A Three Enzyme Pathway for 2-Amino-3-Hydroxycyclopent-2-Enone Formation and Incorporation in Natural Product Biosynthesis

*Wenjun Zhang, Megan L. Bolla, Daniel Kahne, and Christopher T. Walsh**

Department of Biological Chemistry & Molecular Pharmacology, Harvard Medical School and Department of Chemistry & Chemical Biology, Harvard University

E-mail: Christopher_Walsh@hms.harvard.edu

SUPPORTING INFORMATION

Table of Contents

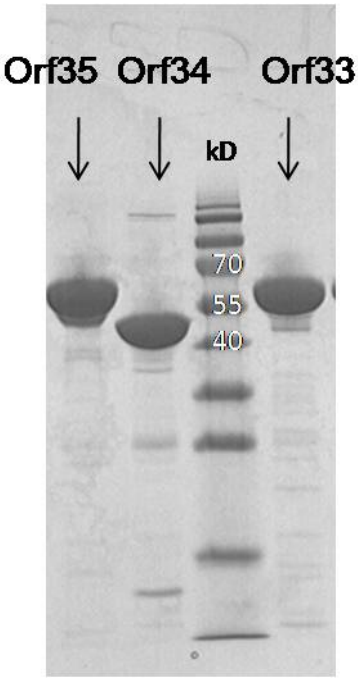


Figure S1. SDS-PAGE of ORF33-35 purified from *E. coli.* Ready Gel Tris-HCl Gel (12% precast, Biorad) was used.

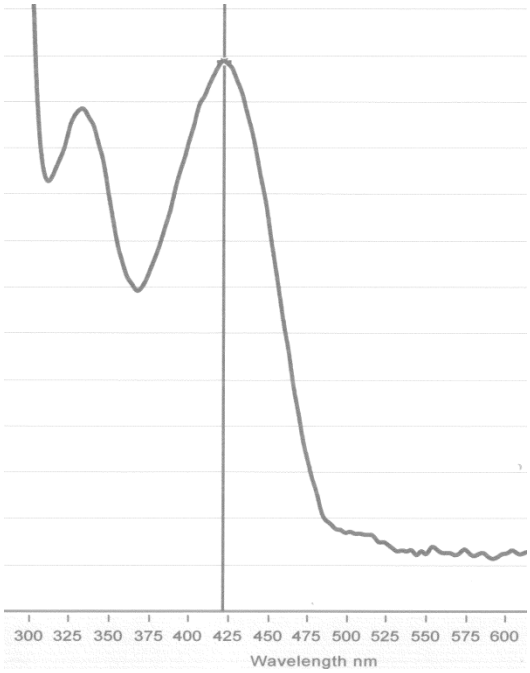


Figure S2. UV profile of purified ORF34 from *E. coli*. UV absorption spectra identified a characteristic λ_{max} at 420 nm, indicating the presence of an enzyme-bound PLP Schiff base.

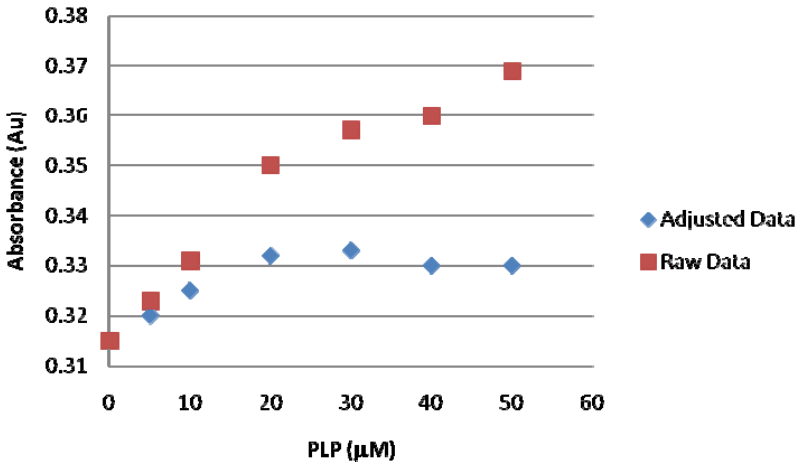


Figure S3. Titration of *E. coli* purified ORF34 with PLP. The experiments were performed as previously described.[1] Reactions were carried out with 50 µM ORF34 in 100 µL of buffer containing 50 mM HEPES, pH 8.0. PLP was added in 1 µL aliquots (500 µM, 0.1 equivalents), and monitored by UV absorbance at 420 nm to generate the raw data. A PLP standard curve (not shown) was generated by adding 1 µL aliquots of 500 µM PLP to a 100 µL 50 mM HEPES solution while monitoring at 420 nm. The adjusted data curve was generated by subtracting the PLP standard curve from the titration raw data. The adjusted data revealed a break-point in the linear phase around 20 µM (0.4 enzyme equivalents), indicating the fraction of protein isolated in the holo form was roughly 60%.

