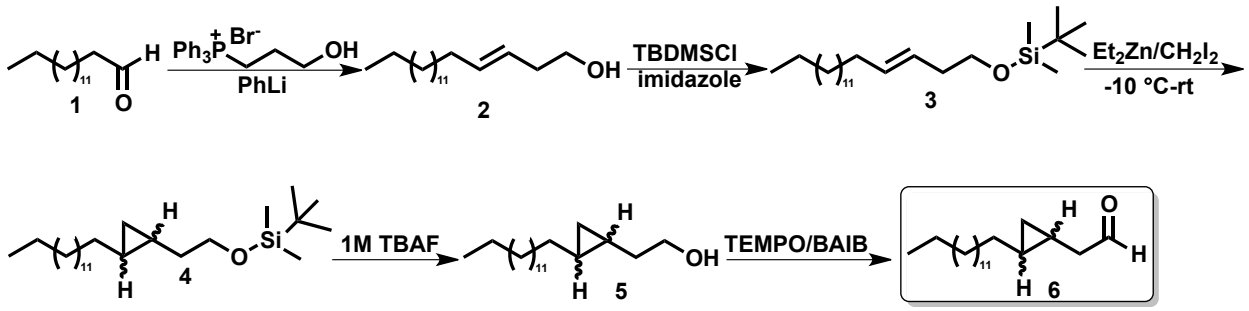
*Corresponding author: e-mail: nmarsh@umich.edu

Table of Contents

Synthesis and Characterization 2-(2-tetradecylcyclopropyl)acetaldehyde (6):

Synthesis of 2-(2-tetradecylcyclopropyl)acetaldehyde **(6)** was achieved in five steps with an overall yield of 10% as outlined in Scheme S1:

Scheme S1. Synthesis of 2-(2-tetradecylcyclopropyl)acetaldehyde (**6**):



Materials. Octadecanal, pentadecanal, tetradecane, heptadecane, 1-octadecene, phenazine methosulfate (PMS), ferrous ammonium sulfate and NADH were obtained from Acros Organics. Potassium chloride, HEPES were from Fisher chemicals. D_2O (99.9%) and DMSO-d_6 (99.9%) were from Cambridge Isotope Laboratories, Inc. All other reagents were of the purest grade commercially available and used without further purification.

Synthesis of Pentadecanal (1). Oxidation of 1-pentadecanol to pentadecanal **(1)** was performed using TEMPO/BAIB strategy according to a literature procedure.[1] To a solution of 1-pentadecanol (1 g, 4.4 mmol) were added (2,2,6,6-Tetramethylpiperidin-1-yl)oxyl (TEMPO, 68.6 mg, 0.44 mmol) and bis(acetoxy)iodobenzene (BAIB, 1.7 g, 5.3 mmol) in dichloromethane at room temperature. The reaction was stirred at room temperature for 3 hours. The solvent was removed under reduced pressure using a rotatory evaporator and the crude mixture was applied to a silica-gel column equilibrated in n-hexane. The column was developed by slowly increasing the polarity of the solvent using a gradient of 0 to 1 % ethylacetate in hexane to obtain **1** (850 mg, 85%). The compound was pure as judged by NMR (Figure S1). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.84 (t, *J* = 6.8 Hz, 1H), 2.47 (t, *J* = 6.8 Hz, 2H), 1.40 (t, *J* = 7.2 Hz, 22H), 1.38-1.32 (m, 22H), 0.94 (t, *J* = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$), δ 202.92, 43.89, 31.89, 29.65, 29.63, 29.62, 29.61, 29.55, 29.40, 29.35, 29.33, 29.14, 22.66, 22.06, 14.09.

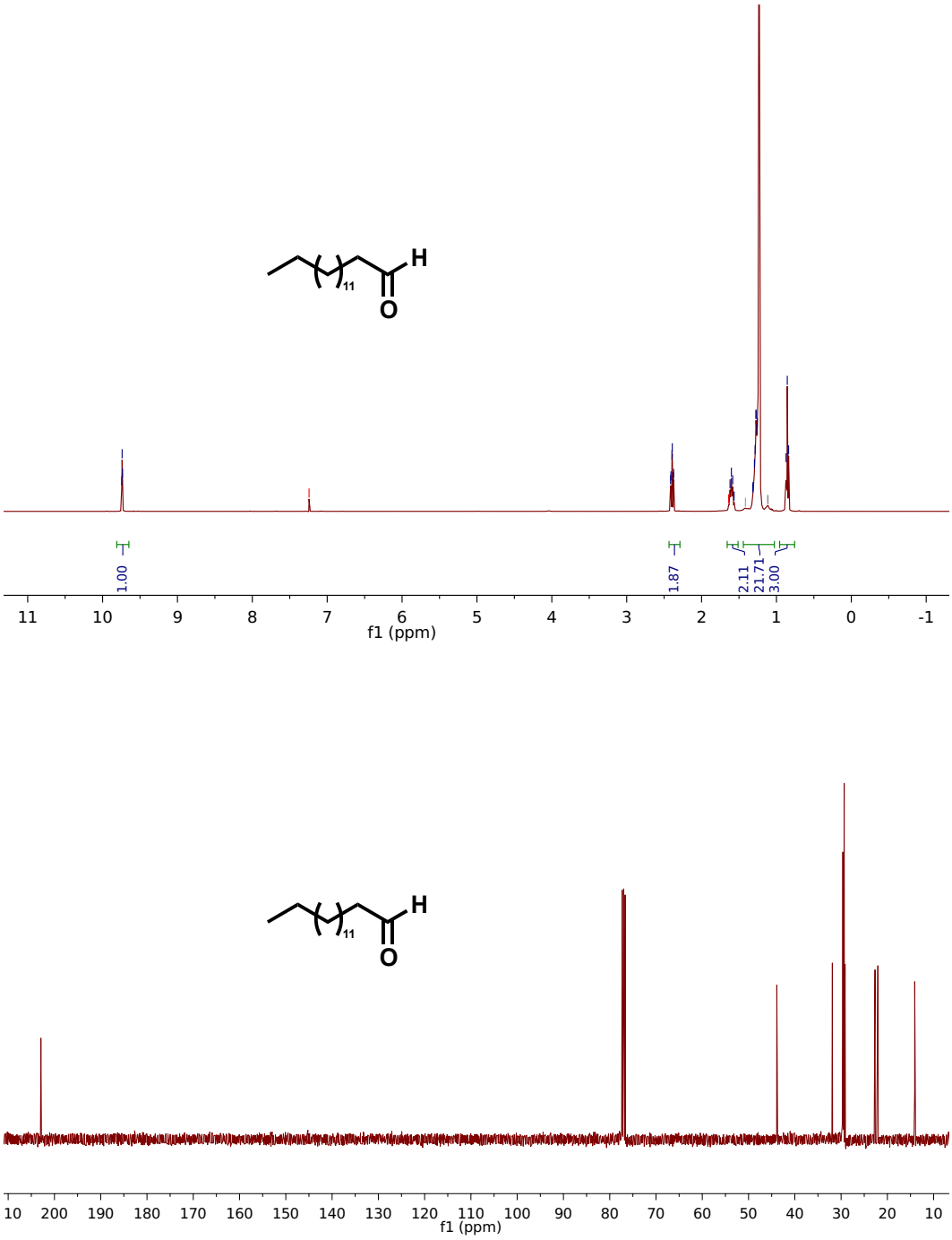


Figure S1. ^{1}H and ^{13}C NMR spectrum of compound **1** in $CDCl_3$.

Synthesis of *trans*-octadec-3-en-1-ol (2). The Wittig reaction of **1** with 3-hydroxypropyl-triphenylphosphosonium bromide was performed according to a literature procedure.[2] To a suspension of 3-hydroxypropyl-triphenylphosphosonium bromide (2 g, 4.98 mmol) in tetrahydrofuran (THF), 1.5 equivalents (3.7 ml) of phenyllithium (1.2 M dissolved THF) was added dropwise under nitrogen at room temperature. The reaction was stirred at room temperature for 30 min. The reaction mixture was cooled to -78 °C in a dry ice/acetone mixture bath for 15 min, followed by dropwise addition of pentadecanal (563 mg, 2.5 mmol) (**1**) (dissolved in minimum volume of THF). After the addition was complete, the reaction mixture was warmed to -30 °C, followed by addition of 0.5 equivalent of phenyl lithium (1.2 ml) to obtain predominantly the *trans*-product**.** The reaction was allowed to warm to room temperature with stirring over 30 min. Finally, the reaction was cooled to -78 °C and quenched by addition of 1.8 ml 6M hydrochloric acid and 5 ml of water. After warming to room temperature the aqueous phase was extracted with diethyl ether (3 x 20 ml), washed with brine, dried over sodium sulphate and concentrated under reduced pressure by rotatory evaporator. The crude mixture was subjected to silica gel column chromatography in n-hexane/EtoAc mixture in which the gradient of the solvent was gradually increased from neat n-hexane to 4 % EtoAc/n-hexane to obtain **2** (460 mg, 68 %) as a predominant *trans* isomer. The compound was pure as judged by NMR (Figure S2) and TLC. Its identity was confirmed by high resolution electron-impact MS (*m/z*): calculated 268.2766; observed 268.2761. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.54-5.48 (m, 1H), 5.42-5.37 (m, 1H), 3.60 (t, J = 6.3 Hz, 2H), 2.24 (q, J = 6.5 Hz, 2H), 1.99 (q, J = 7.0 Hz, 2H), 1.32-1.26 (m, 23H), 0.86 (t, J = 6.8 Hz, 3H);

^{13}C NMR (100 MHz, $CDCl_3$) δ 134.42, 125.49, 61.99, 35.96, 32.66, 31.90, 29.68, 29.67, 29.63, 29.61, 29.55, 29.60, 29.49, 29.44, 29.34, 29.18, 22.67, 14.10.

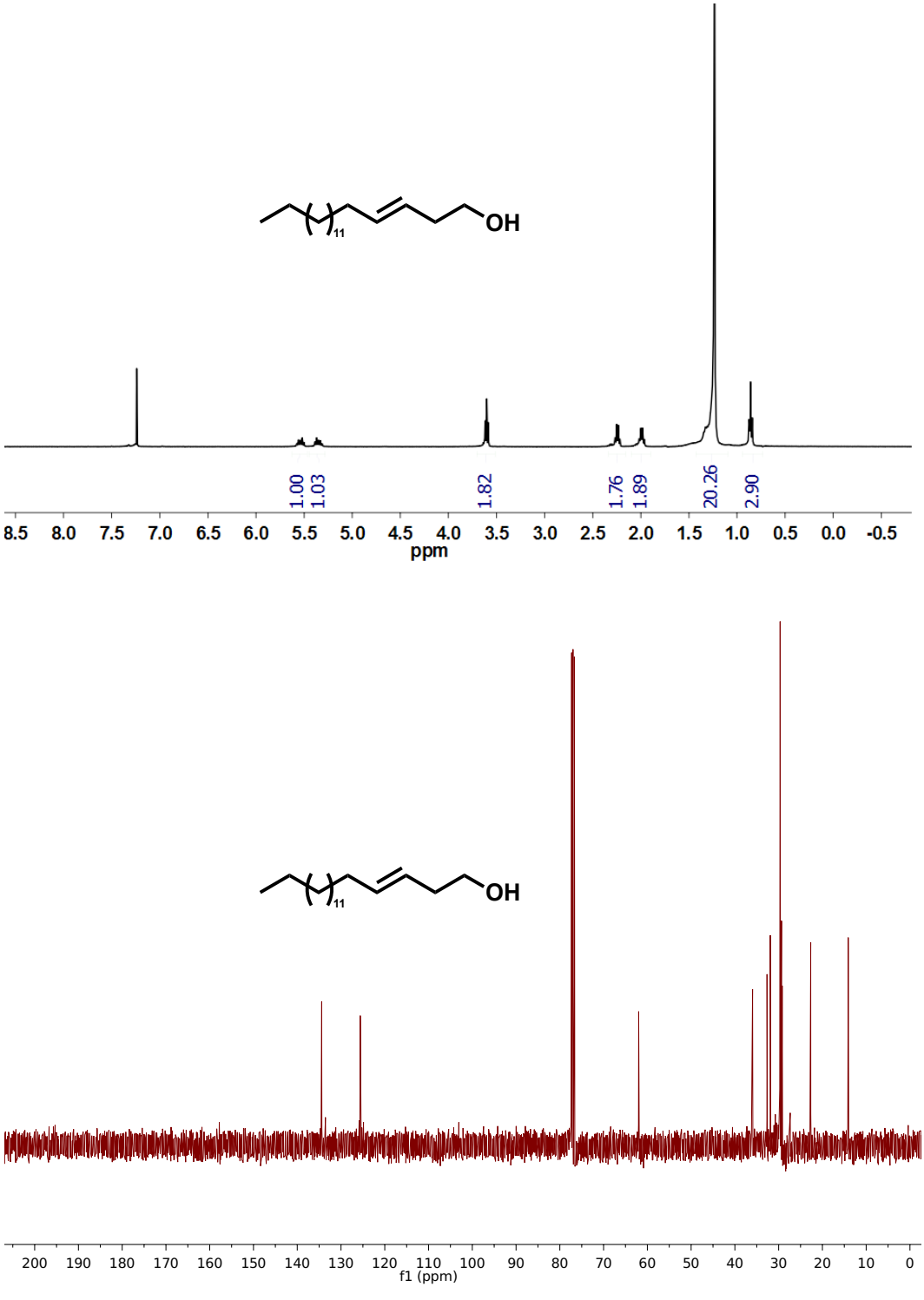


Figure S2. ^{1}H and ^{13}C-NMR spectrum of compound **2** in $CDCl_3$.

