

The genetic incorporation of thirteen novel non-canonical amino acids†

Cite this: *Chem. Commun.*, 2014, 50, 2673

Received 27th November 2013,
Accepted 9th January 2014

DOI: 10.1039/c3cc49068h

www.rsc.org/chemcomm

Thirteen novel non-canonical amino acids were synthesized and tested for suppression of an amber codon using a mutant pyrrolysyl-tRNA synthetase–tRNA_{CUA}^{Pyl} pair. Suppression was observed with varied efficiencies. One non-canonical amino acid in particular contains an azide that can be applied for site-selective protein labeling.

Site-selective installation of non-canonical amino acids (NCAs) at an amber codon is an efficient approach to synthesize proteins with unique functionalities; applications span from basic studies such as protein cellular localization and protein–protein interaction analysis, to biotechnological applications such as the synthesis of heat stable enzymes and therapeutic protein manufacturing.^{1–5} Two aminoacyl-tRNA synthetase–tRNA_{CUA} pairs have been well adapted for the genetic incorporation of NCAs at amber codons in bacteria. One is the tyrosyl-tRNA synthetase–tRNA_{CUA}^{Tyr} pair that was derived from *Methanocaldococcus jannaschii*.^{6–8} The other is the pyrrolysyl-tRNA synthetase (PylRS)-tRNA_{CUA}^{Pyl} pair that naturally occurs in some methanogenic archaea.^{9–12} Due to its broad-spectrum orthogonality from bacteria to human cells and the fact that it can be easily engineered to target a large variety of NCAs, including natural amino acids with posttranslational modifications, the PylRS-tRNA_{CUA}^{Pyl} pair has captivated researchers for the past several years.^{13–28} One of our major contributions to the NCA research field has been the development of PylRS mutants capable of incorporating a number of phenylalanine derivatives, which are substantially different from the structure of pyrrolysine, the native substrate of PylRS.^{29–31} More specifically, we have recently shown that a rationally designed, N346A/C348A mutant of PylRS (PylRS(N346A/C348A)) is capable of incorporating seven *para*- and twelve *meta*-substituted phenylalanine derivatives at amber codons in coordination with tRNA_{CUA}^{Pyl}.^{30,31} This broad substrate scope obviates the need to undergo the arduous task of discovering a new mutant for

each NCA. Herein we demonstrate that PylRS(N346A/C348A) has an even broader substrate scope than previously reported.

Our previous studies revealed a large active site pocket in PylRS(N346A/C348A).³⁰ Removal of the N346 side chain amide dismisses the steric clash that prevents the binding of the aromatic side chain of phenylalanine and the loss of the C348 thiol yields a cavernous pocket capable of binding the *para*- or *meta*-substituted phenylalanine described above. Interestingly, although phenylalanine derivatives with small *para*-substituents have shown to be ineffective substrates for PylRS(N346A/C348A), their isomers with *meta*-substituents act as highly efficient substrates of PylRS(N346A/C348A) for their genetic incorporation at amber codons.^{30,31} In other words, phenylalanine derivatives with *para*-substituents can only be incorporated when they possess large side chains. Upon further inspection, it appears that a majority of the vacancy in the active site pocket of PylRS(N346A/C348A) exists near the *meta* position of phenylalanine. Encouraged by our preliminary work, we reasoned that PylRS(N346A/C348A) could incorporate phenylalanine derivatives with more sterically demanding side chains.

Our investigation began with the synthesis and genetic incorporation of four different *para*-substituted phenylalanine derivatives (**1–4** in Fig. 1A), each with a unique functionality and steric requirement. Synthesis of these derivatives followed the same strategy presented in one of our previous reports of the N346A/C348A mutant,³⁰ with the exception of NCA **1**, which was synthesized using a different approach (see the ESI†). These four NCAs were then tested for their tolerability by PylRS(N346A/C348A). An *E. coli* BL21(DE3) cell that harbours two plasmids, pEVOL-pyIT-PylRSN346A/C348A and pET-pyIT-sfGFP2TAG, was employed for the investigation. pEVOL-pyIT-PylRSN346A/C348A contains genes coding PylRS(N346A/C348A) and tRNA_{CUA}^{Pyl}; pET-pyIT-sfGFP2TAG carries a tRNA_{CUA}^{Pyl} coding gene and a non-sequence-optimized superfolder green fluorescent protein (sfGFP) gene with an amber mutation in position S2 (sfGFP2TAG). The same cells were used in the initial test of the recognition of *para*-substituted phenylalanine derivatives by PylRS(N346A/C348A).³⁰ Growth in minimal media supplemented with 1 mM IPTG and 0.2% arabinose without NCA afforded a minimal expression level of full-length sfGFP (<0.3 mg L^{−1}). Addition of any of **1–4** at 2 mM to the medium

Department of Chemistry, Texas A&M University, College Station, TX 77843, USA.
E-mail: wliu@chem.tamu.edu

† Electronic supplementary information (ESI) available: Synthesis, protein expression, protein labeling, and mass spectrometry analysis. See DOI: [10.1039/c3cc49068h](https://doi.org/10.1039/c3cc49068h)

‡ Current address: Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, Fujian 350002, China.