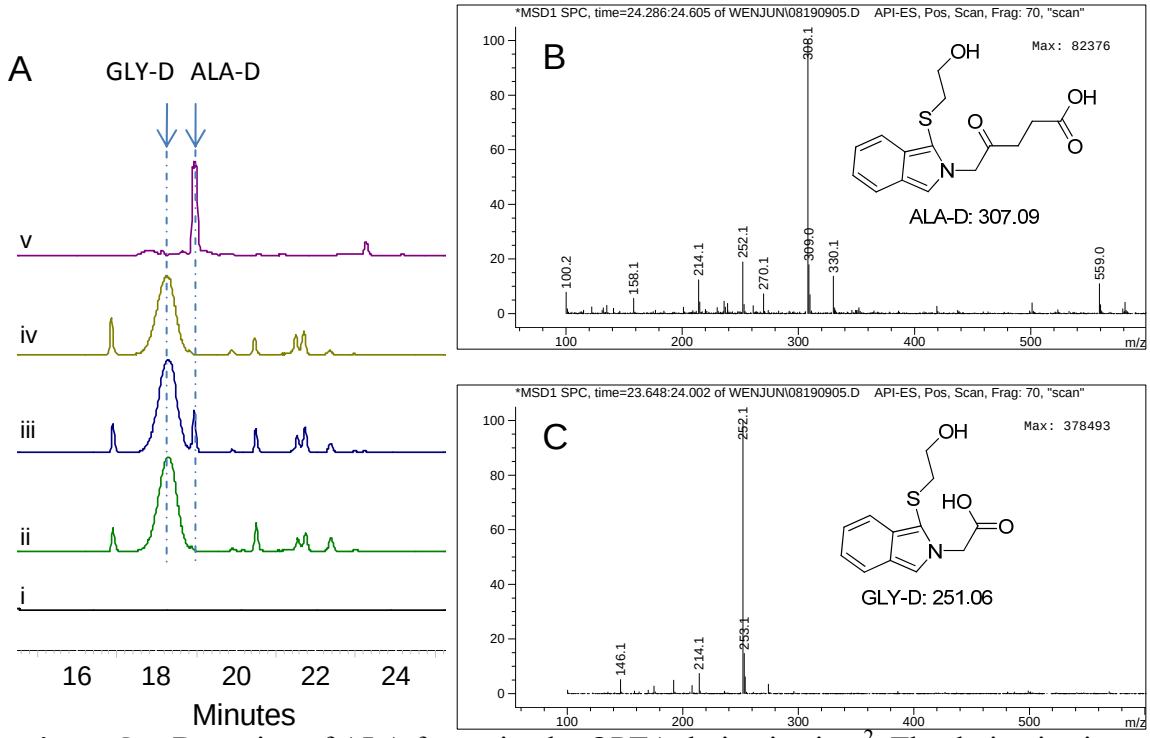


Figure S4. Detection of ALA formation by OPTA derivatization.[2] The derivatization reagent was obtained by mixing 1 mL of 80 mM *o*-phthalaldehyde (OPTA) solution and 1 mL of 98 mM 2-mercaptoethanol (2ME) solution (both prepared in 0.2 M NaOH). The enzyme was removed by centrifugation through a 5 kDa MMCO filter tube at 13000 rpm for 20 min after which, 80 µL of filtered reaction mixture was combined with 80 µL of derivatization reagent. After vortexing for 10 sec, the mixture was incubated for 2 min and subjected to LC-MS analysis. LC-MS analysis was performed with a linear gradient of 5 to 95% CH_3CN (v/v) over 30 min, and 95% CH_3CN (v/v) further for 10 min in H_2O supplemented with 0.1% (v/v) formic acid at a flow rate of 0.5 mL/min. (A) UV traces (340 nm) from LC-MS analysis of the ORF34 reaction (trace iii) and controls with no glycine (trace i), no succinyl-CoA (trace ii, no enzyme control is the same as trace ii), glycine standard(trace iv), and ALA standard (trace v). (B) Mass spectrum of OPTA derivatized ALA measured during LC-MS. (C) Mass spectrum of OPTA derivatized glycine measured during LC-MS. GLY-D and ALA-D are the indicated isoindole derivatives.

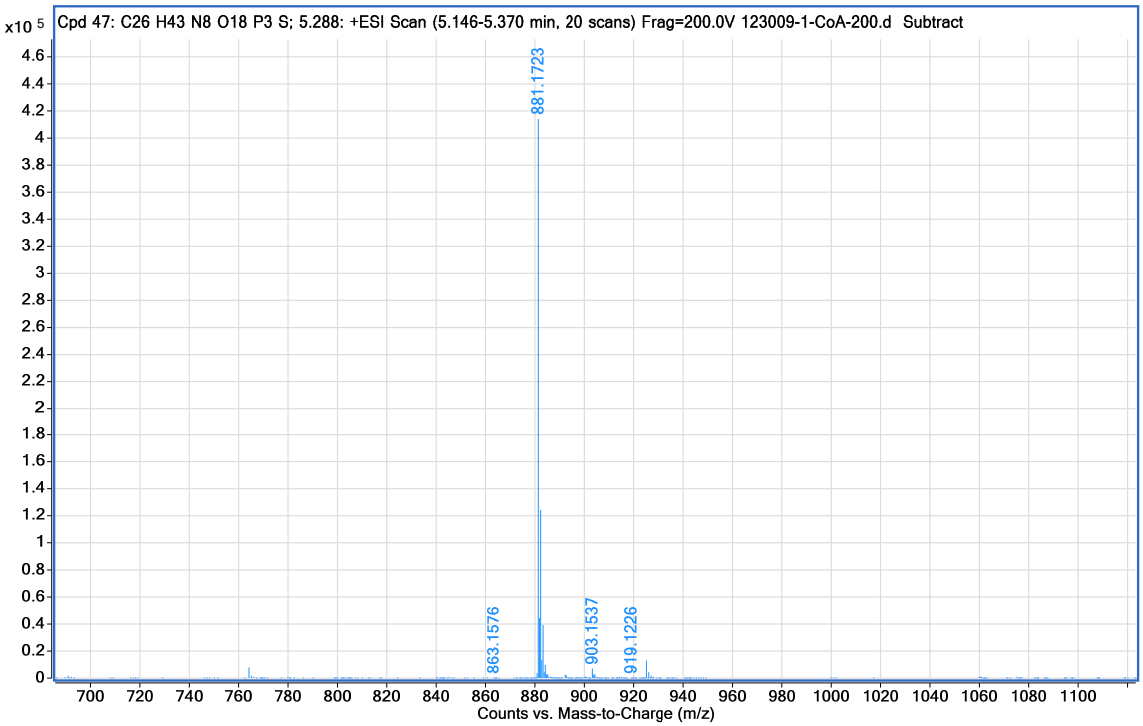


Figure S5. HRMS of ALA-CoA (m/z = 881.1723 $[M + H]^+$, Δ = 2.1 mmu) measured during LC-MS.

