

Supplementary Information for

N-Carboxyanhydride-Mediated Fatty Acylation of Amino Acids and Peptides for Functionalization of Protocell Membranes

Enver Cagri Izgu, Anders Björkbom, Neha P. Kamat, Victor S. Lelyveld, Weicheng Zhang, Tony Z. Jia, and Jack W. Szostak

Table of Contents

1. General Methods	S2
2. Analysis of Peptide Formation Reactions	S3–S5
3. Mechanistic insights into the acylation of amino acids and peptides formed <i>in situ</i>	S6–S8
4. Hydrolysis of Val-NCA	S9
5. Standard curves of FRET efficiencies	S9
6. Synthesis of <i>N</i> ^α -Octanoyl-L-valine	S10
7. Synthesis of <i>N</i> ^α -Oleoyl-L-tryptophan	S10–S11
8. Synthesis of Octanoic-1,1- ¹⁸ O ₂ acid.	S11–S12
9. Synthesis of <i>N</i> ^α -Oleoyl-L-arginine	S13
10. References for Supplementary Information	S13
11. ¹ H and ¹³ C NMR Spectra	S14–S20

General Methods

Anhydrous acetonitrile was purchased from Glen Research (Sterling, VA). Anhydrous dimethylformamide (DMF) and dichloromethane (CH_2Cl_2), amino acids (AAs), oleoyl chloride, and 4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS) were purchased from Sigma-Aldrich (St. Louis, MO). Oleic acid and octanoic acid were purchased from Nu-Chek Prep, Inc (Elysian, MN). Octanenitrile was purchased from Alfa-Aesar (Ward Hill, MA) and H_2^{18}O was purchased from Cambridge Isotope Laboratories (Andover, MA). These reagents were used without further purification unless noted. Purification of the synthesized molecules was performed using a Teledyne Isco combiflash instrument equipped with either a silica gel column or a C18Aq high performance RediSep[®] column. Nuclear Magnetic Resonance (NMR) spectroscopic analyses were carried out on either a Varian INOVA 400 MHz or a Bruker Ascend 400 MHz spectrometer. Chemical shifts (δ) for ^1H NMR spectra were referenced to TMS at $\delta = 0.00$ ppm, or to CHD_2OD at $\delta = 3.31$ ppm. ^{13}C NMR spectra were referenced to CDCl_3 at $\delta = 77.23$ ppm, or to CD_3OD at $\delta = 49.15$ ppm. The following abbreviations were used to describe NMR resonances: s (singlet), d (doublet), t (triplet), q (quartet), quin (quintet), m (multiplet), br (broad), and nfom (non-first order multiplet). Coupling constants (J) were reported in Hz. Low-resolution mass spectrometry (LRMS) analyses were performed using a Bruker Daltonics Esquire 6000 mass spectrometer. Liquid chromatography followed by high-resolution mass spectrometry (LC-HRMS) was carried out on an Agilent 6520 Q-TOF or an Agilent 6230 TOF LC/MS-MS instrument (for further details see *Methods* in the main text). UV measurements were conducted on an Amersham Biosciences, Ultraspec 3100-pro UV/VIS spectrophotometer. Membrane growth assays and microscopy of RNA-membrane localization were performed as described in *Methods*.

Table S1. Oligomerization products observed for the reaction of Val-NCA with the specified AA substrates in the presence of either oleic acid vesicles or octanoic acid. Typical reaction conditions: 0.04 mmoles of Val-NCA (suspension) in 2 mL reaction volume, 20 mM AA, 20 mM fatty acid and 300 mM EPPS buffer at an initial pH of 8.5–8.8. The pH of the reaction mixtures drops by up to 0.3 units during the course of the experiments. Aliquots, which were taken at the 24h time point, were diluted 1000 fold before being injected into the LC-HRMS instrument.

Entry	AA	Fatty acid: oleic acid		Fatty acid: octanoic acid	
		2	OI-3	2	Oc-3
1	Trp (W)	V _n W (n = 1–6)	OI-V _n W (n = 0–3)	V _n W (n = 1–6)	Oc-V _n W (n = 0–5)
2	Arg (R)	V _n R (n = 1–5)	OI-V _n R (n = 0–4)	V _n R (n = 1–3)	Oc-V _n R (n = 0–3)
3	Val (V)	V _n (n = 2–8)	OI-V _n (n = 1–6)	V _n (n = 2–8)	Oc-V _n (n = 1–6)
4	His (H)	V _n H (n = 1–5)	OI-V _n H (n = 0–5)	V _n H (n = 1–5)	Oc-V _n H (n = 0–5)
5	Lys (K)	V _n K (n = 1–5)	OI-K	V _n K (n = 1–4)	Oc-V _n K (n = 0–3)
6	Gln (Q)	V _n Q (n = 1–5)	OI-V _n Q (n = 0–5)*	V _n Q (n = 2–5)	Oc-V _n Q (n = 0–5)
7	Asn (N)	V _n N (n = 2–6)	OI-V _n N (n = 0–5)	V _n N (n = 2–5)	Oc-V _n N (n = 0–4)

*OI-V₃Q was not observed

Table S2. HRMS Analysis of the W-containing products outlined in entry 1, Table S1. Overlay of HRMS spectra are presented in the main text.

A. Oleoylated W-containing Products				B. Octanoylated W-containing Products			
Molecular Formula	Calc. m/z	Obs. m/z	Error (ppm)	Molecular Formula	Calc. m/z	Obs. m/z	Error (ppm)
OI-W: C ₂₉ H ₄₃ N ₂ O ₃ ⁻	467.3279	467.3278	-0.2	Oc-W: C ₁₆ H ₂₀ N ₃ O ₃ ⁻	329.1871	329.1873	0.7
OI-VW: C ₃₄ H ₅₂ N ₃ O ₄ ⁻	566.3963	566.3967	0.7	Oc-VW: C ₂₁ H ₂₉ N ₄ O ₄ ⁻	428.2555	428.2555	<0.1
OI-V ₂ W: C ₃₉ H ₆₁ N ₄ O ₅ ⁻	665.4647	665.4658	1.6	Oc-V ₂ W: C ₂₆ H ₃₈ N ₅ O ₅ ⁻	527.3239	527.3242	0.6
OI-V ₃ W: C ₄₄ H ₇₀ N ₅ O ₆ ⁻	764.5332	764.5335	0.4	Oc-V ₃ W: C ₃₁ H ₄₇ N ₆ O ₆ ⁻	626.3923	626.3917	-1.0
				Oc-V ₄ W: C ₃₆ H ₅₆ N ₇ O ₇ ⁻	725.4607	725.4607	<0.1
				Oc-V ₅ W: C ₄₁ H ₆₅ N ₈ O ₈ ⁻	824.5291	824.5301	1.2

C. V _n W Peptides			
Molecular Formula	Calc. m/z	Obs. m/z	Error (ppm)
VW: C ₁₆ H ₂₀ N ₃ O ₃ ⁻	302.1510	302.1503	-2.4
V ₂ W: C ₂₁ H ₂₉ N ₄ O ₄ ⁻	401.2194	401.2184	-2.6
V ₃ W: C ₂₆ H ₃₈ N ₅ O ₅ ⁻	500.2878	500.2867	-2.3
V ₄ W: C ₃₁ H ₄₇ N ₆ O ₆ ⁻	599.3563	599.3556	-1.1
V ₅ W: C ₃₆ H ₅₆ N ₇ O ₇ ⁻	698.4247	698.4255	1.2
V ₆ W: C ₄₁ H ₆₅ N ₈ O ₈ ⁻	797.4931	797.4931	<0.1