Synthesis of *trans-tert*-butyldimethyl(octadec-3-en-1-yloxy)silane (3). The alcohol functional group in **2** was protected by *tert*-Butyl-dimethylsilyl group (TBDMS) using a literature procedure.[3] To a solution of **2** (100 mg, 0.37 mmol) in anhydrous dicholoromethane (DCM) at 0 °C were added imidazole (40 mg, 0.59 mmol), and catalytic amount of DMAP. TBDMS chloride (66 mg, 0.44 mmol) was added to the reaction mixture resulting in the formation of white suspension. The reaction mixture was gradually warmed to room temperature and stirred overnight. The reaction mixture was diluted with DCM (50 ml) and washed with water (2 X 5 ml). The organic layer was separated, dried over anhydrous sodium sulphate and concentrated on a rotatory evaporator. The crude mixture was purified by silica-gel chromatography using 1 % ethylacetate/n-hexane as the eluting solvent to yield **3** (120 mg, 85%). The compound was pure as judged by NMR (Figure S3) and TLC. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.52-5.46 (m, 1H), 5.42-5.35 (m, 1H), 3.62 (t, J = 7.0 Hz, 2H), 2.22 (q, J = 6.8 Hz, 2H), 2.11 (q, J = 6.8 Hz, 2H), 1.32-1.26 (m, 22H), 0.95-0.85 (m, 12H), 0.07 (s, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 132.65, 131.91, 126.27, 125.43, 63.39, 63.06, 33.86, 32.72, 31.96, 31.15, 29.74, 29.73, 29.72, 29.70, 29.39, 25.97, 25.95, 22.72, 18.99, 18.37, 14.13, -5.25, -5.26.

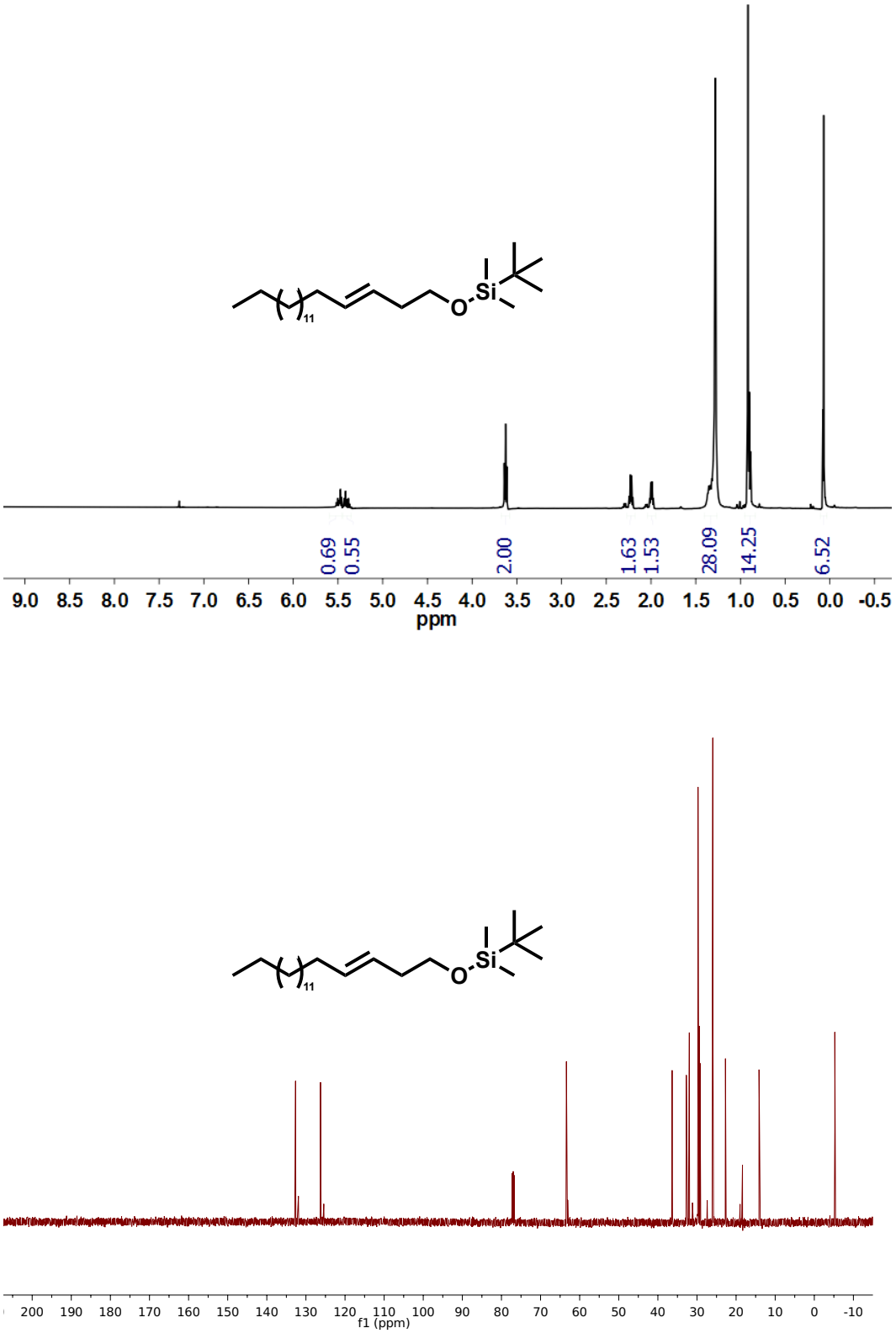


Figure S3. ^{1}H and ^{13}C-NMR spectrum of compound **3** in $CDCl_3$.

Synthesis of *tert*-butyldimethyl(2-(2-tetradecylcyclopropyl)ethoxy)silane (4). 3 was converted to **4** based on literature procedures.[4,5] To a solution of **3** (120 mg, 0.31 mmol) in anhydrous dichloromethane at -10 °C were added diethyl zinc (70 μl, 0.7 mmol) and diiodomethane (100 μl, 1.2 mmol). The reaction mixture was warmed to room temperature and stirred for 48 hours. Upon completion of the reaction as judged by TLC, the crude mixture was concentrated on a rotatory evaporator. **4** was purified by silica-gel column chromatography using 1% ethylacetate/n-hexane as the eluting solvent and was obtained as predominant the *trans*-stereoisomer (40 mg, 30 %). The compound was > 90% pure as judged by NMR (Figure S4) and TLC. ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.67-3.58 (m, 2H), 2.15-2.11 (m, 2H), 1.36-1.25 (m, 22H), 0.90-0.85 (m, 12H), 0.43-0.37 (m, 1H), 0.19-0.13 (m, 2H), 0.06-0.03 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 63.35, 37.69, 34.25, 31.95, 30.32, 29.73, 29.69, 29.65, 29.60, 29.55, 29.50 29.40, 29.22, 26.03, 25.99, 25.97, 22.72, 18.51, 18.38, 15.29, 14.15, 11.45, -5.26.

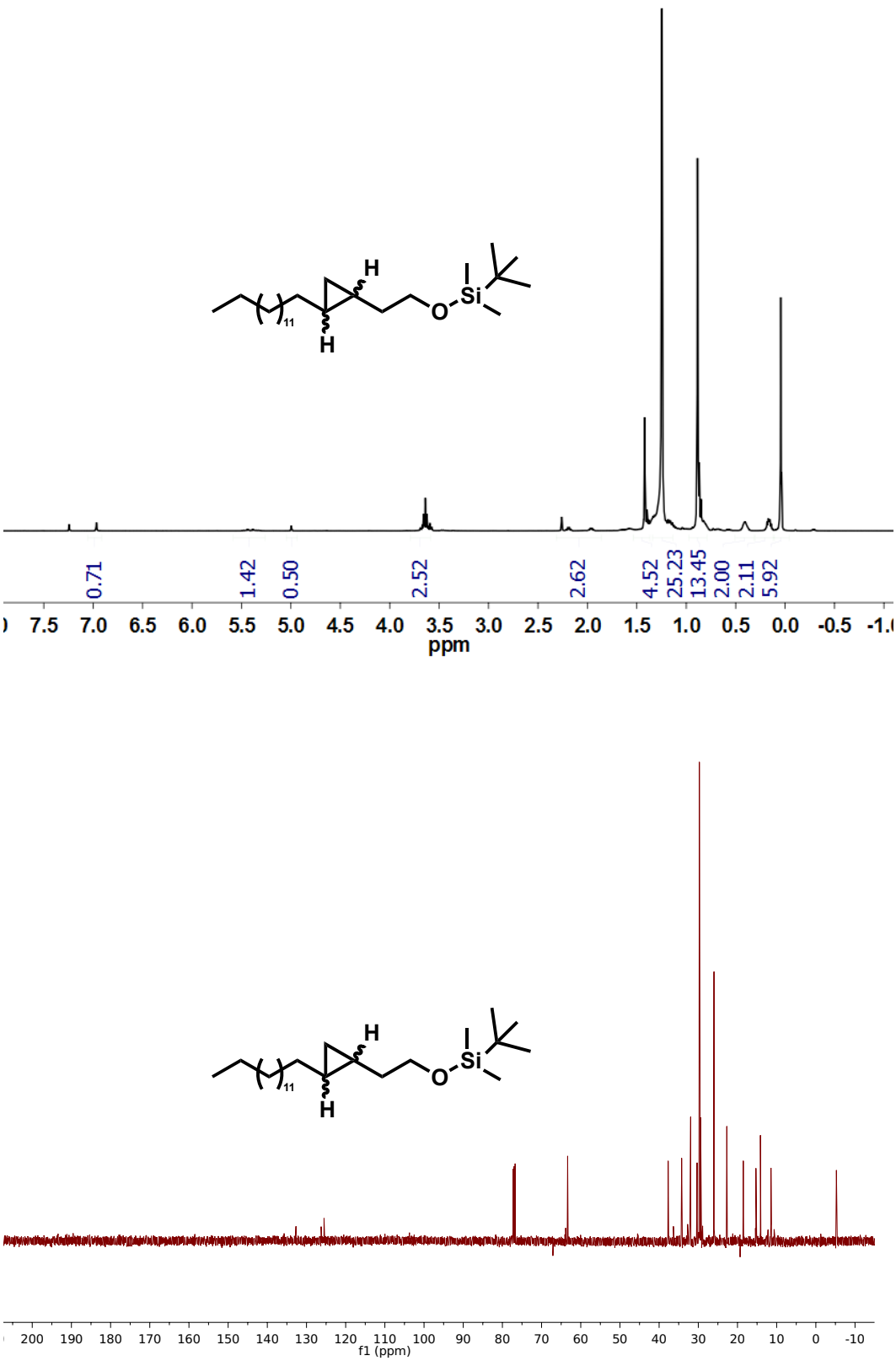


Figure S4. ^{1}H and ^{13}C NMR spectrum of compound **4** in $CDCl_3$.