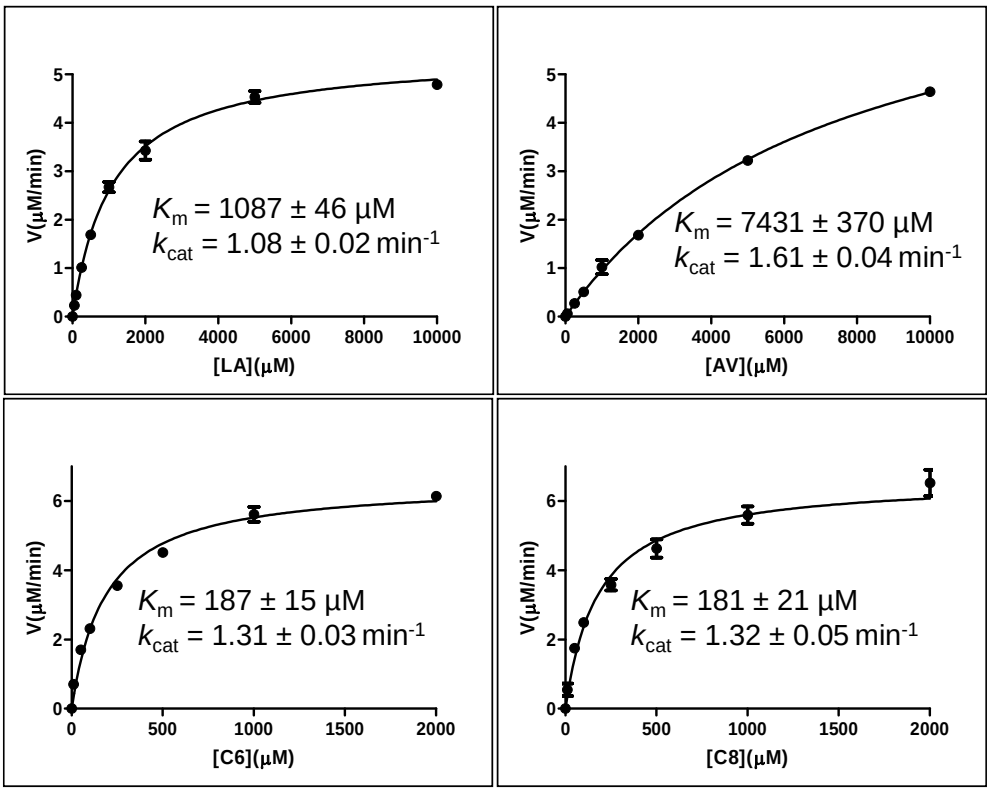


Figure S6. Determination of ORF35 kinetic parameters for various acid substrates by ATP-PP_i release assay. LA: levulinic acid; AV: 5-aminovaleric acid; C6: hexanoic acid; C8: octanoic acid. The inorganic pyrophosphate released by enzymatic reaction was measured continuously using EnzChek Pyrophosphate Assay Kit (Invitrogen). A typical assay contained, in a total volume of 100 μL, 0-10 mM acid substrate, 5 mM ATP, 2 mM $MgCl_2$, and 5 μM ORF35 in 50 mM HEPES, pH 8.0. MESG substrate, purine nucleoside phosphorylase and inorganic pyrophosphatase were added according to the protocol. Reactions were initiated by the addition of acid substrate and monitored at 360 nm on a SPECTRAmax plus 384 96-well plate reader. Initial velocities were calculated using a standard curve for inorganic pyrophosphate. The data were fitted to the Michaelis-Menten equation in GraphPad Prism to obtain estimates for k_{cat} and K_m. Error bars represent standard deviations from at least three independently performed experiments.