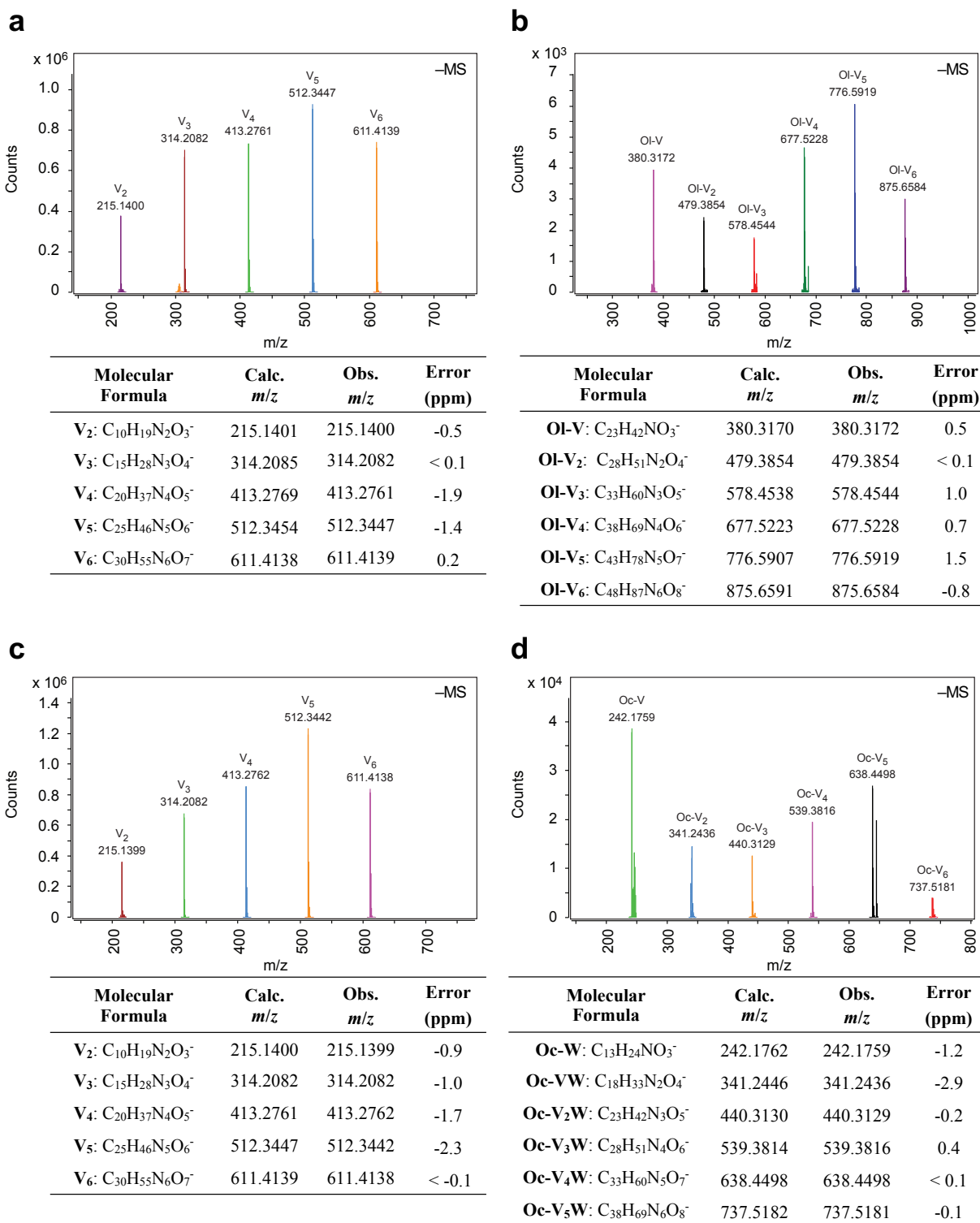


Figure S1. Selected ion-HRMS spectra ($[M-H]^{-1}$) of reaction mixtures (24 h) initially consisting of Val-NCA (20 mM), tryptophan (W) (20 mM), EPPS (300 mM) and either oleic acid (20 mM) (**a** and **b**), or octanoic acid (20 mM) (**c** and **d**). HRMS experiments were carried out in negative mode. Panels (**a**) and (**b**) show the oligo-V peptides observed in the presence of oleic acid. Panels (**c**) and (**d**) show the oligo-V peptides observed in the presence of octanoic acid. Ol- is oleoyl and Oc- is octanoyl.

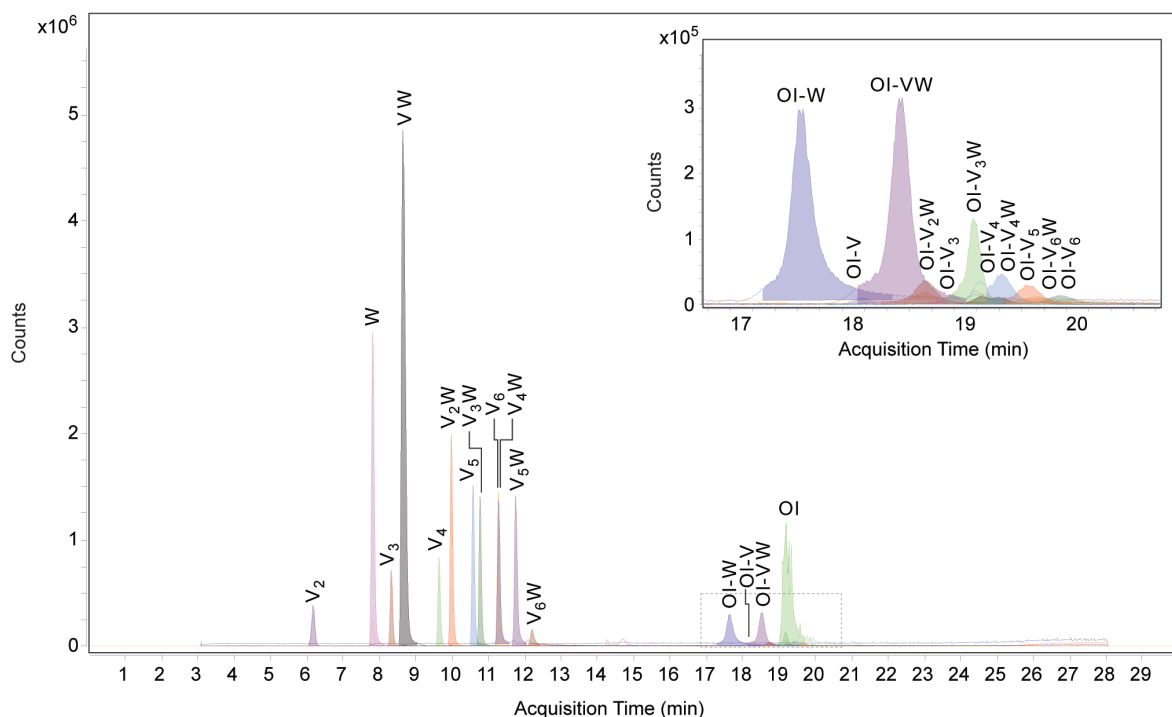


Figure S2. Overlay of HPLC extracted ion chromatogram of a reaction mixture initially containing Val-NCA, oleic acid and Trp, analysis of which is outlined in Fig. 1b-c in the main text.

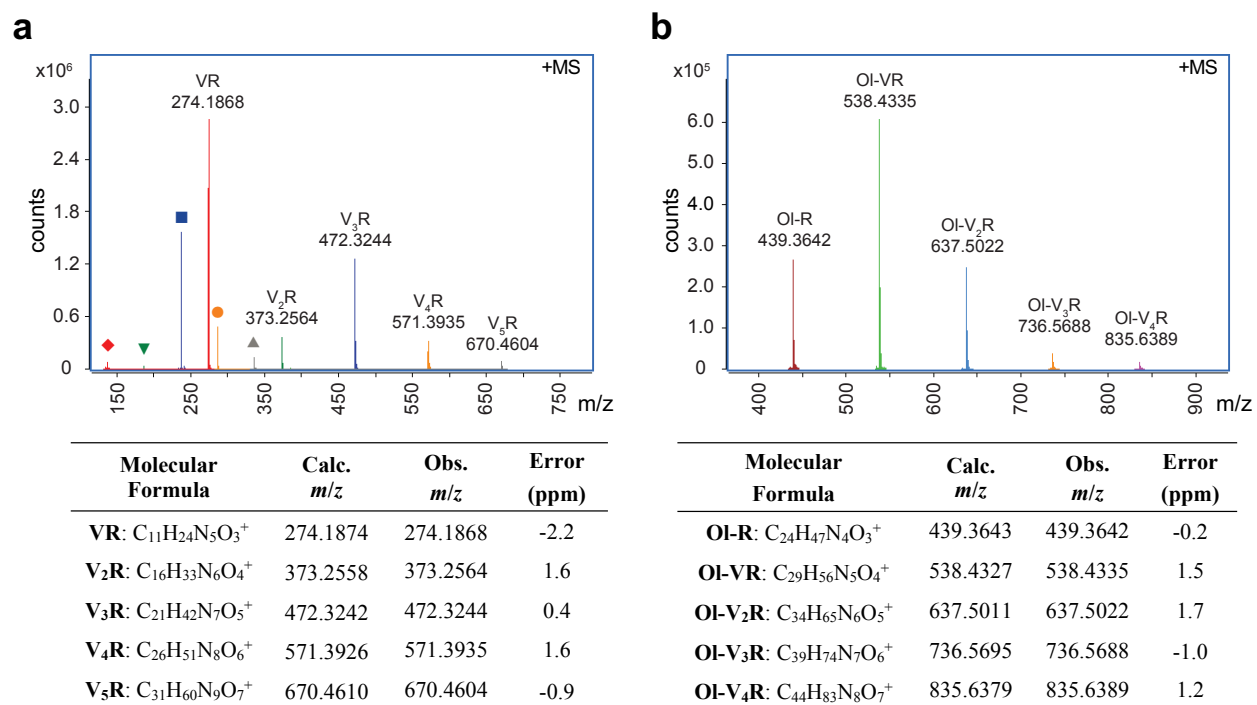


Figure S3. Selected ion-HRMS analyses of reaction mixtures (24 h) initially consisting of Val-NCA (20 mM), arginine (R) (20 mM), EPPS (300 mM) in oleic acid (20 mM) (a and b). HRMS experiments were carried out in positive mode. Panel (a) highlights the observed R-containing unacylated peptides in the presence of oleic acid. Panel (b) highlights the observed R-containing oleoylated (OI-) species. (a) symbols represent the observed second charge state m/z values of the peptide products (♦: VR, ▼: V₂R, ■: V₃R, ●: V₄R, ▲: V₅R).