Synthesis of 2-(2-tetradecylcyclopropyl)ethanol (5). 1 ml of 1.0 M TBAF in THF was added to a solution of **4** (120 mg, 0.304 mmol) in anhydrous THF at room temperature. The reaction mixture was stirred for 3 h under nitrogen. The reaction mixture was diluted with 50 ml dichloromethane and washed with water (2 x 5 ml) and brine (2 x 2 ml). The organic layer was dried over sodium sulphate and concentrated by rotatory evaporation to afford the crude product as yellow oil that was further subjected to silica-gel chromatography using 1% ethylacetate/n-hexane as the eluting solvent and yielded **5** as a white solid (70 mg, 85%). The compound was > 85% pure as judged by NMR (Figure S5) and TLC. ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.73-3.68 (m, 2H), 2.15-2.11 (m, 2H), 1.49-1.47 (m, 2H), 1.45-1.36 (m, 22H), 0.90 (t, J = 6.8 Hz, 3H), 0.47-0.42 (m, 2H), 0.24-0.21 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 63.52, 63.18, 62.00, 37.22, 35.98, 34.18, 32.69, 31.93, 31.74, 30.09, 29.71, 29.67, 29.64, 29.54, 29.52, 29.47, 29.41, 29.38, 29.22, 28.86, 26.11, 25.65, 22.70, 18.38, 18.38, 15.16, 15.11, 14.13, 12.12, 11.33, 10.57.

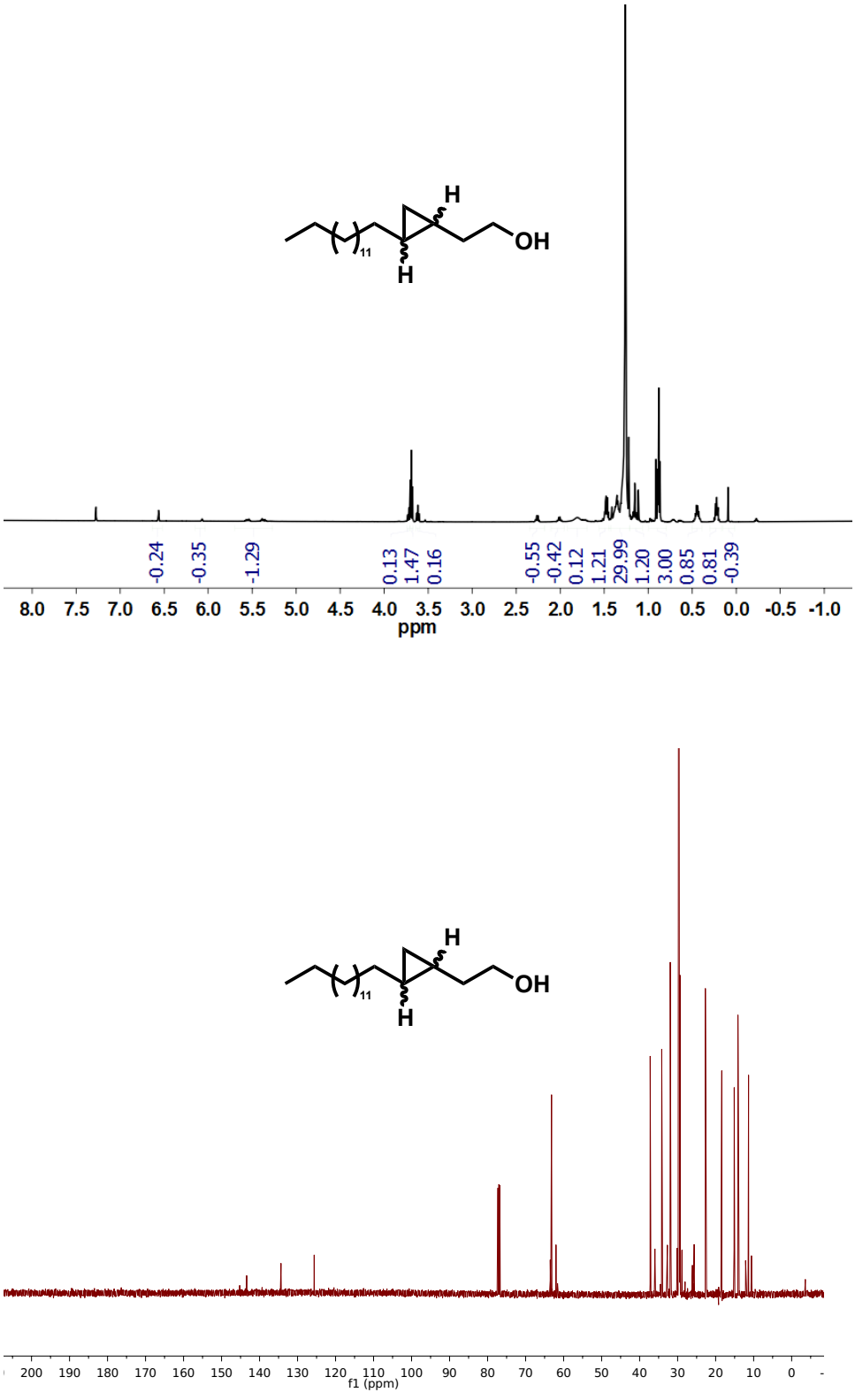


Figure S5. ^{1}H and ^{13}C NMR spectrum of compound **5** in $CDCl_3$.

Synthesis of 2-(2-tetradecylcyclopropyl)acetaldehyde (6). To a solution of **5** (100 mg, 0.34 mmol) in dichloromethane at room temperature were added TEMPO (5 mg, 0.034 mol) and BAIB (120 g, 0.37 mmol). The reaction was stirred at room temperature for 3 hours under nitrogen. After completion, the crude reaction mixture was concentrated using a rotatory evaporator and subjected to silica-gel column chromatography in n-hexane/ethylacetate in which the gradient of the solvent was gradually increased from neat n-hexane to 0.5 % ethylacetate/n-hexane to obtain **6** (60 mg, 65%). The compound was pure as judged by NMR (Figure S6) and TLC. The identity of the compound was confirmed by high resolution electron-impact MS (*m/z*): calculated 280.2760; observed 280.2766. GC-MS analysis of **6** indicated that the compound was a mixture of ~ 70:30 *trans*- to *cis*-stereoisomers. ^{1}H NMR (400 MHz, $CDCl_3$) of **6** (major stereoisomer) δ 9.76 (t, *J* = 7 Hz, 1H), 2.30-2.26 (m, 2H), 1.39-1.24 (m, 22H), 0.86 (t, *J* = 6.8 Hz, 3H), 0.70-0.66 (m, 1H), 0.55-0.50 (m, 1H), 0.41-0.30 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 202.52, 48.20, 47.30, 43.22, 33.87, 32.69, 31.89, 29.80, 29.67, 29.65, 29.63, 29.53, 29.45, 29.39, 29.33, 29.13, 28.95, 22.66, 18.54, 15.04, 14.09, 11.70, 11.51, 10.74, 8.95.

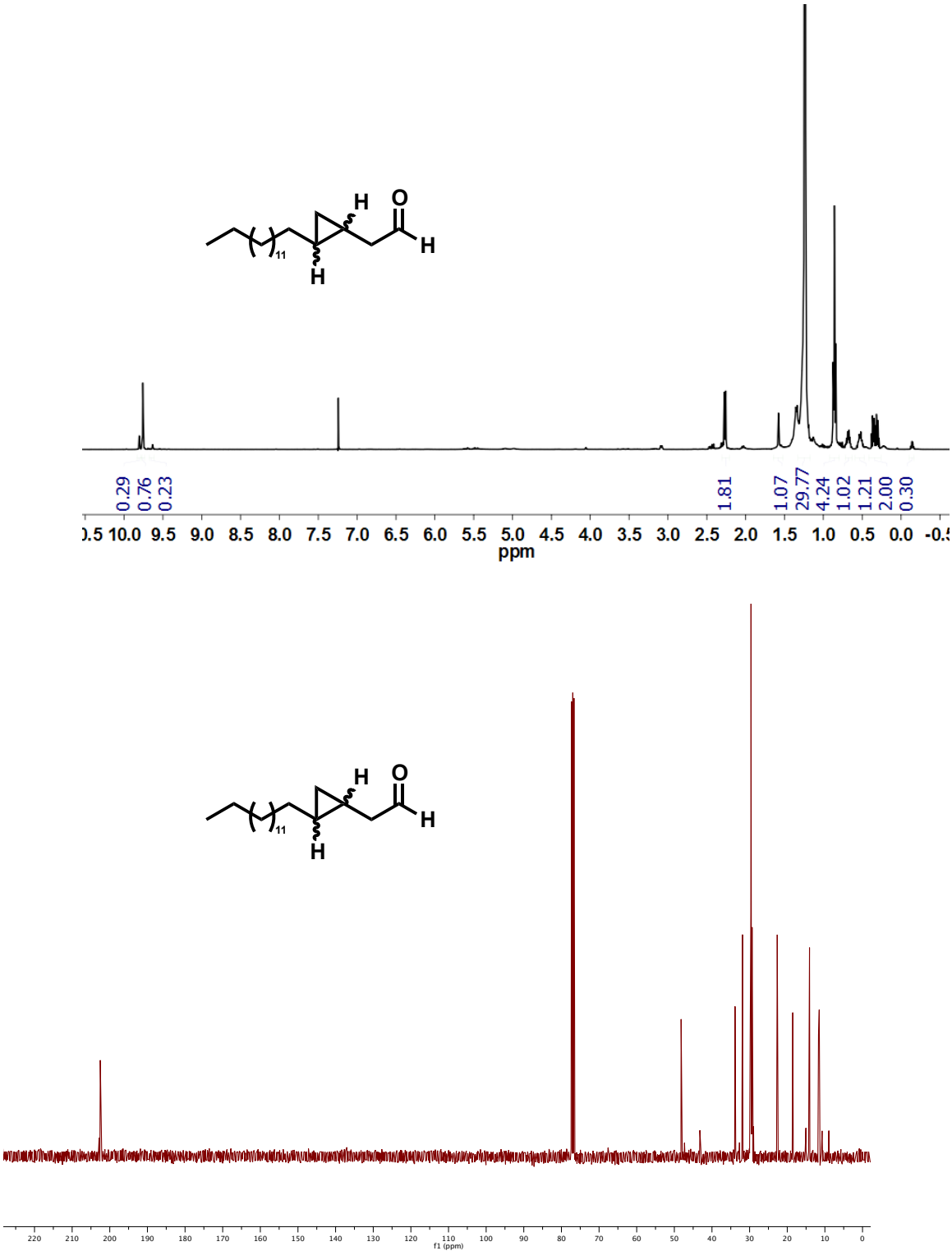


Figure S6. ^{1}H and ^{13}C-NMR of cyclopropyl compound **6**.

Enzyme assay.

Assays were performed in 100 mM HEPES buffer, pH 7.2, containing 100 mM KCl and 10% glycerol as described previously.[6,7] Aldehydes substrates were made up to a stock solution in DMSO. A typical assay contained 10 μM cAD, 20 μM ferrous ammonium sulfate, 400 μM aldehyde substrate, 100 μM phenazine methosulfate (PMS) and 1 mM NADH in a total volume of 500 μL. Assays were shaken for 1 hr at 37 ºC at 200 rpm. Reactions were quenched by addition of 500 μL ethyl acetate and vortexed well to extract hydrocarbon products and un-reacted substrate. The ethyl acetate layer was separated and 8 μL sample of the ethyl acetate layer was injected into GC-MS for analysis. Enzymatic production 1-octadecene was quantified using a calibration plot of 1-octadecene.