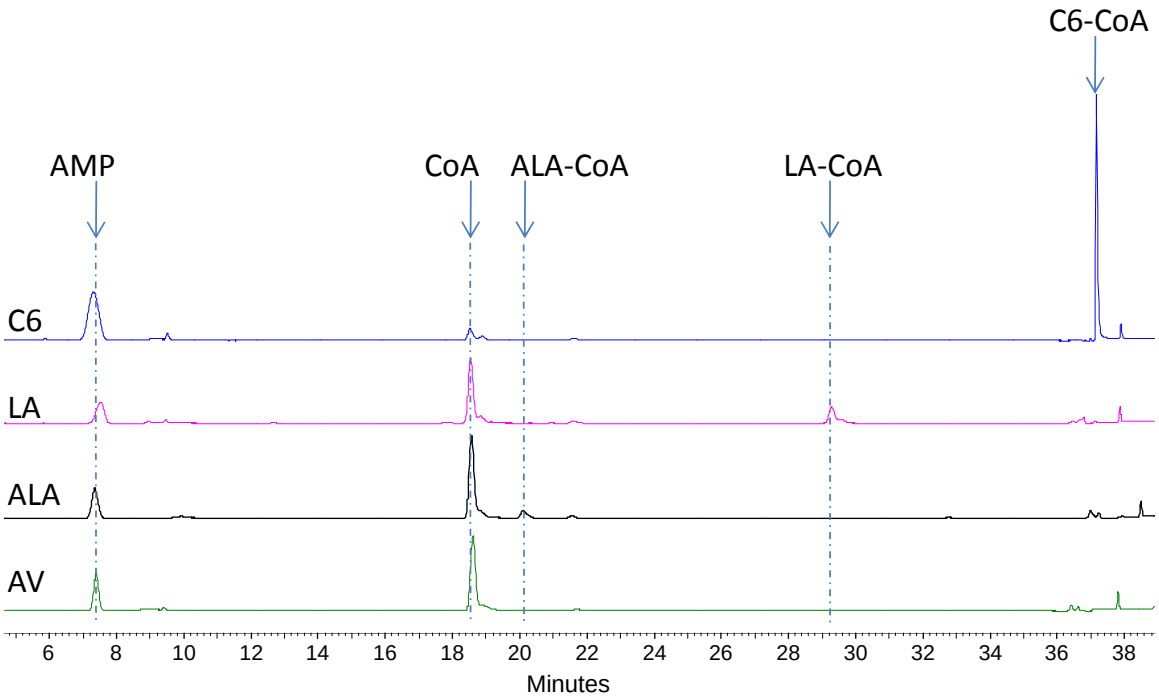
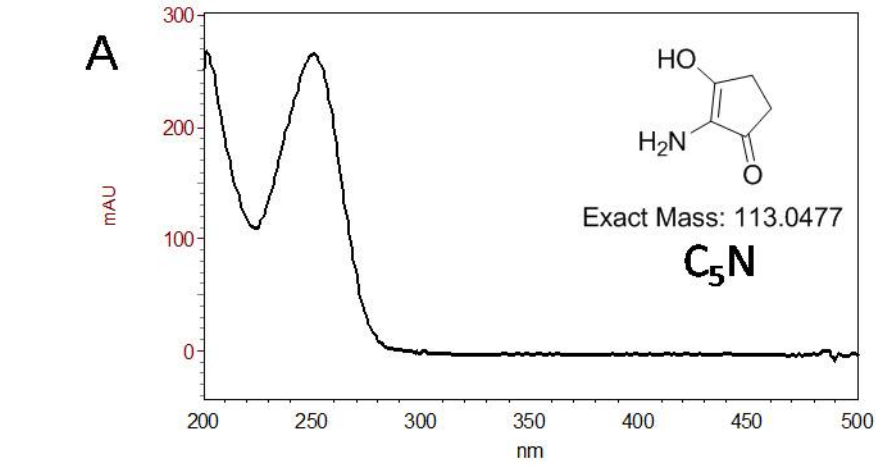
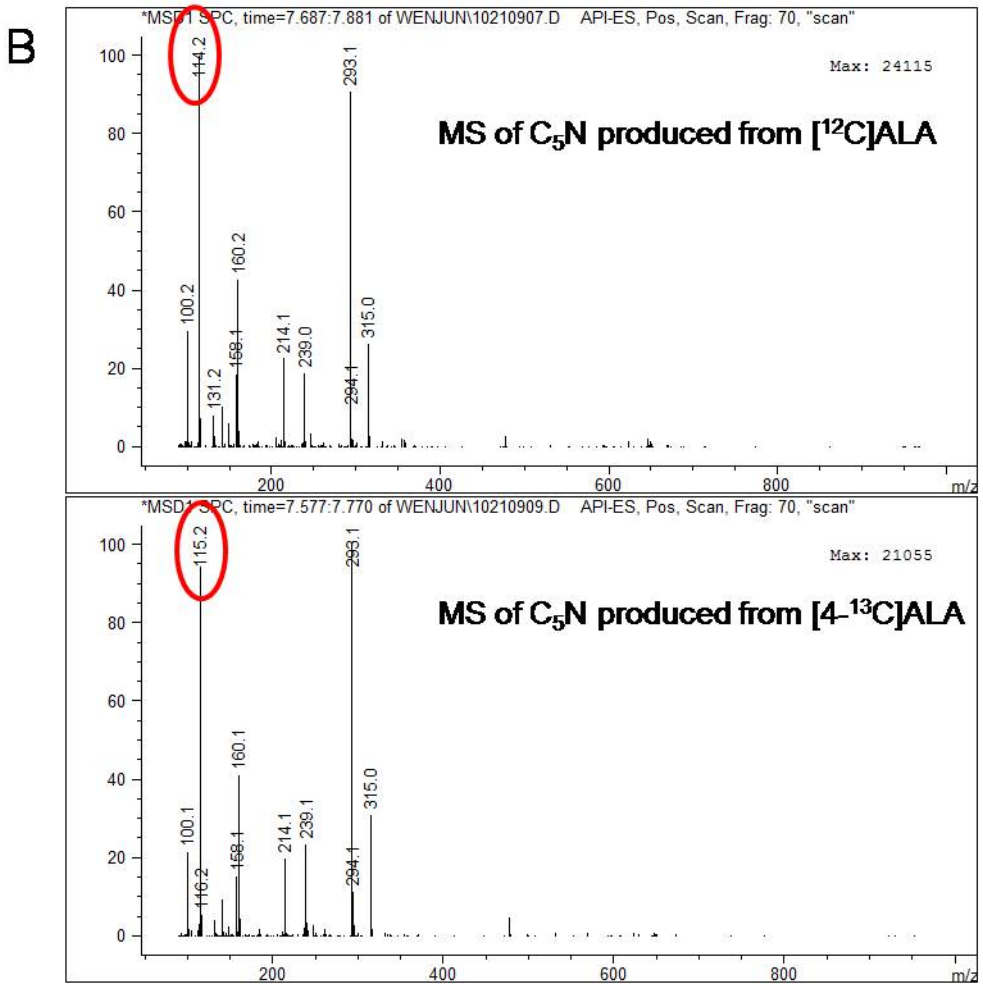


Figure S7. HPLC traces (260 nm) of ORF35 reactions with selected acid substrates. ALA: 5-aminolevulinic acid; LA: levulinic acid; AV: 5-aminovaleric acid; C6: hexanoic acid. The reaction was carried out by incubating 50 μM ORF35 with 2-5 mM acid substrates, 2 mM ATP, 2 mM CoA and 2 mM $MgCl_2$ for 30 min. A concentration of 2 mM was used for C6 to minimize the DMSO concentration in the assay, while a concentration of 5 mM was used for all other acids dissolved in water. The reactions were stopped by addition of tricholoacetic acid (TCA) (final concentration of 5% v/v) and subjected to both LC-MS and HPLC analysis.