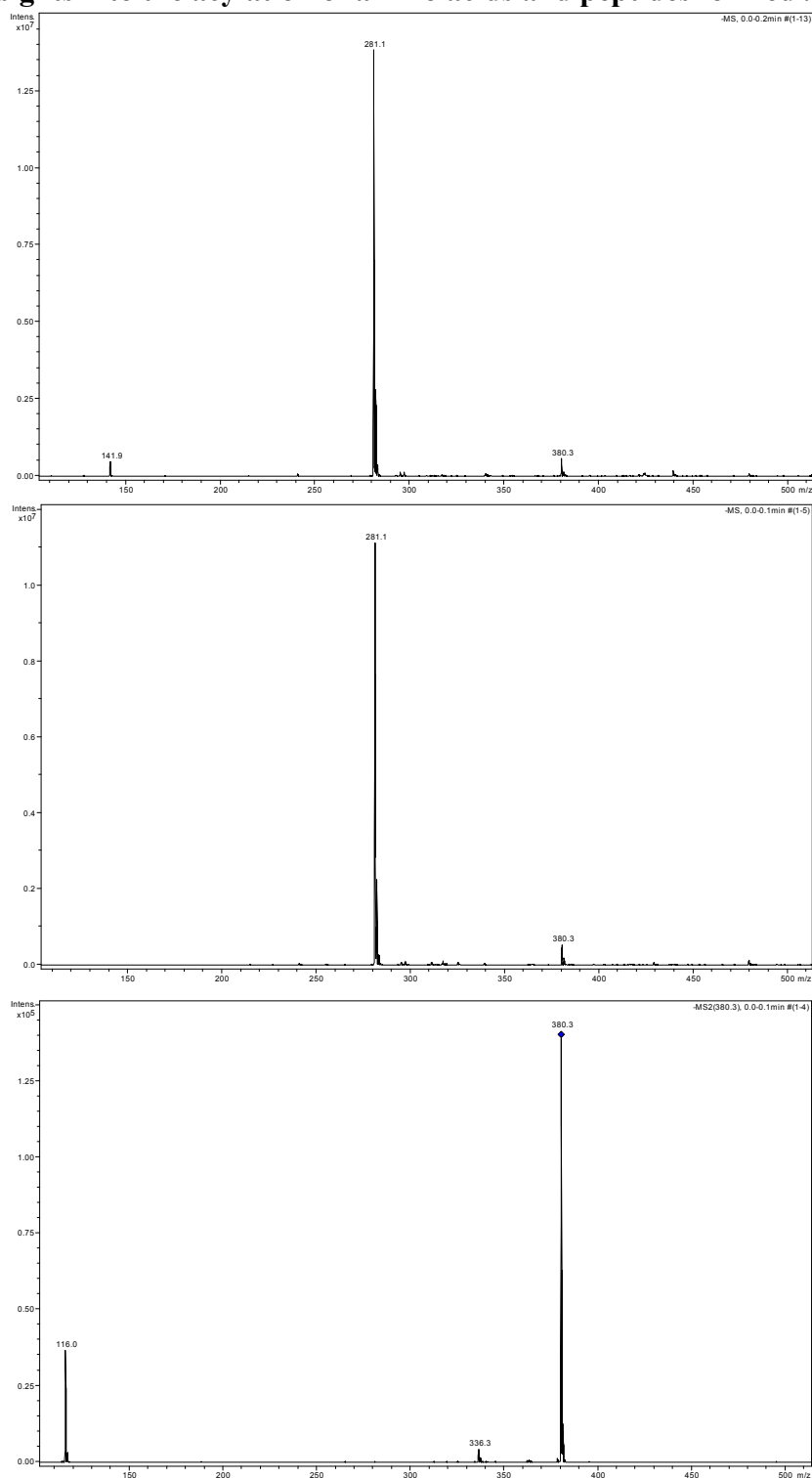
Mechanistic insights into the acylation of amino acids and peptides formed *in situ*

Figure S4. Mass spectroscopic investigation (in negative mode) of the formation of the carboxy-anhydride from Val-NCA, **1** (40 mM) and a mixture of oleic acid (1.5 M) and trimethylamine (1.5 M). ESI spectra of the crude reaction aliquot taken after (top) 15 min and (middle) 6 h of mixing. (Bottom) MS/MS product ion spectrum highlighting the isolated 380 m/z species and its fragmentation products. $[M - H]^{-1}$ m/z values; Val-NCA (**1**): 142, oleic acid: 281, N^{α} -oleoylvaline: 380, valine: 116, and decarboxylated N^{α} -oleoylvaline: 336.

Below in Fig. S5 we show MS analyses of lower concentration aqueous reactions in which we used 40 mM ^{18}O -enriched octanoic acid, 40 mM TEA, and ~ 20 mM Val-NCA at initial pH 8.7. We observed the $[\text{M} - \text{H}]^{-1} = 246$ m/z (Fig. S5, top) and $[\text{M} - \text{H}]^{-1} = 244$ m/z (bottom) signals in negative ionization mode. The existence of the 246 m/z ($^{18}\text{O}_2$ -oleylvaline) species, which fragments to yield a y-ion of 118 m/z (mono- ^{18}O enriched valine carboxylate) from amide bond dissociation and a decarboxylation product with 200 m/z (mono- ^{18}O enriched), suggests that the intramolecular rearrangement (path *i*) occurred. This result is in agreement with Fig. 3c of the main text. Interestingly, when isolating individual peaks within the isotope distribution around this species, we observed that the fragmentation of the 244 m/z signal (^{18}O -oleylvaline) yields y-ions of both 116 and 118 m/z, as well as the decarboxylated ions 200 and 198 m/z, respectively. Partitioning of ^{18}O with the y-ion from fragmentation of ^{18}O -oleylvaline would be expected for an anhydride that arises from ^{18}O -oleate, since the enriched oxygen could be in either the bridging or non-bridging position in the mixed anhydride intermediate. The mono- ^{18}O -oleate needed to generate this species could arise from an impurity in the enriched preparation of $^{18}\text{O}_2$ -oleic acid (Fig. S6, bottom) or from the hydrolysis of the anhydride generated by the reaction of $^{18}\text{O}_2$ -oleate with Val-NCA. In fact, during

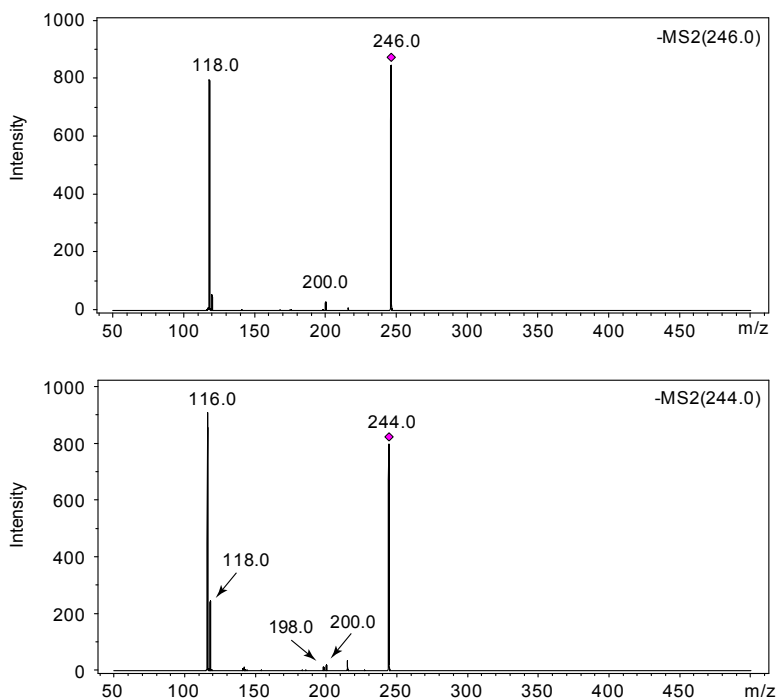


Figure S5. Mechanistic investigation of the formation of a mixed anhydride from Val-NCA and ^{18}O -enriched octanoic acid. (Top and bottom) MS/MS product ion spectra highlighting the isolated ^{18}O -enriched species and their fragmentation products.

the direct MS injection of a crude reaction (Fig. S6), we did observe an increase in the abundance of the 145 m/z peak corresponding to mono- ^{18}O -oleate, suggesting that a pathway for hydrolysis of the mixed anhydride exists in a reaction containing Val-NCA.