GC-MS analysis. Gas chromatography analysis was performed using a Shimadzu QP-2010S GC-MS instrument equipped a quadrupole mass detector. A DB-5 column (Restek, 30m x 0.25 mm x 0.25 μm) was used for elution. The flow rate of the helium carrier gas was kept constant at 1 mL/min and the inlet temperature was maintained at 200 °C. The interface temperature was maintained at 250 °C. Injections were made in splitless mode. Oven temperature was held initially at 70 °C for 2 min and then gradually increased to 300 °C at 20 °C/min and finally maintained at 300 °C for 5 min. Data analysis was performed by GC-MS PostRun analysis software.

Formation of 1-Octadecene from 6 by cAD.

The formation of 1-Octadecene from the cyclopropyl substrate **6** by cAD was confirmed by GC-MS. 1-octadecene was characterized by a peak at 9.84 min (Figure S7) with a mass of *m/z* = 252.3 (Figure S8). Enzymatically formed 1-octadecene was co-chromatographed with an authentic standard of 1-octadecene. The amount of 1-

octadecene formation was quantified by reference to a standard curve. The presence of PMS, NADH and O_2 was essential for the activity of cAD (Figure S10).

Non-enzymatic formation of 1-ethyl-2-tetradecylcyclopropane from 6.

Interestingly, a minor peak at 10.0 min with *m/z* = 252.3 was also evident that co-chromatographed with an authentic standard of 1-ethyl-2-tetradecylcyclopropane, the non-rearranged product from the decarbonylation of **6** (Figure S7). The non-rearranged product formation was independent of the presence of cAD and formed in similar amount when cAD was omitted from the assay (Figure S9).

The small amount of cyclopropyl product formed has prevented us from determining whether the C1 carbon is converted to formate, as observed in the cAD-catalyzed reaction, or possibly to CO or CO_2, as observed in the insect and plant decarbonylases. In the absence of this information, a mechanism for this unusual side reaction cannot be postulated. However, the fact that the reaction is independent of Fe^{2+} suggests that it occurs by a mechanism of rather different from that of cAD. O_2 is known to react with reduced PMS to generate a variety of reactive oxygen species[8] that could conceivably attack an aldehyde. The observation that the non-enzymatic reaction it does not result in ring opening implies that the non-enzymatic reaction either does not involve a radical intermediate or that the radical is extremely short lived. Given that the stabilization of reactive species is a ubiquitous feature of enzyme catalysis, it is perhaps not surprising that the enzyme-generated cyclopropylcarbinyl radical is much longer lived. The non-enzymatic cleavage of the C1–C2 bond of **6** may be made easier due to stabilization of the intermediate radical (or transition state) through hyper-conjugation

with the cyclopropyl ring. We note that, under the same conditions, non-enzymatic decarbonylation was not observed with simple aldehydes that are substrates for cAD.

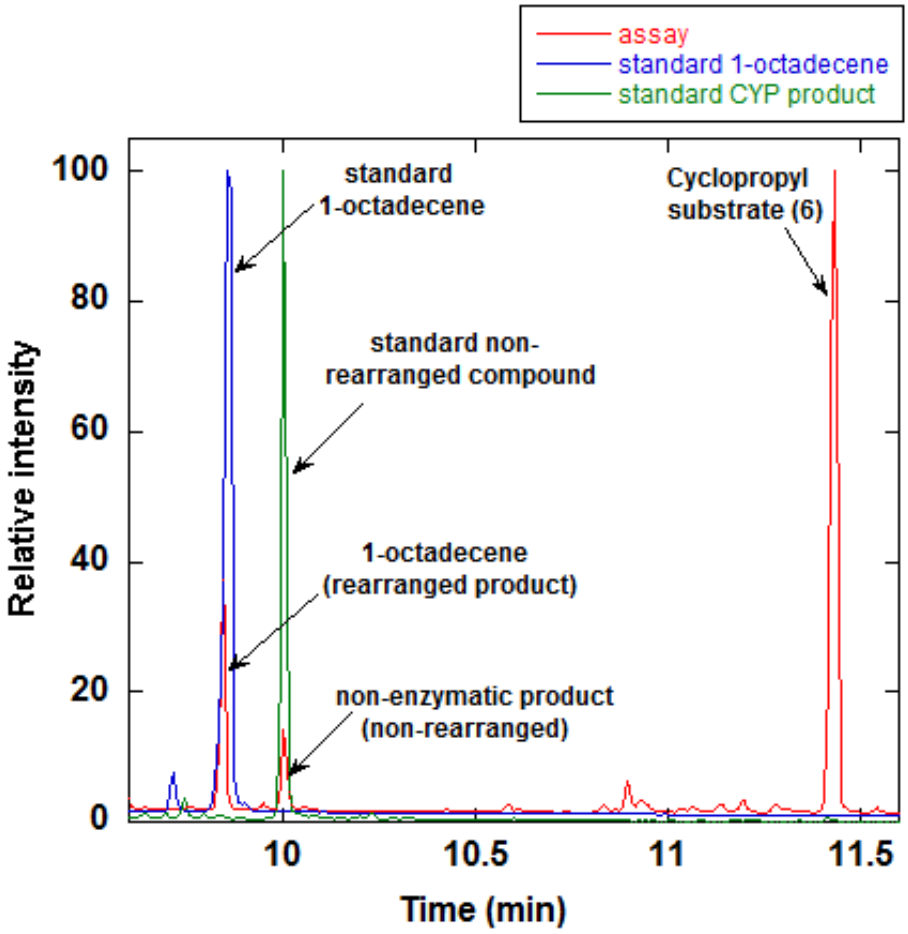
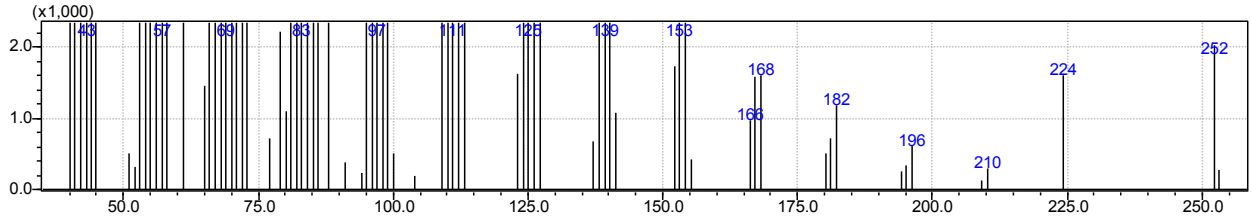


Figure S7. Gas chromatographs of 1-octadecene formed from **6** by *Np* cAD (in red), an authentic standard 1-octadecene (in blue) and an authentic standard of 1-methyl-2-tetradecylcyclopropane, formed non-enzymatically (in green). Retention times of 1-octadecene and 1-methyl-2-tetradecylcyclopropane products are at 9.84 min and 10 min, respectively.

A



B

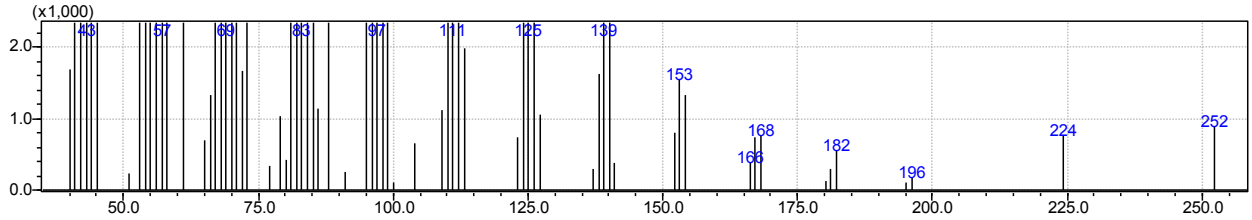


Figure S8. Electron Impact mass spectral analysis of authentic standard 1-octadecene (A) and enzymatically obtained 1-octadecene (B). Molecular ion peak of *m/z* = 252 is characteristic of 1-octadecene.

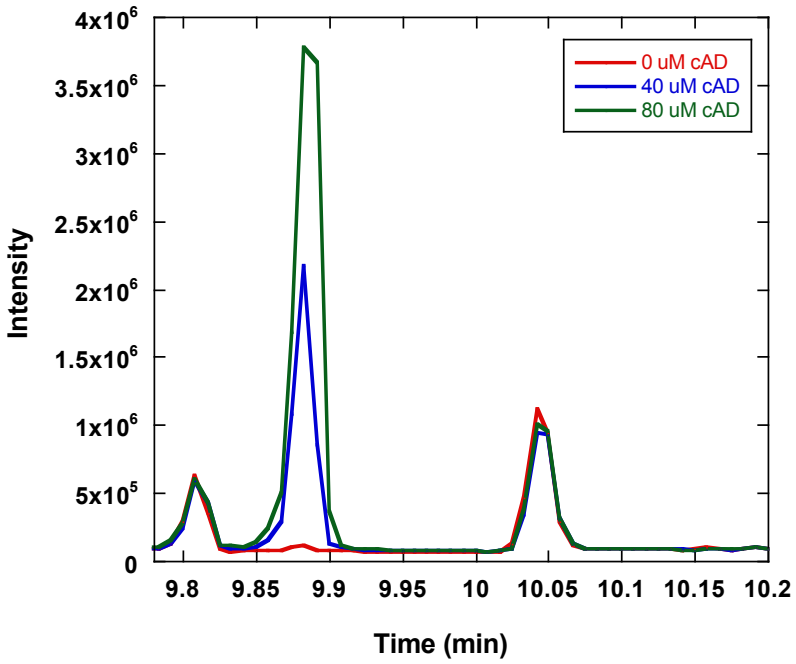


Figure S9. Formation of 1-methyl-2-tetradecylcyclopropane (retention time = 10.05 min) from **6** is independent of cAD. The formation of octadecene increases linearly with enzyme concentration (blue and green traces); in the absence of cAD no octadecene is formed (red trace), whereas the amount of 1-methyl-2-tetradecylcyclopropane formed is constant.

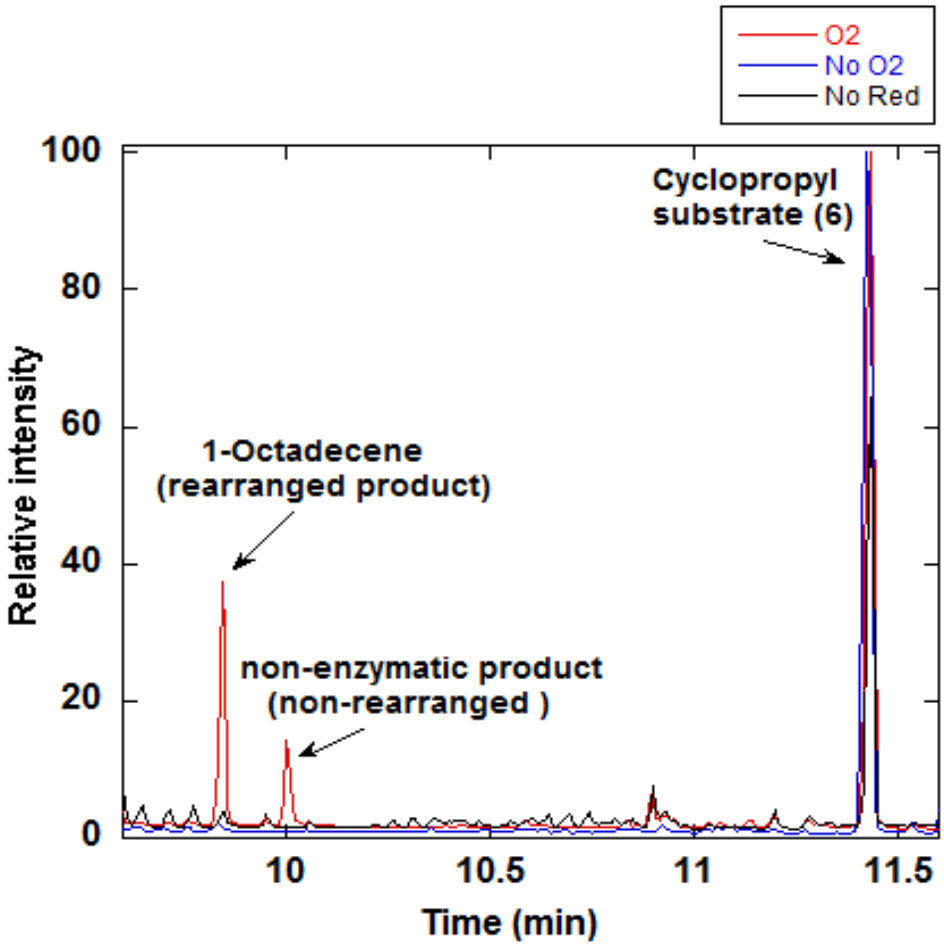


Figure S10. Overlaid chromatographs showing the formation of 1-octadecene and the non-enzymatically formed cyclopropyl alkane product in the presence of cAD, O_2 and PMS/ NADH (red trace). In the absence of O_2 (blue trace) or in the absence of reducing system PMS/ NADH (black trace), neither 1-octadecene nor the cyclopropyl alkane product is observed.