A
300
200
100
0
mAU
200
250
300
350
400
450
500
nm
HO
H2N
O
Exact Mass: 113.0477
C5N


B
*MSD1 SPC, time=7.687:7.881 of WENJUN\10210907.D API-ES, Pos, Scan, Frag: 70, "scan"
Max: 24115
MS of C5N produced from [12C]ALA
114.2
293.1
100.2
160.2
214.1
239.0
315.0
131.2
158.1
294.1
*MSD1 SPC, time=7.577:7.770 of WENJUN\10210909.D API-ES, Pos, Scan, Frag: 70, "scan"
Max: 21055
MS of C5N produced from [4-13C]ALA
115.2
293.1
100.1
160.1
214.1
239.1
315.0
116.2
158.1
294.1
m/z

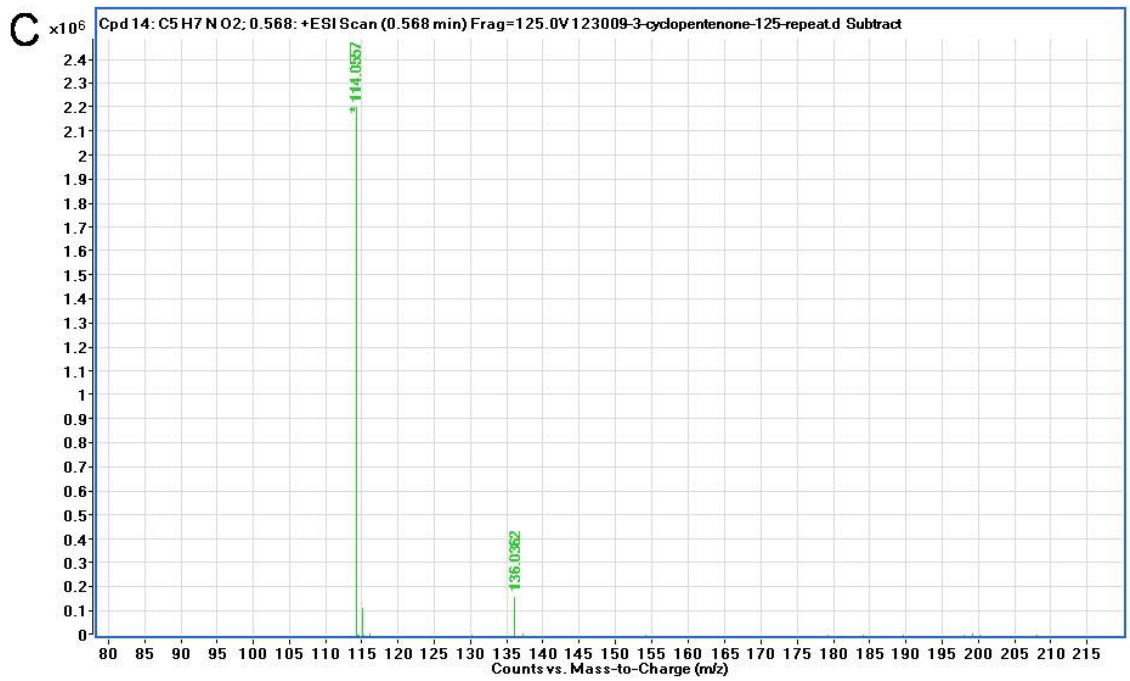


Figure S8. Spectroscopic characterization of C_5N. (A) C_5N UV profile (λ_{max} = 250 nm). (B) LRMS of C_5N using [^{12}C]ALA (top panel) and [4-^{13}C]ALA (bottom panel) starting material. (C) HRMS of C_5N using [^{12}C]ALA starting material (m/z = 114.0557 $[M + H]^+$, Δ = 0.7 mmu).

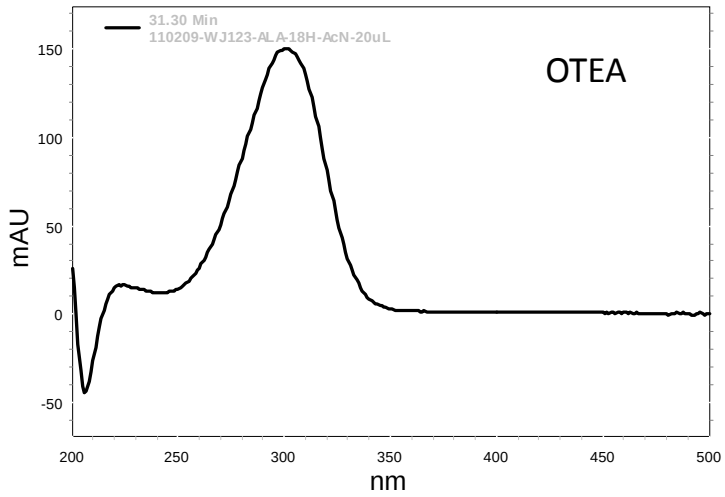


Figure S9. UV profile of 2,4,6-octatrienoic acid (OTEA) showing a λ_{max} of 300 nm measured during HPLC.