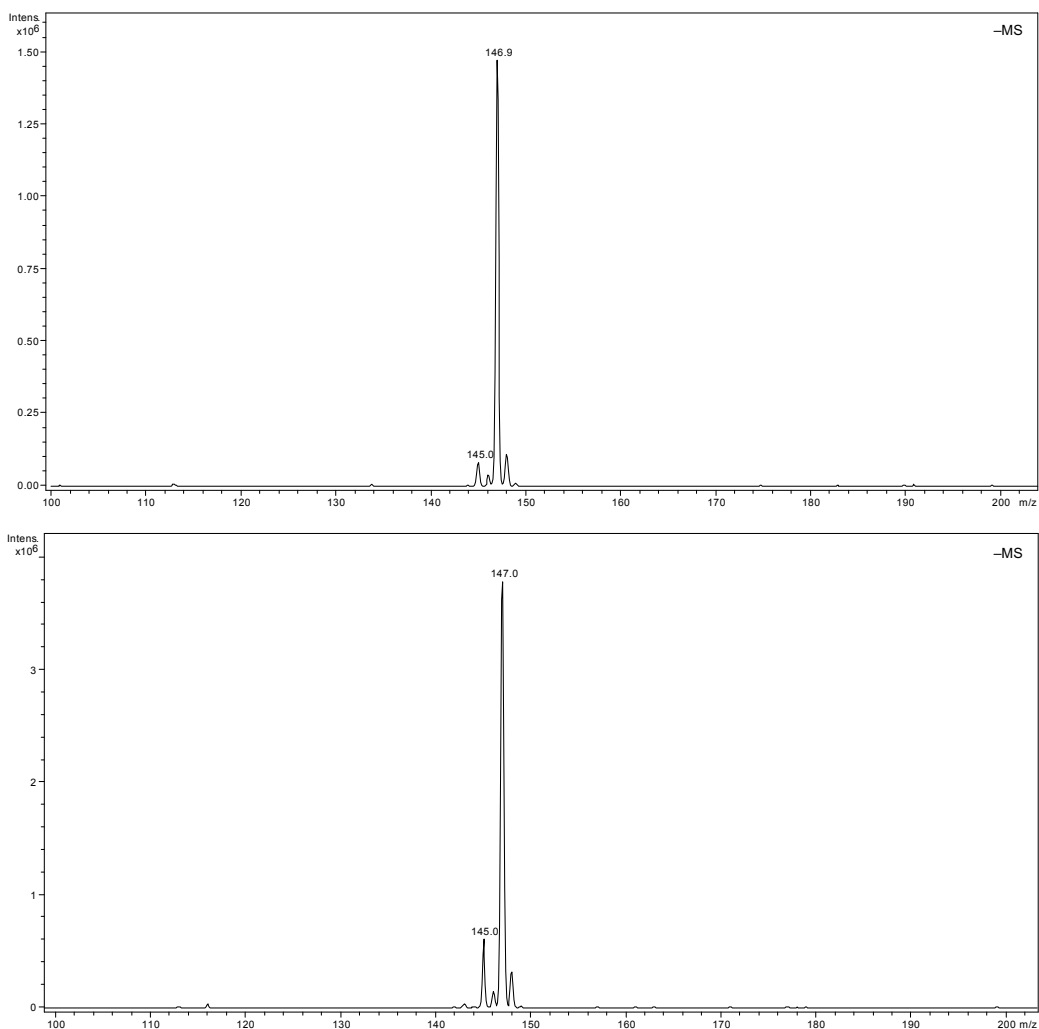


Figure S6. Low-resolution MS spectra of $^{18}\text{O}_2$ -oleic acid (top) before the reaction was initiated and (bottom) after being mixed (as a triethylammonium salt at ca. pH 8.5 and 20 °C) with Val-NCA.

Hydrolysis of Val-NCA

Val-NCA (**1**) hydrolyzes rapidly under our reaction conditions. In aqueous buffer at pH 8.5, in the absence of both the fatty acid and free AA, the observed rate constant for hydrolysis, $k_{\text{obs-h}}$, of **1** was measured as ca. 1 min^{-1} by ^1H NMR spectroscopy (Fig. S7).¹ Our hydrolysis study was carried out as follows: At $15 \text{ }^\circ\text{C}$, a 0.14 M d_8 -THF solution of **1** (0.6 mL) was added quickly into a vigorously stirred NaEPPS solution (300 mM , pH 8.5, 11.4 mL) to give a 0.007 M solution of **1**, an initial concentration for which the NCA polymerization would be negligible.¹ Reaction aliquots (1 mL) taken over a 2 min period were quenched by addition to 0.05 mL TFA in a 4-mL glass vial, and the unreacted **1** was recovered by CDCl_3 (0.6 mL) extraction (vortex for 1 min). For each sample, the organic layer was removed using a gas-tight Hamilton syringe and transferred into a 1-mL vial containing 50 mg Na_2SO_4 . After being dried over Na_2SO_4 , the CDCl_3 solution (ca. 5 mL) was transferred into an NMR tube, and the ^1H NMR spectrum was acquired over 256 scans. The area under the $-\text{CH}_3$ signals of **1** (two apparent doublets, δ ca. 1.1 ppm) was used to determine the relative mol% of **1**, by normalizing to the peak area of the d_8 -THF residual solvent peak (δ ca. 1.7 ppm).

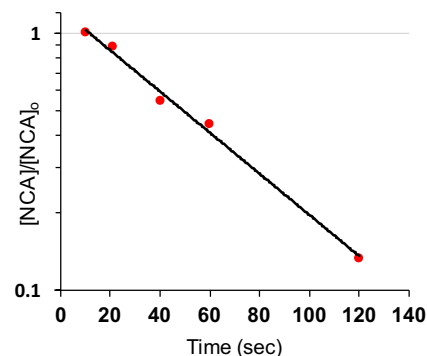


Figure S7. Hydrolysis of Val-NCA (**1**) at $15 \text{ }^\circ\text{C}$ and pH 8.5. Observed rate constant for hydrolysis, $k_{\text{obs-h}}$ of **1** is 0.018 sec^{-1} . ($R^2 = 0.993$).

Standard curves of FRET efficiencies

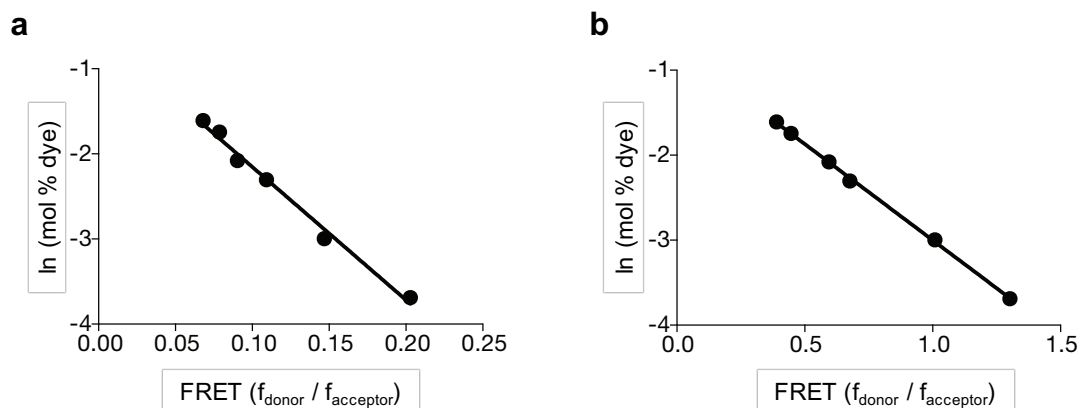
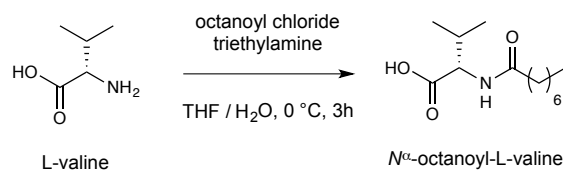


Figure S8. Standard curve of FRET efficiency as a function of mol% FRET dyes in (a) oleate and (b) POPC vesicles ($R^2 > 0.99$).

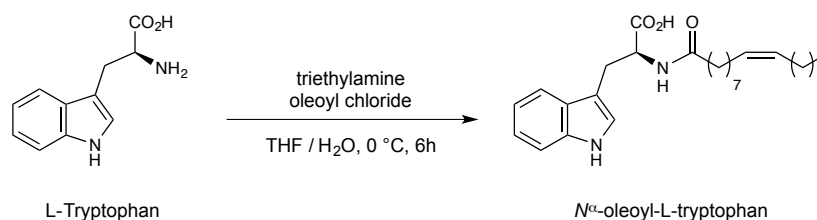
Synthesis of *N*^α-octanoyl-L-valine



To a stirred suspension of L-valine (175 mg, 1.5 mmol, 1.5 equiv) in a mixture of THF (8 mL) and H₂O (1.5 mL) was added octanoyl chloride (0.17 mL, 1.0 mmol, 1 equiv) and triethylamine (0.6 mL, 4.0 mmol, 4 equiv) at 0 °C. The reaction mixture was stirred for 3 h, then concentrated under reduced pressure. The resulting material was dissolved in CH₂Cl₂ (100 mL) and washed with water. The organic phase was concentrated under reduced pressure, and the resulting crude material was purified by combiflash silica gel chromatography (CH₂Cl₂ / MeOH step gradient). The product fraction was concentrated under reduced pressure to afford *N*^α-octanoyl-L-valine (120 mg, 51% yield, based on octanoyl chloride).