Identification of formate as co-product in reaction of 6 with cAD.

The detection of formate as a co-product in cAD assay was confirmed by derivatizing the products of the enzyme reaction with 2-nitrophenylhydrazine (2-NPH) followed by reverse phase HPLC as described in the literature.[6,9] 400 μl of the products of the enzyme reaction was mixed with 40 μl of 120 mM 2-NPH solution (aqueous solution in 0.25 M HCl) and 40 μl of EDC working solution (300 mM EDC in 1:1 pyridine:HCl). After vortexing for ~30 s, the reaction mixtures were incubated at 60 °C for 30 min. The samples were then centrifuged to remove precipitated protein and insoluble reaction products. 400 μl of the clear supernatant was subjected to chromatography on a Nucleosil C18 RP HPLC column (250 mm x 4 mm, 5 μM, 120 Å). The column was equilibrated in 50 % water, (acidified with 0.05% AcOH) and 50 % methanol (acidified with 0.05% AcOH) and compounds were eluted with a gradient of 30% methanol to 90% methanol over 45 min at 0.7 mL/min. The hydrazide derivative of formate was eluted at ~28 min and detected at 230 nm (Figure S11). The presence of formate in the derivatized assay samples was confirmed by comparison with the retention time of an authentic standard. Analysis of this peak by UV-visible and mass spectrometry further confirmed the identity of the compound as the 2-NPH derivative of formate.

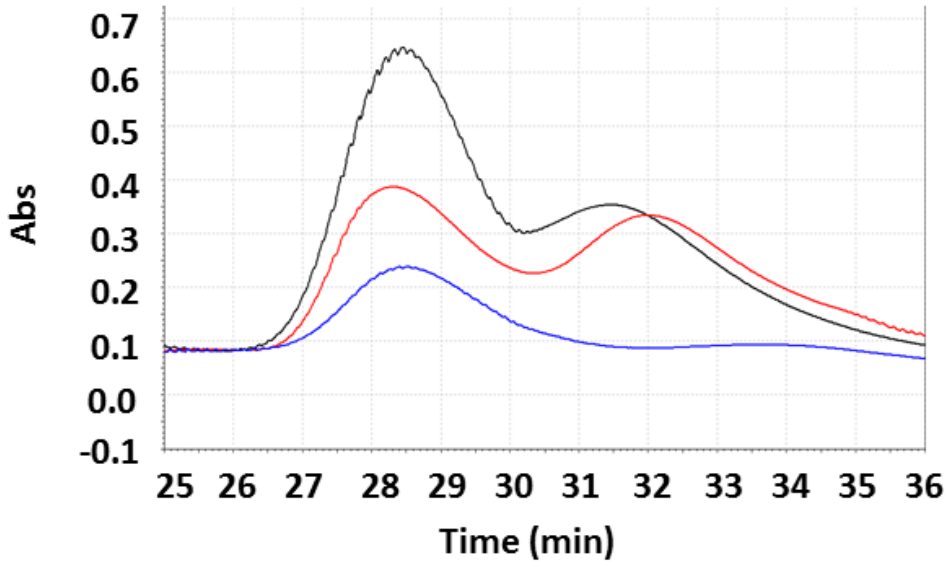


Figure S11: Overlaid HPLC traces of 2-NPH derivatives of authentic formate (in blue), *Np* cAD reaction product with cyclopropyl compound **6** (red) and *Np* cAD reaction product with octadecanal (black) at 230 nm. 2NPH-formate derivative elutes at retention time of ~ 28 min. Fractions were collected and studied by ESI-MS (negative mode). Identity of each formate derivative was confirmed by obtained mass of *m/z* = 180.1.

Inhibition of cAD by cyclopropyl aldehyde, 6

To determine whether the enzyme was irreversibly inactivated by **6**, cAD was incubated with **6,** 400 μM, in presence of Fe^{2+}, PMS, NADH and O_2 as described above. The reaction mixture was shaken for 1 hr at 37 ºC and then the authentic substrate, octadecanal, 400 μM, was added. The assay mixture was shaken for an additional 1 h, hydrocarbons extracted with ethyl acetate and product formation was quantified by GC-MS. ~1-equivalent (w.r.t cAD) of 1-octadecene was formed and whereas less than 0.1 equivalents of heptadecane were formed. As a control experiment, cAD was similarly incubated with the alternative substrate pentadecanal, 400 μM, for 1 hr and then assayed with octadecanal for an

additional 1 hr. ~6.5 equivalents of tetradecane and ~4 equivalents of heptadecane were formed, demonstrating that inactivation of cAD was not due to non-specific loss of activity during turnover (see main text).

Identification of covalent modification site in cAD inactivated by 6:

Covalent modification of cAD by cyclopropyl aldehyde **6** was investigated using various mass spectroscopic techniques including LC-MS, MALDI-TOF and LC-MS-MS. After incubation with **6**, *Np* cAD was found to be covalently modified by a mass of 251±0.5 Da. that corresponds to decarbonylated **6.** As discussed below, F107 (shown in the sequence of cAD from *Nostoc punctiformis* in red, below) was determined to be the most likely site of covalent modifications.

Amino acid sequence of His-tagged *Np* cAD:

MHHHHHHSSGVDLGTENLYFQSNAQQLTDQSKELDFKSETYKDAYSRINAIVIE
GEQEAHENYITLAQLLPESHDELIRLSKMESRHKKGFEACGRNLAVTPDLQFAKE
FFSGLHQNFQTAAAEGKVVTCLLIQSLIIEC**F**AIAAYNIYIPVADDFARKITEGVVK
EEYSHLNFGEVWLKEHFAESKAELELANRQNLPIVWKMLNQVEGDAHTMAME
KDALVEDFMIQYGEALSNIGFSTRDIMRLSAYGLIGA

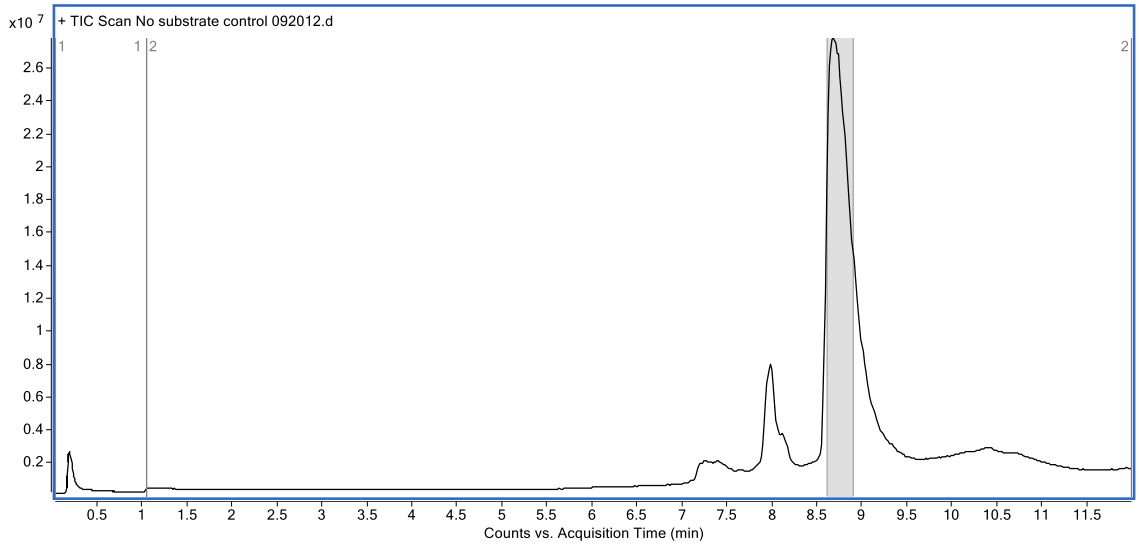
LC-MS analysis. Covalent modifications on cAD were analyzed by LC-MS using an Agilent 6520 LC - accurate-mass Q-TOF MS system. After incubation with either cyclopropyl substrate **6** or with octadecanal, the protein was recovered, desalted and reconstituted with 0.1% formic acid. 5 μl of the sample was injected into a Poroshell 300SB-C8 column equilibrated with 0.1% formic acid and 5% acetonitrile. Proteins were eluted for 5 min with 95% water: 5% acetonitrile followed by an increase in gradient to 100% acetonitrile over 7 min at a constant flow rate of 0.5 ml/min. Eluting proteins were

detected at 280 nm; cAD eluted with a retention time of 8.6-9.5 min. Mass data were obtained using intact protein mode and analyzed using Agilent MassHunter Qualitative Analysis software. The raw data was deconvoluted with respect to maximum entropy.

The mass of the cAD prior to reaction with **6** was determined as 28911 ± 0.5 Da (Figure S12), in excellent agreement with the calculated molecular weight. The mass of cAD incubated with octadecanal was also 28911 ± 0.5 Da (Figure S13). However, incubation of cAD with **6**, resulted in between 60 - 80 % of the recovered enzyme eluting from the column as a species characterized by a slightly longer retention time and a molecular weight of 29162 ± 0.5 Da (Figure S14). The increase in molecular weight of 251 ± 0.5 Da is consistent with the formation of a covalent adduct between decarbonylated **6** and cAD.

LC-MS of as isolated *Np* cAD:

A



B.

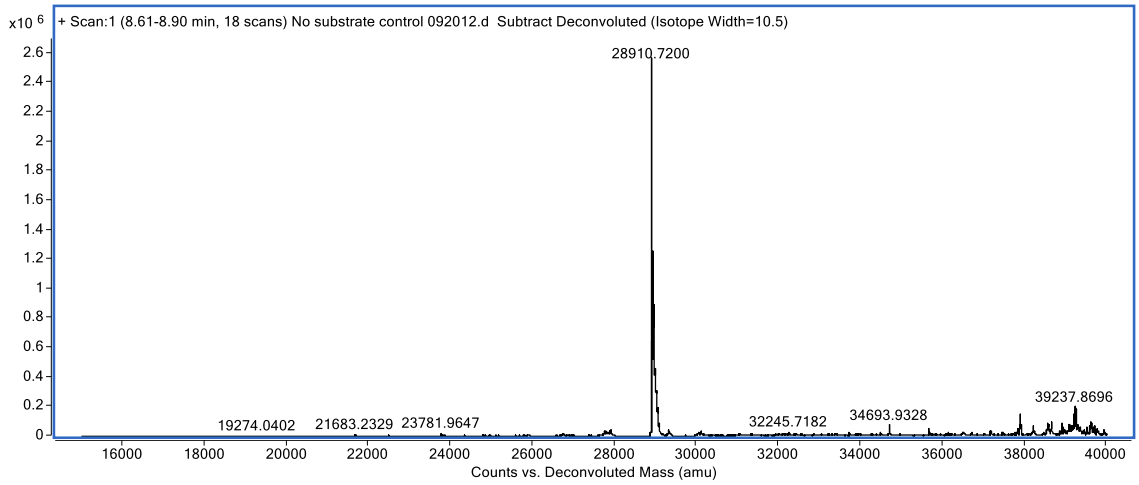


Figure S12. Reverse phase liquid chromatogram and mass spectral analysis of *Np* cAD. A. Total ion chromatogram (TIC) B. Deconvoluted mass spectrum of as isolated *Np* cAD. Highlighted region of A was extracted for mass spectral analysis. *Np* cAD elutes as a single peak with retention time 8.6 to 9.2 min. Mass of *Np* cAD is 28911 ± 0.5 Da.