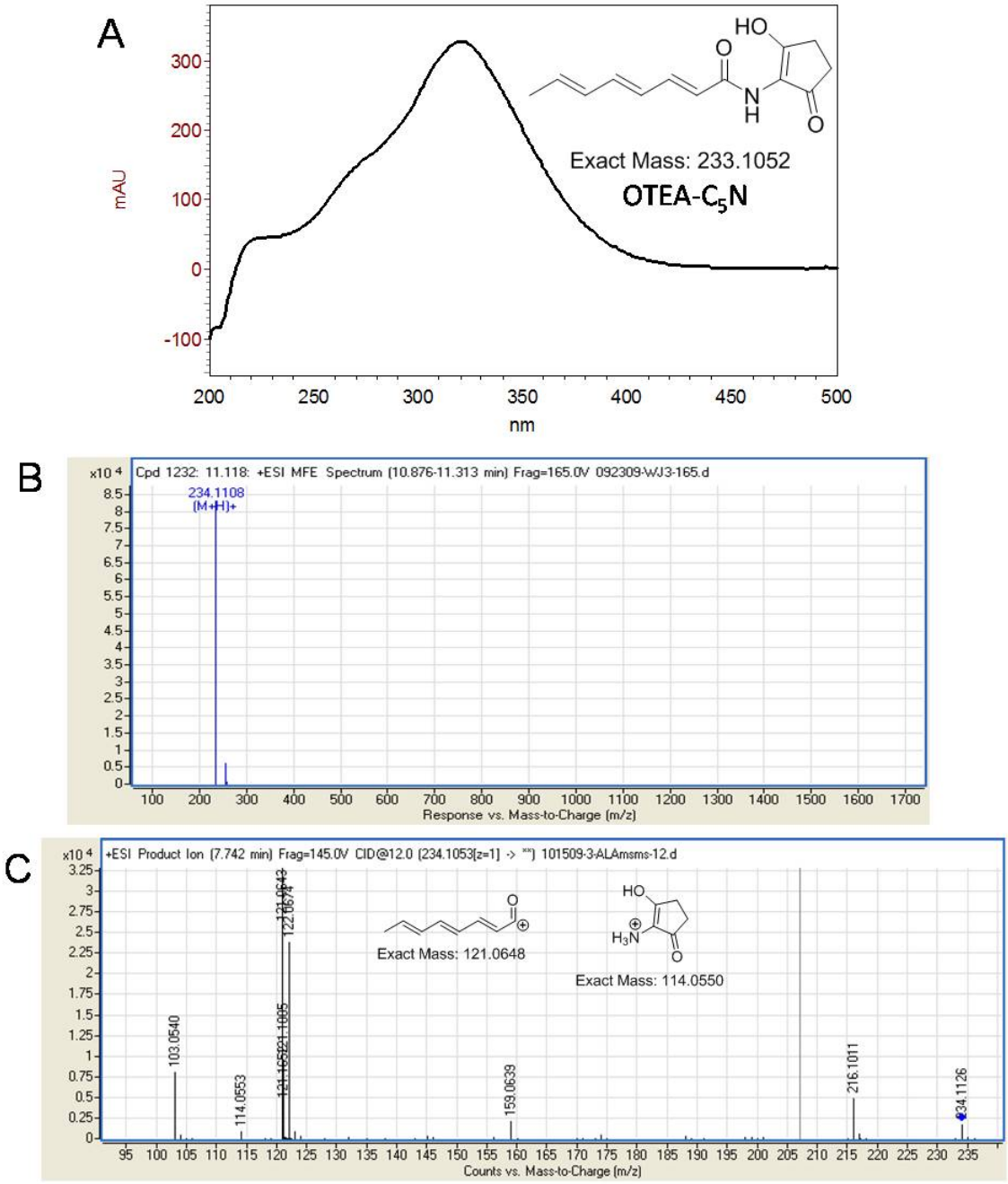
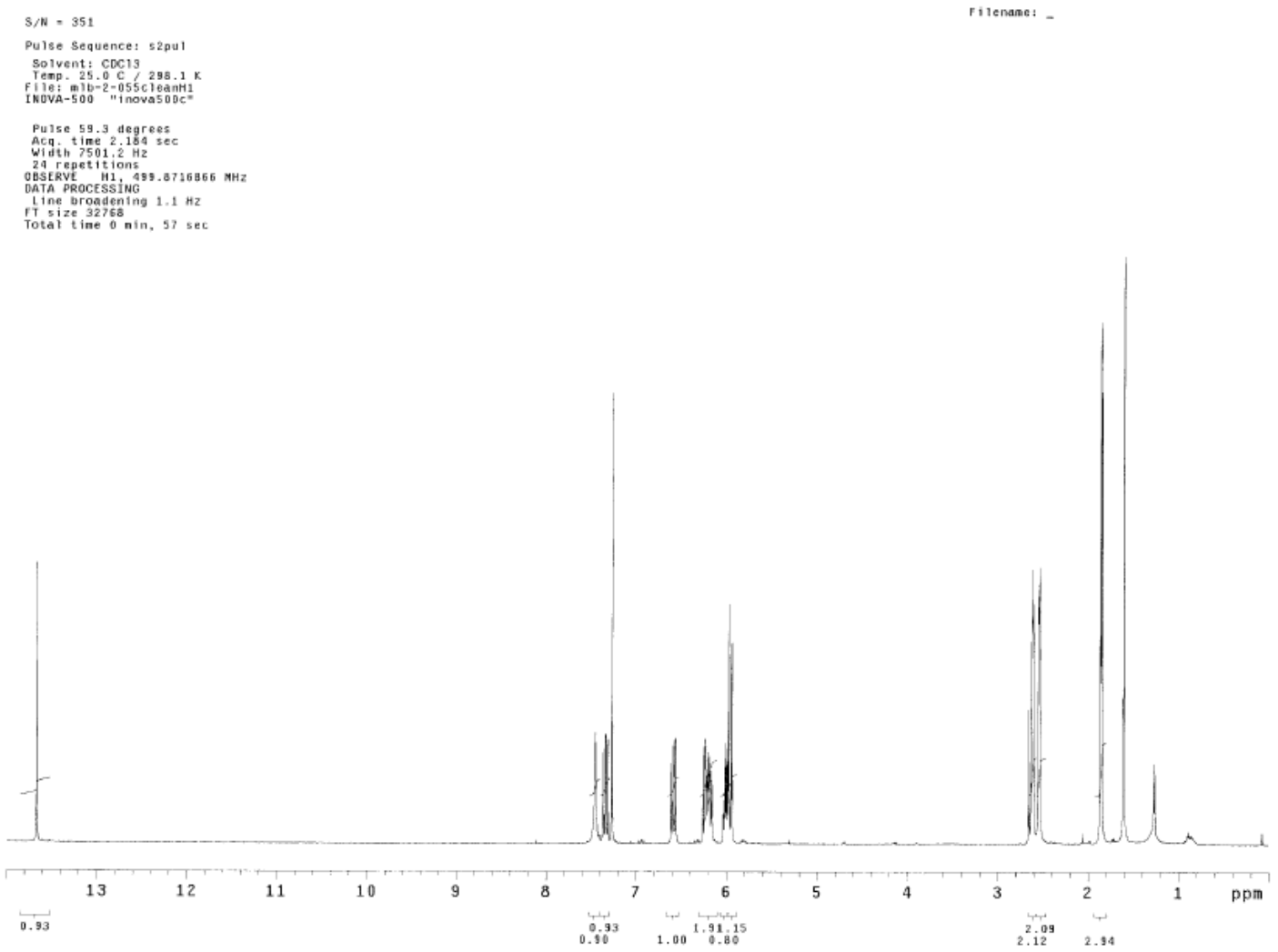