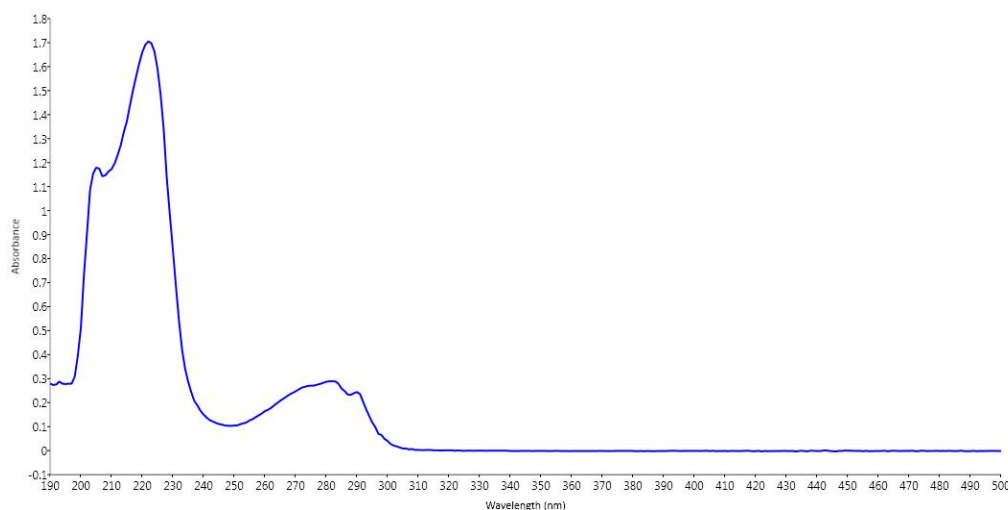
¹H NMR (400 MHz, CDCl₃) δ 7.90 (br s, 1H), 6.15 (d, *J* = 8.5 Hz, 1H), 4.59 (dd, *J* = 8.5, *J* = 4.5 Hz, 1H), 2.27 (t, *J* = 7.0 Hz, 2H), 2.24 (m, 1H), 1.64 (br quin, *J* = 7.0 Hz, 2H), 1.36–1.21 (m, 8H), 0.98 (d, *J* = 7.0 Hz, 3H), 0.95 (d, *J* = 7.0 Hz, 3H), and 0.88 (br t, *J* = 7.0 Hz, 3H); ¹³C NMR (100 MHz, CDCl₃) δ 175.3, 174.1, 57.0, 36.7, 31.7, 31.0, 29.2, 29.0, 25.7, 22.6, 19.0, 17.7, and 14.0; LRMS (ESI): calcd for [C₁₃H₂₅NO₃ – (H⁺)][–] 242.1762, found 242.0.

Synthesis of *N*^α-oleoyl-L-tryptophan



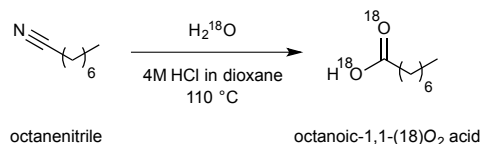
To a stirred suspension of L-tryptophan (200 mg, 1.0 mmol, 1.5 equiv) in a mixture of THF (8.0 mL) and H₂O (1.5 mL) was added oleoyl chloride (0.22 mL, 0.7 mmol, 1 equiv) and triethylamine (0.6 mL, 4.2 mmol, 6 equiv) at 0 °C. The reaction mixture was stirred for 6 h, then concentrated under reduced pressure. The resulting material was dissolved in CH₂Cl₂ (100 mL) and washed with water. The organic phase was concentrated under reduced pressure, and the resulting crude material was purified by combiflash silica gel chromatography (CH₂Cl₂/MeOH

step gradient). The product fraction was concentrated under reduced pressure to afford N^{α} -oleoyl-L-tryptophan (205 mg, 63% yield, based on oleoyl chloride). The $^1\text{H-NMR}$ spectrum of the isolated product acquired in CDCl_3 matched the reported spectrum.² **LRMS** (ESI): calcd for $[\text{C}_{29}\text{H}_{44}\text{N}_2\text{O}_3 - (\text{H}^+)]^-$ 467.3279, found 467.3. The molar extinction coefficient at 280 nm absorbance (ϵ_{280}) of N^{α} -oleoyl-L-tryptophan dissolved in MeOH was measured as $5700 \text{ M}^{-1}\text{cm}^{-1}$ (see the UV spectrum below). In addition, ϵ_{280} of L-tryptophan dissolved in MeOH/ H_2O (1:1) was measured as $5980 \text{ M}^{-1}\text{cm}^{-1}$.



These ϵ_{280} values were used to determine the yields of the N^{α} -acylated and unacylated Trp-containing products as described in the main text. Purification was achieved using combiflash reversed-phase C18 chromatography ($\text{H}_2\text{O}/\text{CH}_3\text{CN}$ or $\text{H}_2\text{O}/\text{MeOH}$ gradient).

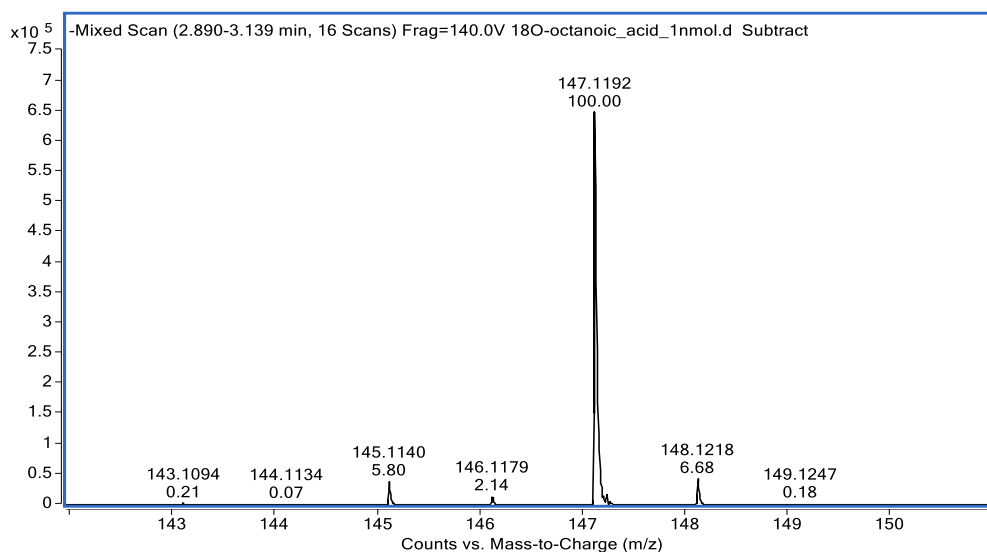
Synthesis of Octanoic-1,1- $^{18}\text{O}_2$ acid



Inside of a glove box, a 15 mL high pressure reaction vessel equipped with a magnetic stir bar, was charged with octanenitrile (550 μL , 3.6 mmol, 1 equiv), H_2^{18}O (97%, 500 μL , 25 mmol, 7 equiv) and 4M HCl in dioxane (3.5 mL). The vessel was sealed tightly with a Teflon cap, then taken out of the glove box and placed in a 110 $^\circ\text{C}$ -oil bath. After being stirred for 4 h, the reaction mixture was allowed to cool to room temperature and filtered through a PTFE filter (pore size 0.22

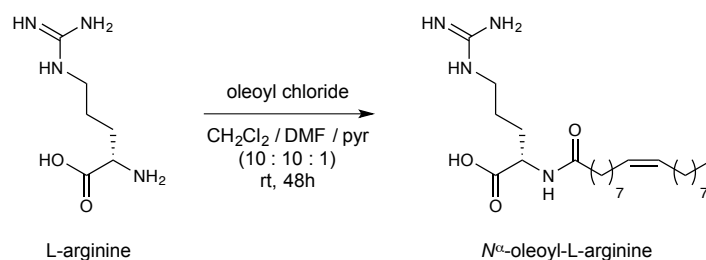
μm). Diethyl ether (8 mL) was added to the filtrate, and the resulting mixture was vortexed in a Falcon tube for ca. 1 min. The organic layer was transferred into another Falcon tube and cooled to 0 °C. Anhydrous triethylamine (1 mL) was added dropwise to remove most of the excess HCl present in the organic liquid as an insoluble triethylammonium chloride salt. The Falcon tube was then vortexed for ca. 1 min and placed in a –30 °C freezer. After 30 min, the tube was centrifuged to pellet the solid material. The supernatant was collected, dried over Na_2SO_4 and concentrated under reduced pressure. The resulting oil was subjected to combiflash silica gel chromatography (with 210–215 nm UV detection) using a hexane/diethyl ether elution mixture (95:5 to 75:25 gradient) to afford octanoic-1,1- $^{18}\text{O}_2$ acid (320 mg, 61% yield, 87% enrichment by HRMS).