LC-MS of octadecanal treated *Np* cAD:

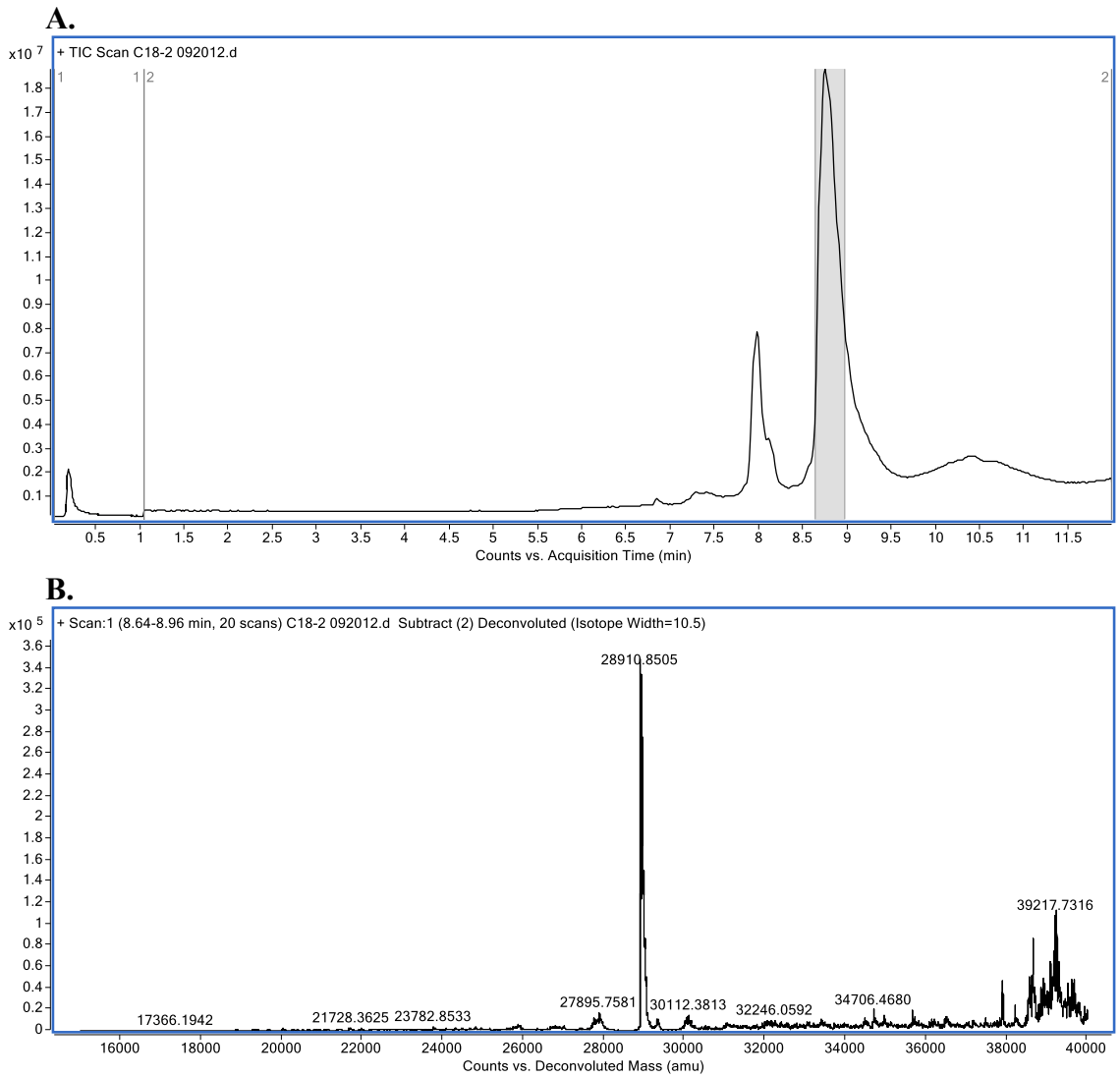


Figure S13. Reverse phase liquid chromatogram and mass spectral analysis of *Np* cAD reacted with 1-octadecanal. A. Total ion chromatogram (TIC) B. Deconvoluted mass spectrum of 1-octadecanal-reacted *Np* cAD. Highlighted region of A was extracted for mass spectral analysis. Octadecanal-reacted *Np* cAD elutes as a single peak with retention time 8.6 to 9.2 min. The M_r of octadecanal-reacted *Np* cAD was 28911 ± 0.5 Da which was indistinguishable from "as isolated" *Np* cAD.

LC-MS of cyclopropyl substrate, 6, reacted with *Np* cAD:

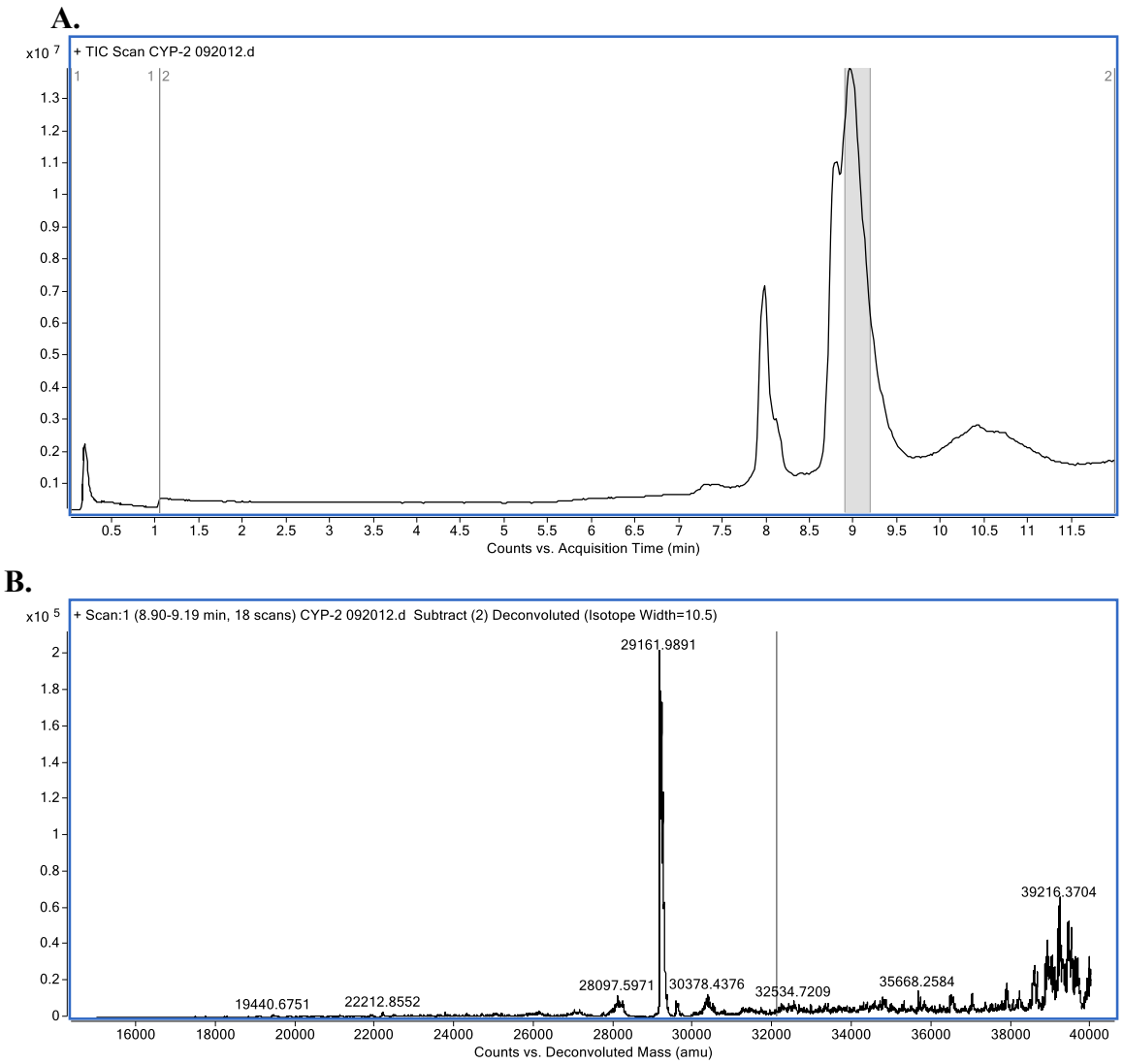


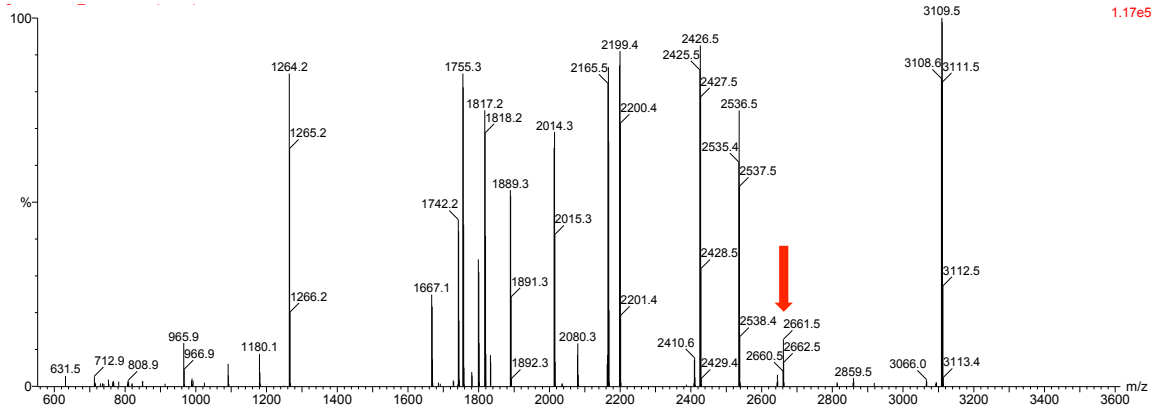
Figure S14. Reverse phase liquid chromatogram and mass spectral analysis of *Np* cAD incubated with cyclopropyl substrate **6.** A. Total ion chromatogram (TIC) showing both unmodified and modified protein. B. Deconvoluted mass spectrum of *Np* cAD reacted with **6**. Highlighted region of A was extracted for mass spectral analysis. *Np* cAD reacted with **6** shows two peaks; one peak with the same retention time as unmodified *Np* cAD and another peak with slightly higher retention time with a higher M_r of 29162 ± 0.5 Da.

The increase in molecular weight of 251 ± 0.5 Da is consistent with the formation of a covalent adduct between decarbonylated **6** and cAD.

MALDI-TOF data analysis. To determine the location of the covalent modification, samples of the inactivated and unmodified enzyme were subjected to proteolytic digestion with either trypsin or Glu-C and studied by MALDI-TOF (Micromass TofSpec-2E). α-cyano-4-hydroxycinnamic acid was used as a matrix. The instrument was calibrated using an external reference of five standard peptide of known mass.