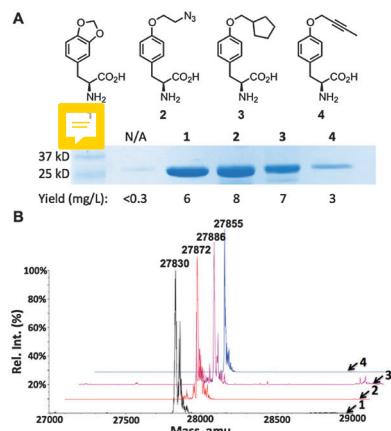


Fig. 1 (A) Structures of **1–4** and their site-specific incorporation into sfGFP at its S2 position. (B) Deconvoluted ESI-MS spectra of sfGFP variants incorporated with **1–4**. Their theoretical values are 27 832 Da for **1**, 27 873 Da for **2**, 27 886 Da for **3**, and 27 856 Da for **4**. Satellite signals are largely due to metal ion adducts (*i.e.* Li, Na, K).

all promoted full-length sfGFP expression (Fig. 1A). The expression levels for **1–3** are comparable to that for *para*-propargyloxyphenylalanine (7.8 mg L⁻¹),³⁰ and the electrospray ionization mass spectrometry analysis of four purified sfGFP variants displayed molecular weights that agreed well with the theoretical values corresponding to full-length proteins with the first methionine (Fig. 1B).

The results obtained for **1–4** demonstrate that PylRS(N346A/C348A) tolerates phenylalanine derivative substrates with rigid and bulky substituents at the *para* position. However, ESI-MS data for compound **3** also show a small side peak corresponding to the incorporation of phenylalanine, a result we have observed previously. The remaining satellite peaks for these compounds correspond to common metal adducts in ESI-MS. Additionally, results obtained for **1** demonstrated that both *meta* and *para* positions can be occupied without detriment to expression levels. These results, coupled with our previous endeavours, led us to wonder if phenylalanine derivatives with long-chain *meta*-substituents could serve as substrates of PylRS(N346A/C348A) for genetic incorporation as well. To investigate this hypothesis, a series of *meta*-alkoxy and *meta*-acylphenylalanines with substituent chain lengths of up to six carbons were synthesized. We chose these specific derivatives because the parent NCAAs *meta*-methoxy-phenylalanine and *meta*-acetylphenylalanine act as efficient substrates for PylRS(N346A/C348A). The synthesis of *meta*-alkoxy-phenylalanines was straightforward, starting with a published route to obtain protected *meta*-tyrosine, at which point the intermediate was subjected to various alkyl halides to afford different derivatives. Acidic deprotection then yielded free amino acids as racemic chloride salts. The synthesis of *meta*-acylphenylalanines was more divergent. Alkyl Grignards were added to a solution of *meta*-tolunitrile, which afforded acylbenzenes upon acidic workup. Radical bromination and then displacement with diethylacetamidomalonate afforded protected *meta*-acyl-phenylalanines that were deprotected in 6 M HCl to obtain free amino acids. More detailed synthetic routes can be found in the ESI.[†]

With the desired NCAs in hand, we thenceforth tested their incorporation efficacies at amber codons using the