^1H NMR (400 MHz, CDCl_3) δ 2.36 (t, $J = 7.5$ Hz, 2H), 1.64 (br q, $J = 7.5$, 2H), 1.37–1.24 (m, 8H), and 0.89 (t, $J = 7.0$ Hz, 3H); ^{13}C NMR (100 MHz, CDCl_3) δ 179.8, 34.2, 31.9, 29.2, 29.1, 24.9, 22.8 (6x CH_2) and 14.3; HRMS (ESI): calcd for $[\text{C}_8\text{H}_{16}^{18}\text{O}_2 - (\text{H}^+)]^-$ 147.1162, found 147.1191.



m/z calc	m/z obs	natural	enriched	enriched norm.
143.1078	143.1094	91.15	0.18	0.21
144.1112	144.1134	8.11	0.06	0.07
145.1132	145.1140	0.69	5.04	5.80
146.1159	146.1179	0.04	1.86	2.14
147.1184	147.1192	0	86.90	100
148.1209	148.1218	0	5.80	6.68
149.1272	149.1247	0	0.16	0.18

Synthesis of N^α -oleoyl-L-arginine

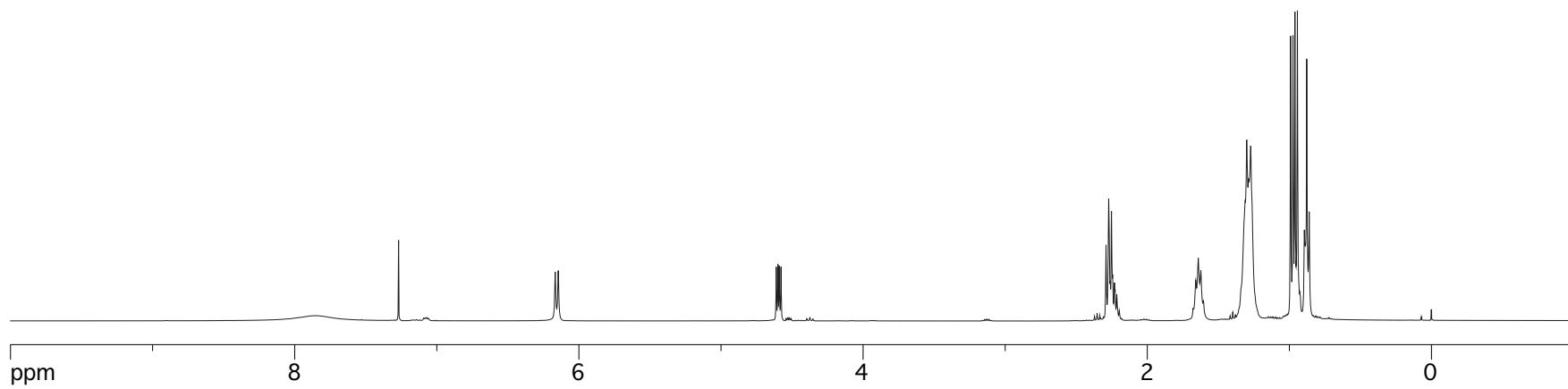
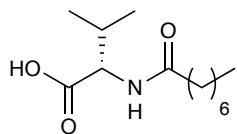


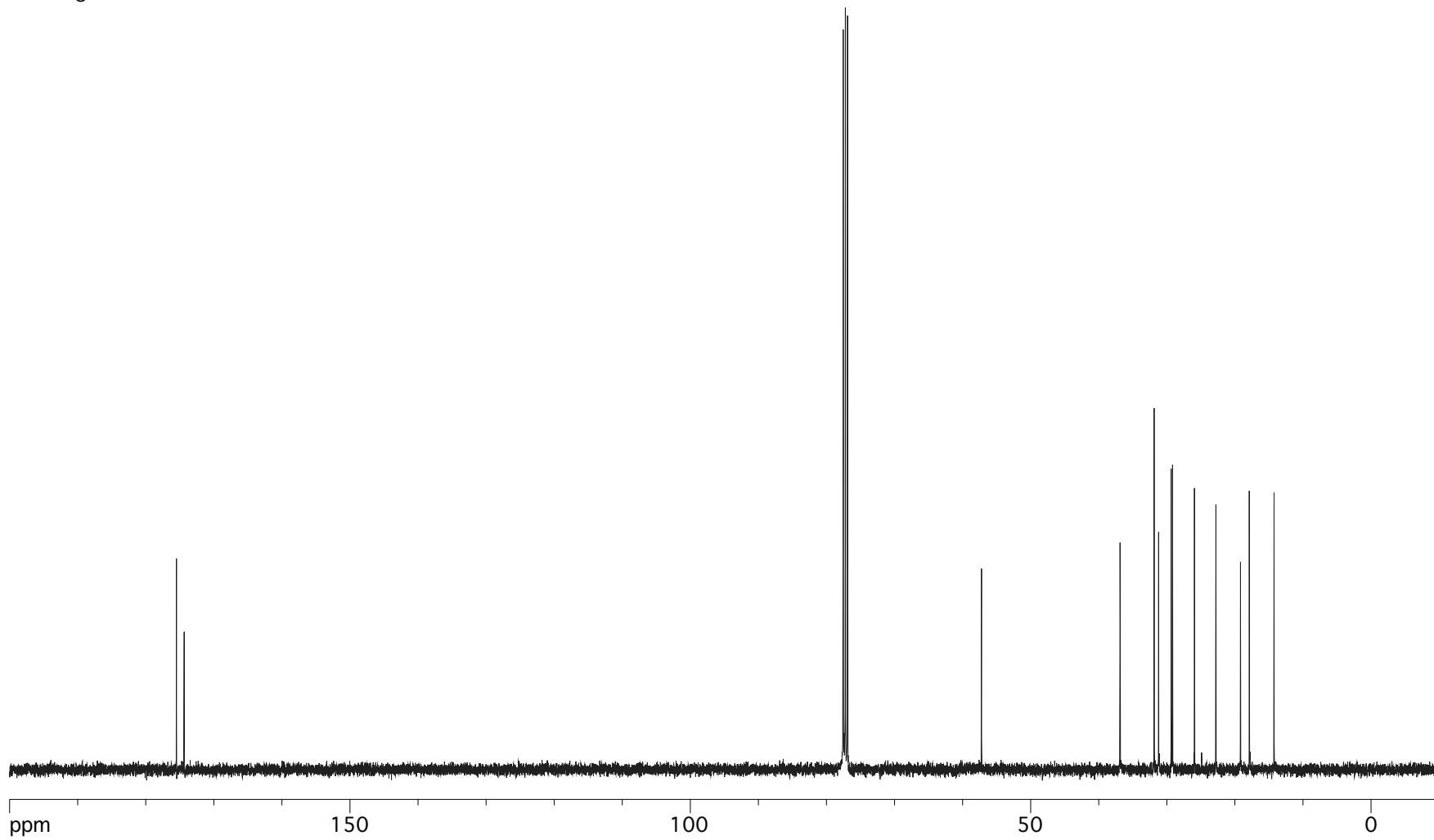
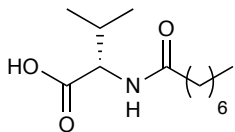
To a stirred suspension of L-arginine (180 mg, 1.0 mmol, 2 equiv) in CH_2Cl_2 (3.5 mL), DMF (3.5 mL) and pyridine (0.4 mL) was added oleoyl chloride (0.17 mL, 0.5 mmol, 1 equiv) at room temperature. The reaction mixture was stirred for 48 h, then diluted with 1N HCl (50 mL) and extracted with CH_2Cl_2 (100 mL). The organic phase was concentrated under reduced pressure, and the resulting crude material was purified by combiflash reversed-phase C18 chromatography (solvent A: water, solvent B: 0.1% TFA/acetonitrile, gradient from 40% to 100% solvent B). The aqueous product fraction was transferred into a Falcon tube and freeze-dried to afford N^α -oleoyl-L-arginine (54 mg, 26% yield).