Figure S10. Spectroscopic characterization of OTEA-C_5N. (A) OTEA-C_5N UV profile (λ_{max} = 321 nm). (B) HRMS of OTEA-C_5N (m/z = 234.1108 $[M + H]^+$, Δ = 1.7 mmu). (C) MS fragmentation of OTEA-C_5N.


S/N = 351
Pulse Sequence: s2pul
Solvent: CDCl3
Temp. 25.0 C / 298.1 K
File: mlb-2-055cleanH1
INOVA-500 "inova500c"
Pulse 59.3 degrees
Acq. time 2.184 sec
Width 7501.2 Hz
24 repetitions
OBSERVE H1, 499.8716866 MHz
DATA PROCESSING
Line broadening 1.1 Hz
FT size 32768
Total time 0 min, 57 sec
Filename: _
13
12
11
10
9
8
7
6
5
4
3
2
1
ppm
0.93
0.93
0.90
1.00
1.91
1.15
0.80
2.09
2.12
2.94

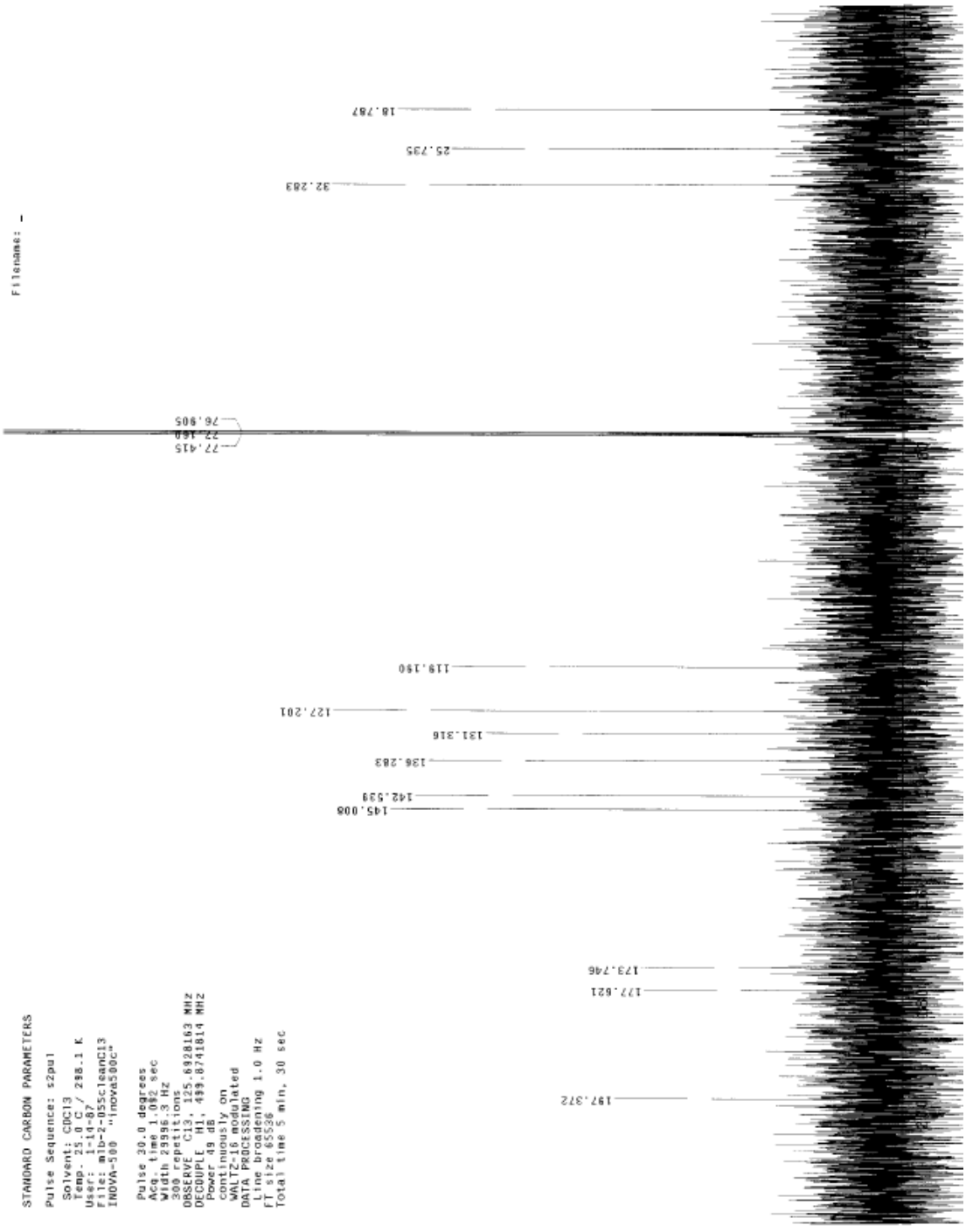


Figure S11. ^{1}H and ^{13}C NMR characterization of synthesized OTEA-C_5N standard.