Proteins were denatured with 8 M urea and reduced with 10 mM DTT prior to alkylation of cysteine residues using 50 mM iodoacetamide. The resulting alkylated protein was diluted to reduce the urea concentration to 1.5 M and treated with sequencing grade modified trypsin (Promega) and Glu-C separately overnight at 37 °C. The proteolytic fragments were analyzed by MALDI-TOF mass spectrometry and the spectra of the modified and unmodified enzyme digests compared. Analysis of the spectra identified two peptides, one Glu-C-derived (Figure S15A), the other trypsin-derived (Figure S16A), which were absent from the spectra of the covalently modified enzyme. Significantly, the peptides overlapped in sequence and encompass a 20-residue segment, CFAIAAYNIYIPVADDFARK, that forms part of the hydrophobic substrate-binding channel of cAD.

A.



B

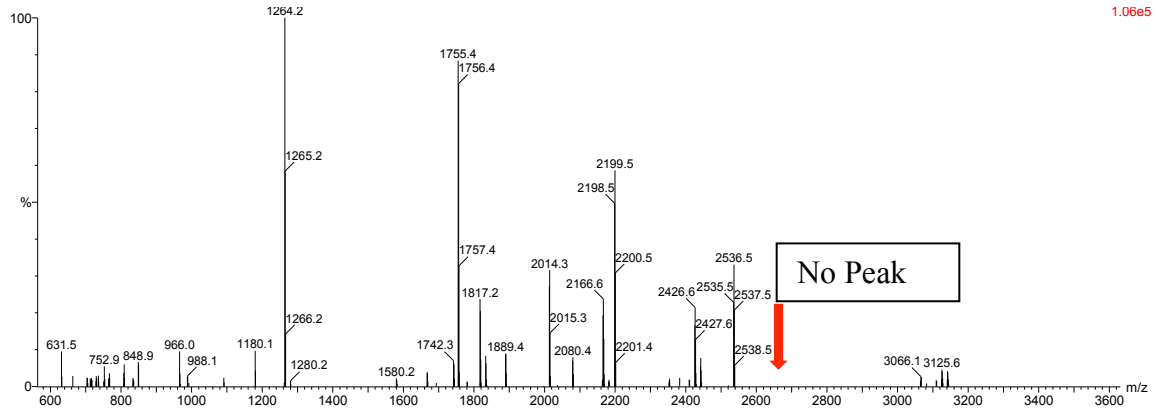


Figure S15. MALDI spectrum of GluC digests of *Np* cAD (A) and **6** treated *Np* cAD (B). Red arrow on spectrum **A** shows the peak of interest with mass 2661.5 Da that represents carbamidomethylated CFAIAAYNIYIPVADDFARKIT peptide fragment that is absent in spectrum **B**.

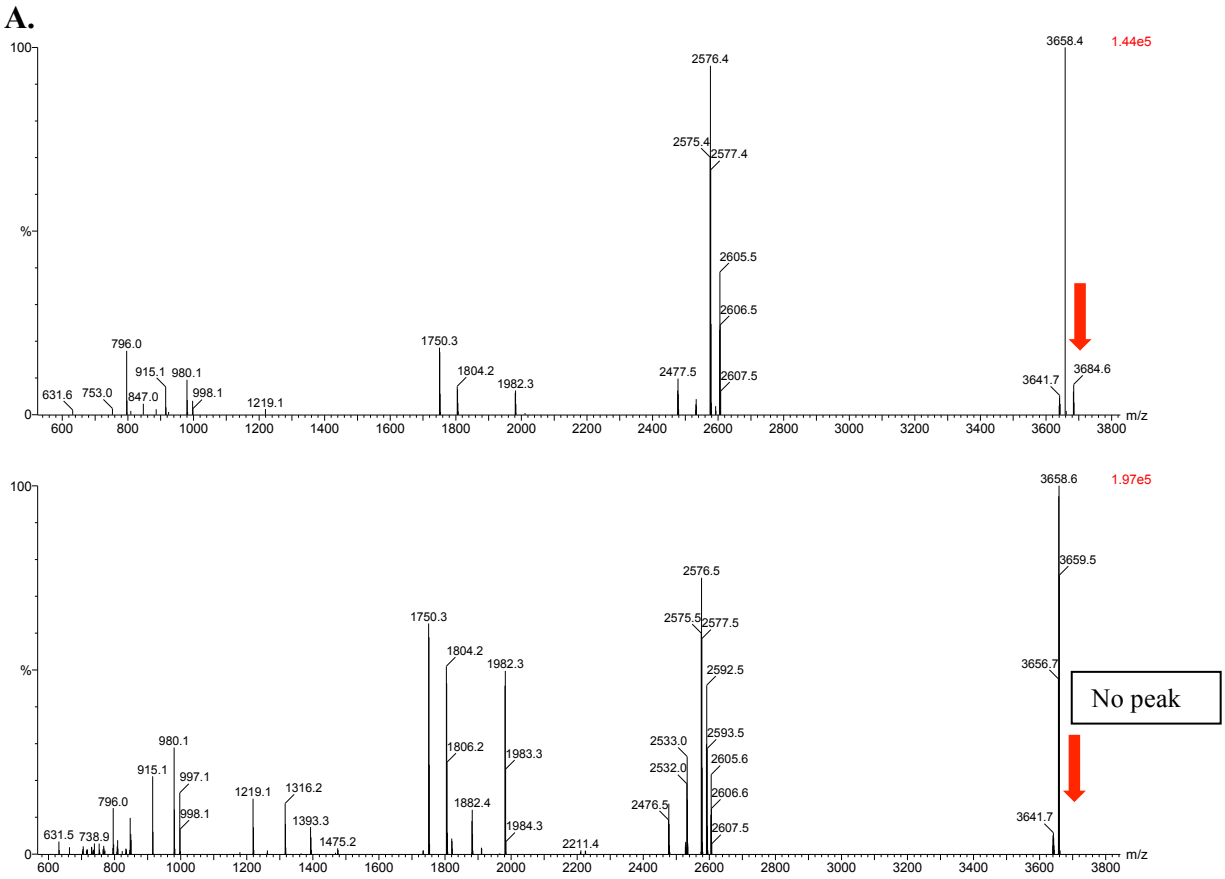


Figure S16. MALDI spectrum of trypsin digests of *Np* cAD (A) and **6** treated *Np* cAD (B). Red arrow on spectrum **A** shows the peak of interest with mass 3684.6 Da that represents VVTCLLIQSLIIECFAIAAYNIYIPVADDFARK peptide fragment that is absent in spectrum **B**.

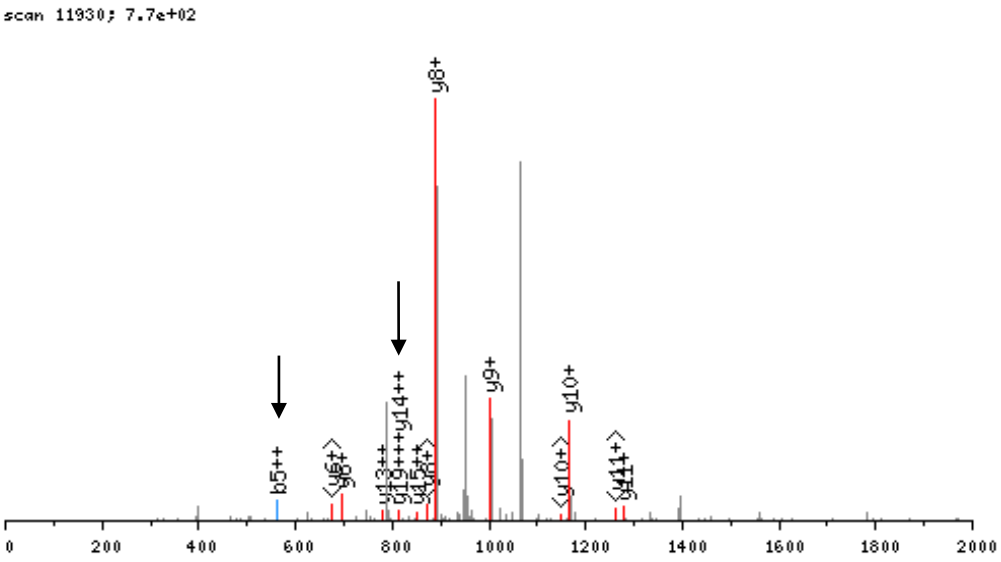
LC-Tandem mass analysis. Tandem mass analysis was performed by ion-trap mass spectrometer (LTQ-XL, ThermoFisher). Proteins were digested as described before for MALDI-TOF analysis. The samples were further acidified with trifluoroacetic acid and peptides were purified using SepPak C18 cartridge (Waters). The resulting sample was injected into a C18 reverse phase column (Aquasil) equilibrated with 5% acetonitrile/1%

acetic acid and peptides eluted with a linear gradient of increasing acetonitrile from 5 % to 60% over 40 min at a flow rate of 300 nl/min. The eluting peptides were directly introduced into ion-trap mass spectrometer (LTQ-XL) equipped with a nano-spray source. A full MS scan (m/z 400-2000) was acquired and the most abundant 6 ions were studied by MS/MS mode (relative collision energy ~ 35%). Raw files were searched against an *E. coli* database appended with the *Np* cAD sequence and a decoy database using X!Tandem (www.thegpm.org). The modification of phenylalanine, tyrosine, histidine, or glutamic acid, by a mass increment of 250.4 Da, together with methionine oxidation (+ 16 Da) and carbamidomethylation of cysteine (+ 57 Da) were considered as variables. The mass tolerance limit was set at 1 Da for precursor peptides and 0.5 Da for fragmented peptides. Results were further analyzed by Trans-Proteomic Pipeline (TPP) analysis, including PeptideProphet and ProteinProphet.[10]

This analysis succeeded in identifying one covalently modified peptide, present in low-abundance, in the tryptic digest. The secondary ion mass spectrum of this peptide displayed a fragmentation pattern that was consistent with Phe107 being modified by an additional mass of 251 Da (Figure S17). Phe107 forms part of the hydrophobic substrate channel of cAD and would be within ~ 5 Å of the putative alkyl radical formed by the opening of the cyclopropyl ring of **6**. These results suggest that the reaction of the product alkyl radical with the phenylalanine ring results in the covalent attachment of the alkyl fragment to the protein, thereby inactivating the enzyme.

A

IECFAIAAYNIYIPVADDFAR, MH+ 2933.2853, m/z 978.4333



B

b^{+}	b^{2+}	#	AA	#	y^{+}	y^{2+}	y^{3+}
114.0913	57.5496	1	I	21			
493.6068	247.3073	2	E	20	2820.2007	1410.6043	940.7388
653.7456	327.3767	3	C	19	2440.6852	1220.8465	**814.2336**
1051.3222	526.1650	4	F	18	2280.5464	1140.7771	760.8540
1122.3594	**561.6836**	5	A	17	1882.9698	941.9888	628.3285
1235.4434	618.2256	6	I	16	1811.9327	906.4702	604.6494
1306.4805	653.7442	7	A	15	1698.8486	**849.9282**	566.9547
1377.5176	689.2627	8	A	14	1627.8115	**814.4097**	543.2757
1540.5810	770.7944	9	Y	13	1556.7744	**778.8911**	519.5967
1654.6239	827.8159	10	N	12	1393.7110	697.3594	465.2422
1767.7080	884.3579	11	I	11	**1279.6681**	640.3380	427.2279
1930.7713	965.8896	12	Y	10	**1166.5841**	583.7959	389.5332
2043.8554	1022.4316	13	I	9	**1003.5207**	502.2643	335.1788
2140.9081	1070.9580	14	P	8	**890.4367**	445.7222	297.4841
2239.9765	1120.4922	15	V	7	793.3839	397.1959	265.1332
2311.0137	1156.0107	16	A	6	**694.3155**	347.6617	232.1104
2426.0406	1213.5242	17	D	5	623.2784	312.1431	208.4313
2541.0675	1271.0377	18	D	4	508.2514	254.6296	170.0890
2688.1360	1344.5719	19	F	3	393.2245	197.1162	131.7467
2759.1731	1380.0904	20	A	2	246.1561	123.5819	82.7239
		21	R	1	175.1190	88.0634	59.0449

Figure S17. Linear trap quadrupole (LTQ) mass spectral analysis of trypsin digests of *Np* cAD after reaction with **6**. **A**. Mass spectrum of the peptide fragment IECFAIAAYNIYIPVADDFAR. **B**. Peptide ions with different charges. Presence of b5++ and y19+++ ions is consistent with F107 residue of *Np* cAD modified with the hydrocarbon chain.