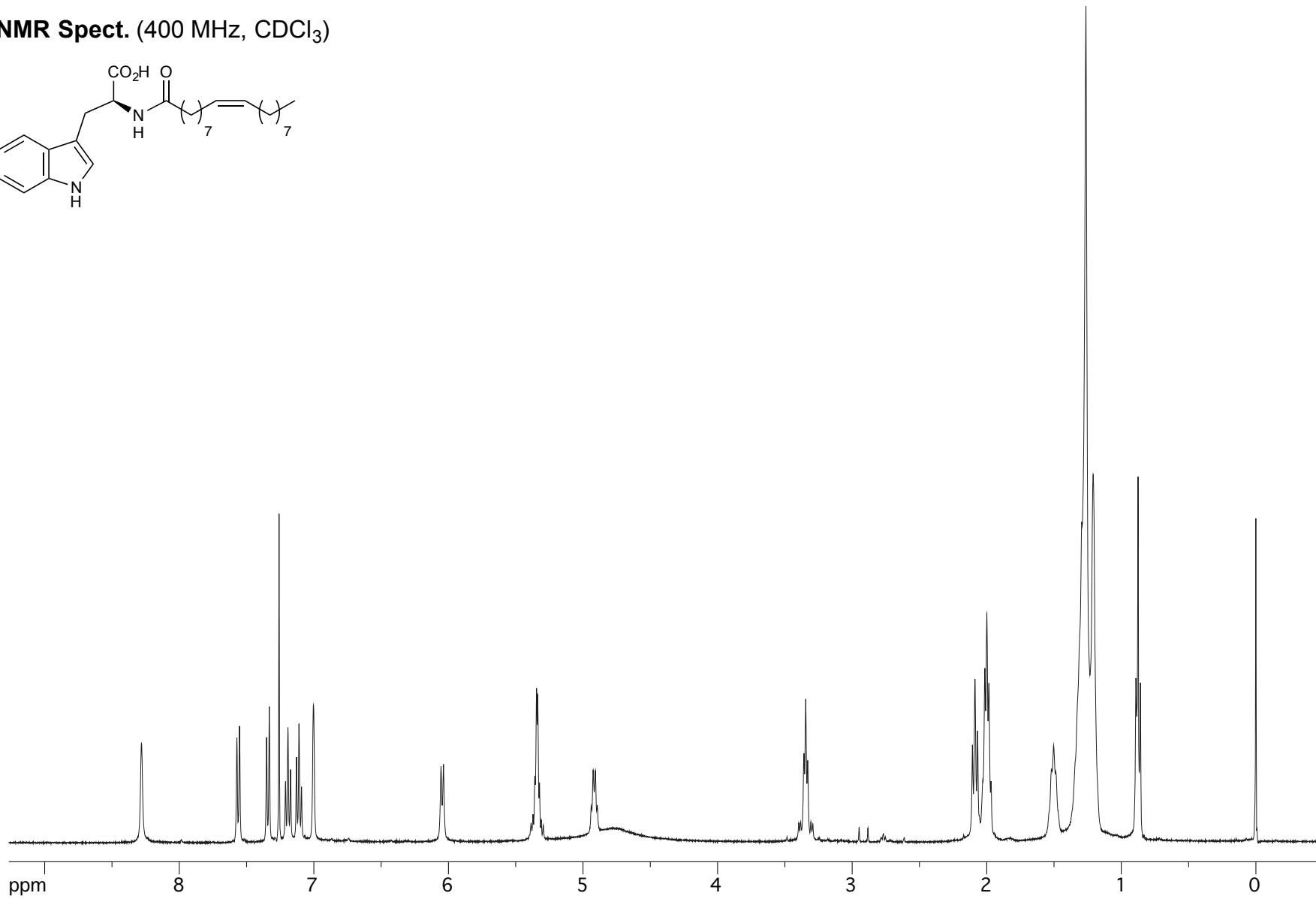
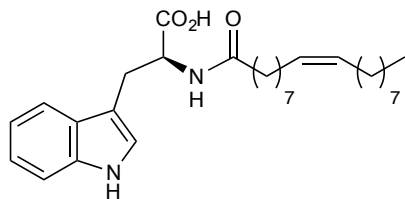
^1H NMR (400 MHz, CD_3OD) δ 5.35 (overlapped m, 2H), 4.43 (dd, $J = 8.5, 5.5$ Hz, 1H), 3.27–3.15 (m, 2H), 2.26 (t, $J = 7.5$ Hz, 2H), 2.08–2.00 (m, 4H), 1.98–1.89 (m, 1H), 1.77–1.70 (m, 1H), 1.70–1.58 (m, 4H), 1.40–1.24 (overlapped m, 20H), and 0.90 (t, $J = 7.0$ Hz, 3H) (exchangeable protons were not observed); **^{13}C NMR** (100 MHz, CD_3OD) δ 176.6, 175.4, 158.8, 131.0, 130.9, 53.2, 42.0, 37.0, 33.2, 31.0⁺, 31.0⁻, 30.8, 30.6, 30.5⁺, 30.5⁻, 30.4, 31.0, 30.0, 28.3⁺, 28.3⁻, 27.1, 26.5, 24.0, and 14.6; **HRMS** (ESI): calcd for $\text{C}_{24}\text{H}_{46}\text{N}_4\text{O}_3 \cdot \text{H}^+$ 439.3643, found 439.3639.

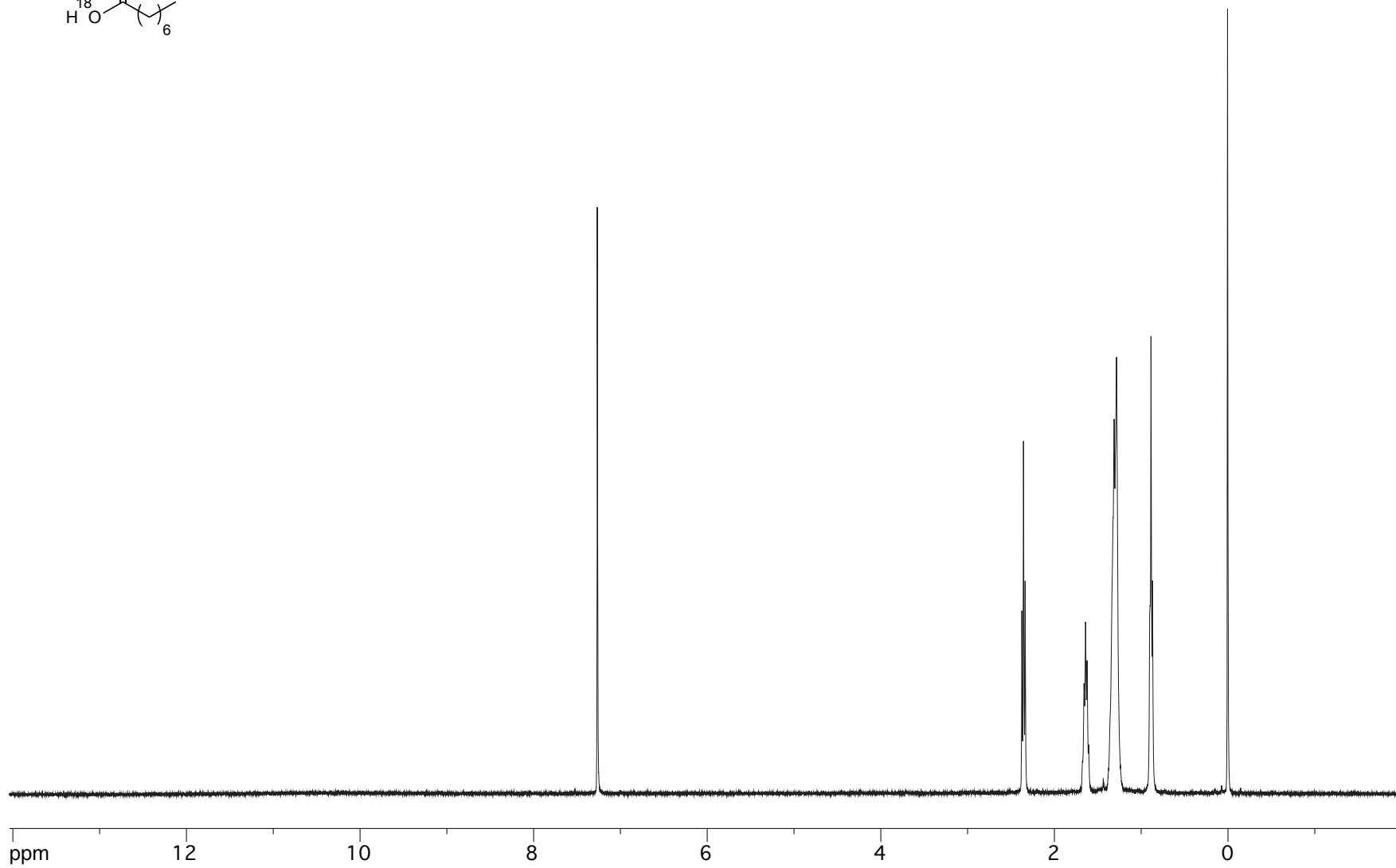
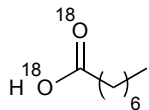
References