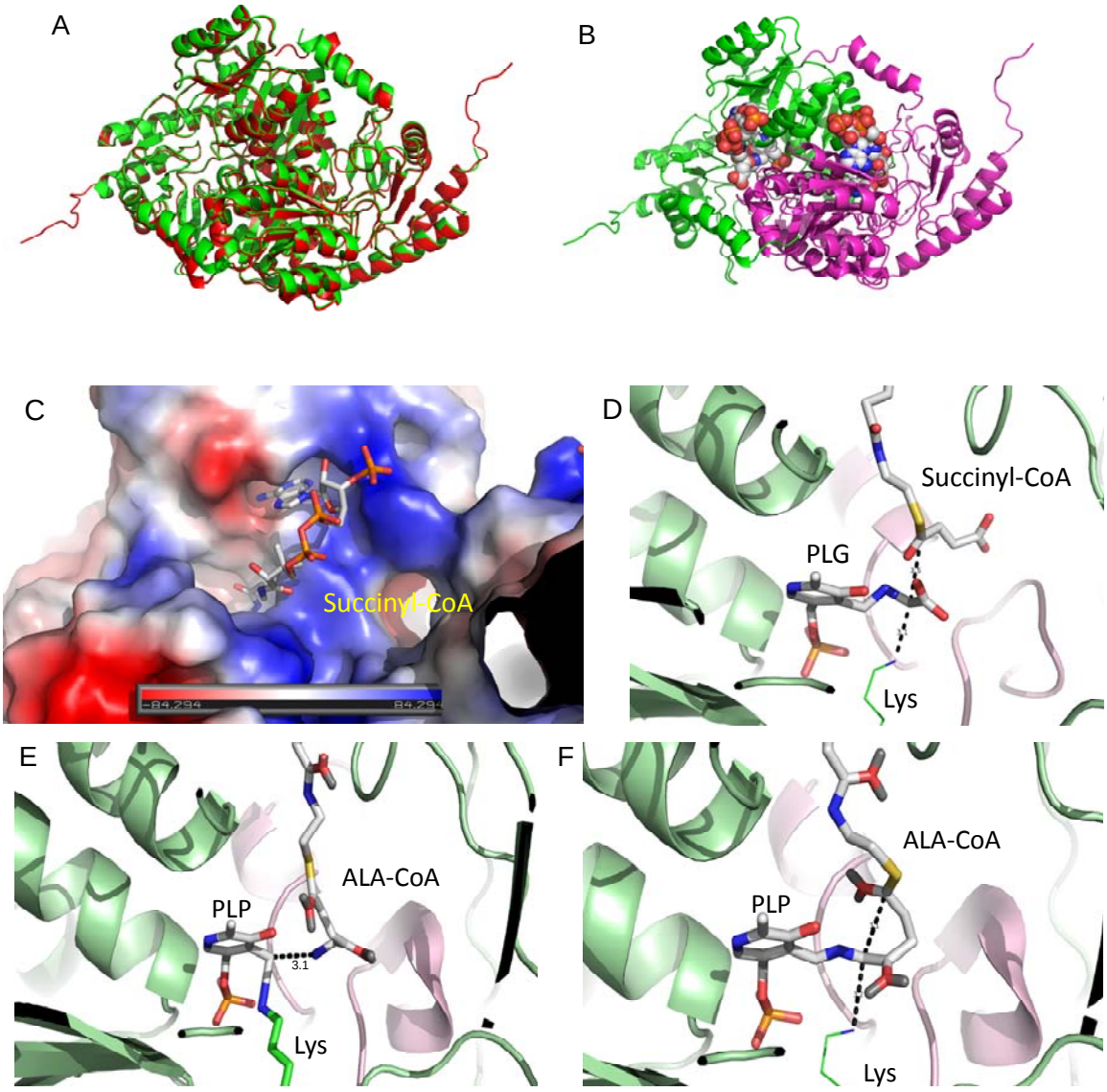


Figure S12. Modeled three-dimensional structure of ORF34. The structure was modeled on *R. capsulatus* ALAS using online program HHpred[3] and the substrates were docked using software GOLD. (A) Overlay of modeled ORF34 dimer (red) onto *R. capsulatus* ALAS structure (green, PDB: 2BWN). (B) Modeled ORF34 dimer bound with PLP, glycine and succinyl-CoA. (C) Electrostatic surface model of ORF34 bound with succinyl-CoA. (D) Details of the active site model of ORF34. The figure shows the relationship between external aldimine Gly=PLP, succinyl-CoA and Lys_{248}. (E) Details of the active site model of ORF34 showing the relationship between internal aldimine Lys=PLP and ALA-CoA. (F) Details of the active site model of ORF34 showing the proposed relationship between external aldimine C5-ALA=PLP and Lys_{248}.

References:

[1] Aron, Z. D.; Dorrestein, P. C.; Blackhall, J.R.; Kelleher, N.L.; Walsh, C. T. *J. Am. Chem. Soc.* **2005**, *127*, 14986-14987.

[2] Morineau, G.; Azoulay, M.; Frappier, F. *J. Chromatogr.* **1989**, *467*, 209-216.

[3] Soding, J.; Biegert, A.; Lupas, A.N. *Nucl. Acids. Res.* **2005**, *33*, W244-248.