The proposed site of modification is consistent with the structure of cAD.

The crystal structure of *Np* cAD has not been solved, however the structure of cAD from *Prochlorococcus marinus* MIT9313 (*Pm* cAD, PDB ID 2OC5A) has been solved as part of a structural proteomics project.[11] An alignment of the sequences of these two proteins reveals 62% sequence identity and 79% sequence similarity, indicating a high degree of structural similarity. F107 in *Np* cAD is conserved in *Pm* cAD (F130) and in all other know cAD sequences. Based on the structure of *Pm* cAD (Figure S18), F107 forms part of the hydrophobic substrate channel and would be within ~ 5 Å of the putative alkyl radical formed by the opening of the cyclopropyl ring of **6.**

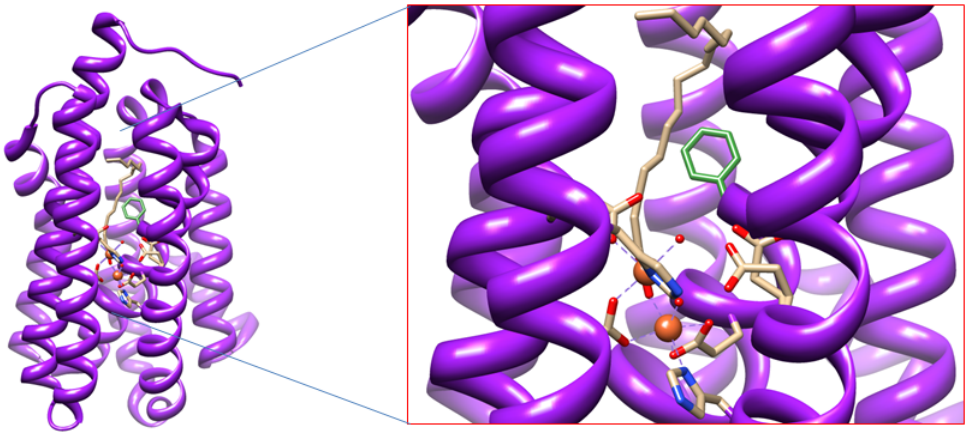


Figure S18: Crystal structure of cAD from *Prochlorococcus marinus* MIT9313 (PDB ID 2OC5A) showing di-iron active site and co-crystalized long chain fatty acid. The phenylalanine residue as shown in green most likely undergoes covalent modification after incubation of cAD with cyclopropyl aldehyde **6**.

Source of proton in 1-octadecene.

We initially investigated the source of the new proton in the product 1-octadecene formed in the reaction of **6** with cAD by performing the reaction in deuterated buffer (pD 7.6). The assay buffer was made as described before except H_2O was replaced with 99.9% D_2O (maximum H_2O in the buffer ~ 2%). Performing the reaction in deuterated buffer resulted in an increase in *m/z* for the 1-octadecene molecular ion from 252.3 to 253.3, consistent with the incorporation of a single deuterium atom. Less intense peaks with *m/z* = 254.3 and 255.3 were also noted that arise through non-enzymatic exchange of protons at the α-carbon of the cyclopropyl aldehyde (Figure S19). Interestingly, a less abundant but still significant peak at *m/z* = 252.3 was also observed. This suggests that some protons may derive from a non-exchangeable position on the protein.

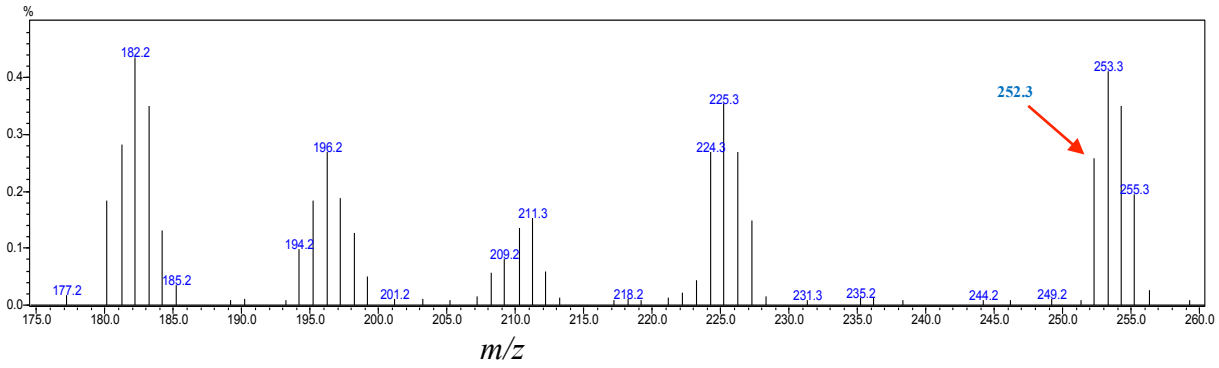
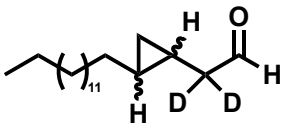


Figure S19. A section of mass spectrum of cyclopropyl product from **6** in deuterated buffer showing molecular ion peak of more intense mono-deuterated 1-octadecene of *m/z* 253.3 and less intense peaks of *m/z* 254.3 and 255.3 that arise through non-enzymatic exchange of the acidic hydrogen at the α-carbon of **6**. The peak of m/z 252.3 derives from non-exchangeable position of the protein.

In further experiments, we prepared di-deuterated **6,** in which the α-carbon was di-deuterated to avoid complications arising from exchange of aldehyde C-α protons during the experiment.

Preparation of di-deuterated 6: **6** (final concentration 10 mM) was dissolved in 90% DMSO-d_6, 10% in D_2O, buffered with 10 mM HEPES at pD 8.4. The solution was shaken for 12 h at 37 °C after which time GC-MS analysis revealed that the α-protons were greater than 95 % exchanged (Figure S20). The stock solution was stored frozen at -80 °C for further use in deuterium incorporation assays.



Di-deuterated 6

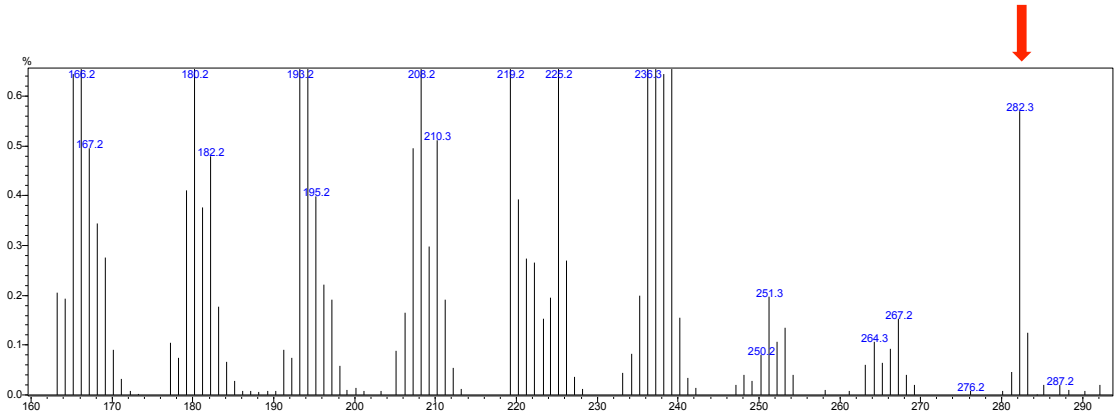


Figure S20. A section of mass spectrum of di-deuterated cyclopropyl substrate **6** showing molecular ion peak of *m/z* 282.3.

When di-deuterated **6** was reacted with cAD in deuterated buffer the predominant molecular ion for 1-octadecene had *m/z* = 255.3 (Figure S21), corresponding to tri-deuterated alkene. This is consistent with the proton coming from the solvent or a solvent-exchangeable group on the enzyme. However, a less abundant but still significant peak at *m/z* = 254.3 was also observed. This suggests that some protons may derive from a non-exchangeable position on the protein.

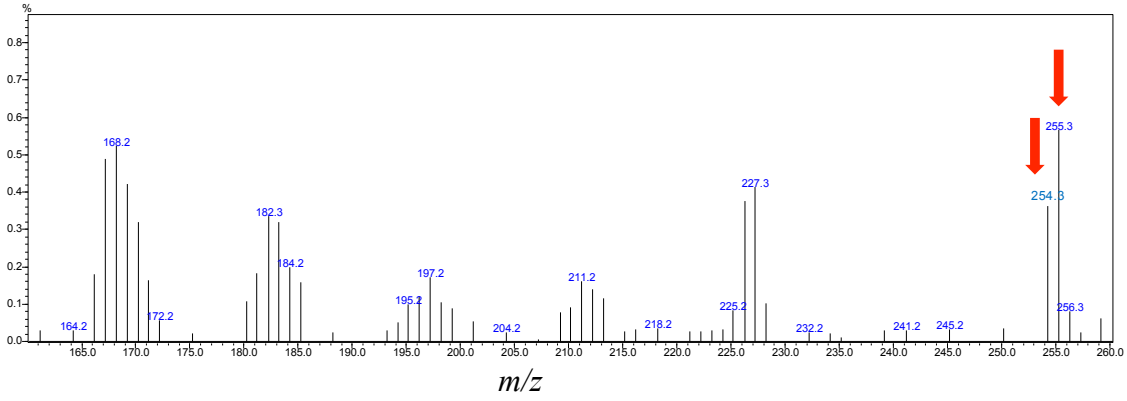


Figure S21. A section of mass spectrum of enzymatically obtained product from di-deuterated **6** in deuterated buffer, showing a more intense tri-deuterated 1-octadecene molecular ion peak of *m/z* 255.3 and a less intense peak of di-deuterated 1-octadecene of

molecular ion peak of *m/z* 254.3. The peak of *m/z* = 254.3 was derived from non-exchangeable protons presumably from the side chains of the protein.

References

1. De Mico, A.; Margarita, R.; Parlanti, L.; Vescovi, A.; Piancatelli, G. *J. Org. Chem.* **1997**, *62*, 6974.
2. Schlosser, M.; Christmann, K.F. *Angew. Chem. Intl. Ed.* **1966**, *5,* 126.
3. Rajesh, M.; Sen, J.; Srujan, M.; Mukherjee, K.; Sreedhar, B.; Chaudhuri, A. *J. Am. Chem. Soc.* **2007**, *129*, 11408.
4. Furukawa, J.; Kawabata, N.; Nishimura, J. *Tetrahedron* **1968**, *24,* 53.
5. Lebel, H. L.; Marcoux, J.-F.; Molinaro, C.; Charette, A. B. *Chem. Rev*. **2003**, *103*, 977.
6. Das, D.; Eser, B.E.; Han, J.; Sciore, A.; Marsh, E.N.G. *Angew. Chem. Intl. Ed.* **2011**, *50,* 7148.
7. Eser, B.E.; Das, D.; Han, J.; Jones, P.R.; Marsh, E.N.G. *Biochemistry* **2011**, *50,* 10743.
8. Halaka, F. G.; Babcock, G. T.; Dye, J. L. *J. Biol. Chem.* **1982**, *257*, 1458.
9. Whitteck, J.T.; Cicchillo, R. M.; van der Donk, W. A. *J. Am. Chem. Soc.* **2009**, *131*, 16225
10. Cao, F.; Chen, Y.; Cierpicki, T.; Liu, Y.; Basrur, V.; Lei, M.; Dou, Y. *PLos one* **2010**, *5*, e14102.
11. Unpublished *structure solved by Joint Center of Structural Genomics (protein database entry PDB|2OC5|A)*.