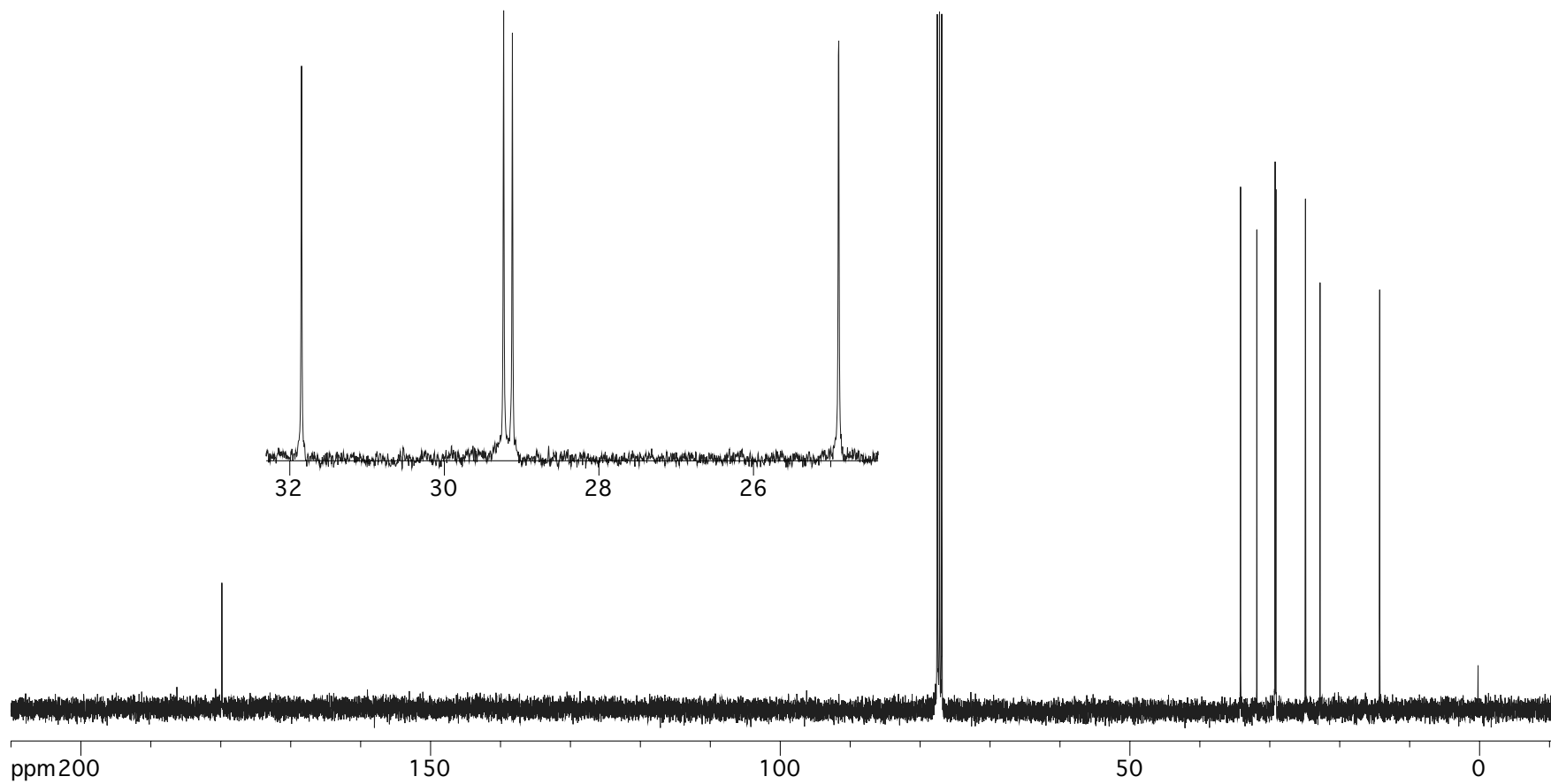
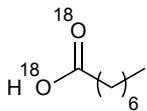
1. At pH 6.5, the $k_{\text{obs-h}}$ of **1** has previously been determined to be 0.07 min^{-1} by capillary electrophoresis: Plasson, R.; Biron, J.-P.; Cottet, H.; Commeyras, A.; Taillades, J. *J. Chromatogr., A* **2002**, *952*, 239–248.
2. Osornio, Y. M.; Uebelhart, P.; Bosshard, S.; Konrad, F.; Siegel, J. S. & Landau, E. M., *J. Org. Chem.* **2012**, *77*, 10583–10595.

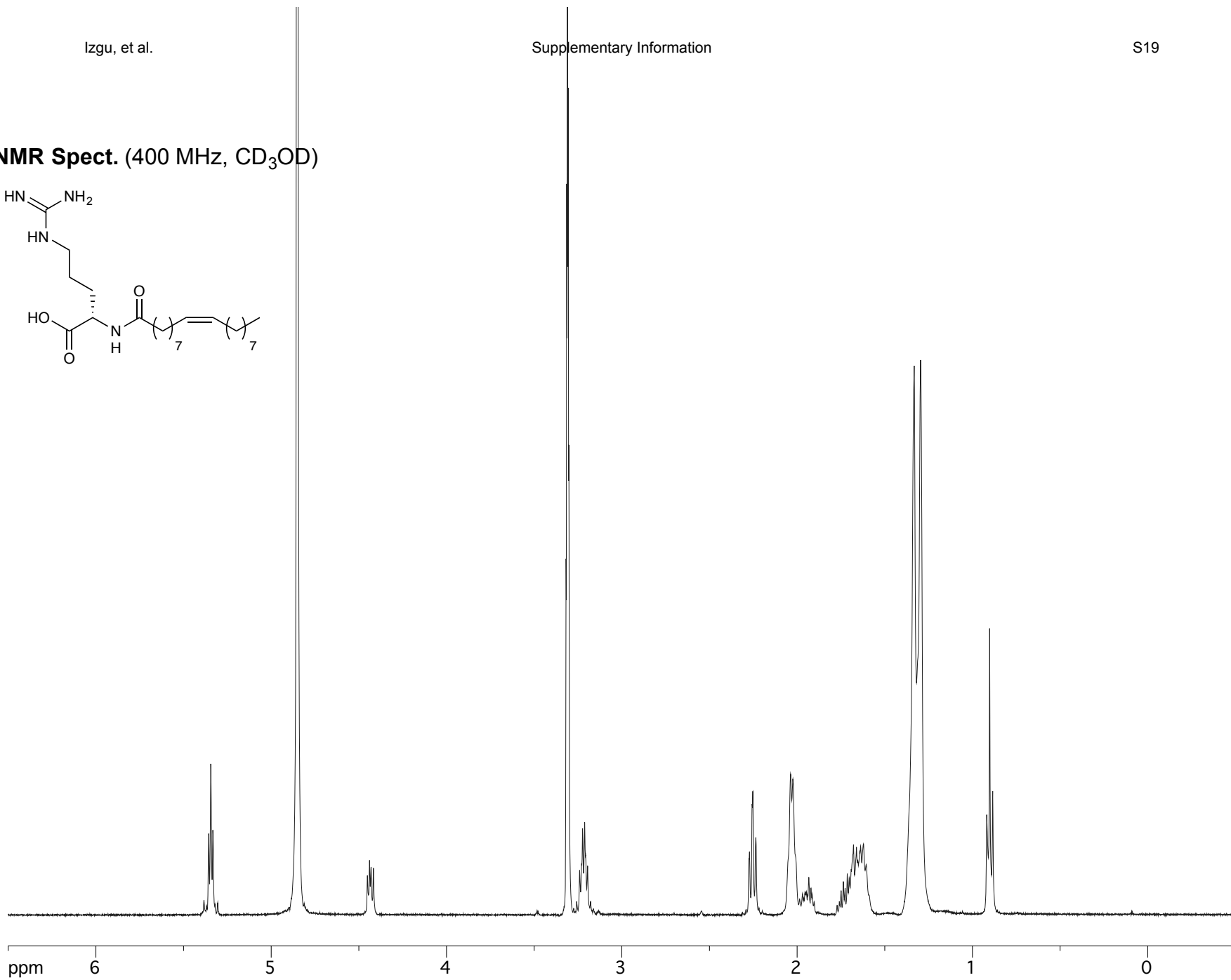
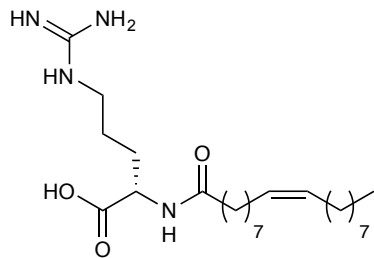
^1H NMR Spect. (400 MHz, CDCl_3)

^{13}C NMR Spect. (100 MHz, CDCl_3)

^1H NMR Spect. (400 MHz, CDCl_3)

^1H NMR Spect. (400 MHz, CDCl_3)

^{13}C NMR Spect. (100 MHz, CDCl_3)

^1H NMR Spect. (400 MHz, CD_3OD)

^{13}C NMR Spect. (100 MHz, CD_3OD)