PylRS(N346A/C348A)-tRNA^{Val}_{CUA} pair. The *E. coli* cells used for these compounds harboured two plasmids, pEVOL-pylT-PylRSN346A/C348A and pET-pylT-sfGFPS2TAG'. pET-pylT-sfGFPS2TAG' contains a sequence-optimized sfGFP with an amber mutation at its S2 position (sfGFPS2TAG'). In comparison to the sfGFPS2TAG gene in pET-pylT-sfGFPS2TAG, sfGFPS2TAG' has one more alanine residue in front of the amber mutation. Growing this cell in minimal media without NCAAs yielded a minimal expression level of full-length sfGFP. However, all ether NCAs **6–10** (2 mM) in the medium promoted the synthesis of sfGFP with a designated NCAAs incorporated (Fig. 2A). In comparison to phenylalanine derivatives with small *meta*-substituents such as **5**, **6–10** apparently have low incorporation levels. Molecular weights of purified sfGFP variants determined by ESI-MS agreed well with the theoretical values corresponding to a designated NCAAs at the S2 position and the first methionine hydrolysed (Fig. 2B). The removal of the first methionine is due to the insertion of alanine after it. A number of smaller signals can be observed, but they largely correspond to common metal adducts; the expected masses were always the major signal. Compounds **8**, **9**, and **10** have low solubility; when added to the medium at 2 mM, compound **10** was observed to precipitate after 12 h of expression. The low sfGFP expression levels for **8**, **9** and **10** may be partially due to the toxicity of the compounds; indeed, smaller pellet sizes are observed for **8** and **9**. Although the sfGFP expression level for **9** was very low, the purified sfGFP displayed an ESI-MS molecular weight that still matched the theoretical value of sfGFP with **9** incorporated at S2, indicating that a low concentration of **9** was still sufficient to observe incorporation of **9** at the amber mutation site.

Overall, addition of ketone derivatives **12–15** at 2 mM to the medium promoted high sfGFP expression yields, and longer alkyl lengths had less of an impact on protein yields in comparison to the ether series **6–10**, though the sfGFP expression levels for **12–15** are lower than that for **11** (Fig. 3A). This series of NCAs are also readily soluble, with no precipitation observed in the medium after overnight incubation. ESI-MS analysis of the purified sfGFP variants confirmed high incorporation fidelities of **12–15** at the S2 site.

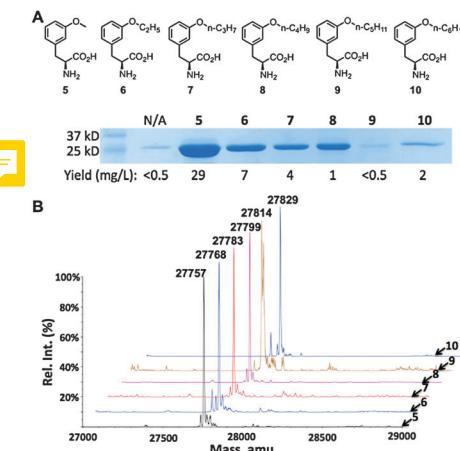


Fig. 2 (A) Structures of **5–10** and their site-specific incorporation into sfGFP at its S2 position. (B) Deconvoluted ESI-MS spectra of sfGFP variants incorporated with **5–10**. Their theoretical values are 27 758 Da for **5**, 27 772 Da for **6**, 27 786 Da for **7**, 27 800 Da for **8**, 27 814 Da for **9**, and 27 828 Da for **10**.

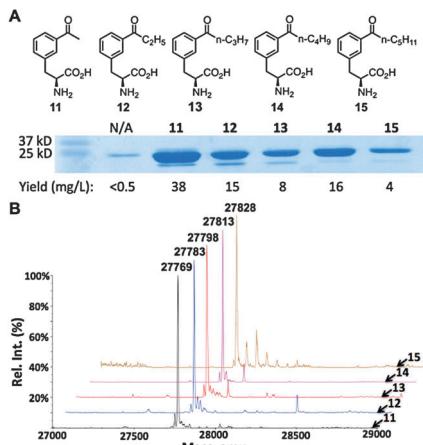


Fig. 3 (A) Structures of **11–15** and their site-specific incorporation into sfGFP at its S2 position. (B) Deconvoluted ESI-MS spectra of sfGFP variants incorporated with **11–15**. Their theoretical values are 27770 Da for **11**, 27784 Da for **12**, 27798 Da for **13**, 27812 Da for **14**, and 27826 Da for **15**. Compounds **13** and **14** show small signals corresponding to an N-terminal methionine on sfGFP. Compound **15** has several small signals attributed to sodium and potassium adducts.

Among all of the novel NCAs that can be taken by PylRS(N346A/C348A), **2** has an active azide functionality for a click reaction with an alkyne³² and **12–15** contain a ketone group that potentially reacts with a hydroxylamine. Both functionalities can be applied for site-selective labeling of proteins incorporated with **2** and **12–15**. Since labeling of sfGFP incorporated with **11** with a hydroxylamine dye was demonstrated previously,³¹ we chose to demonstrate the selective labeling of **2** using a diarylcyclooctyne dye **D1** in this study (Fig. 4). **D1** contains a strained alkyne that undergoes a spontaneous reaction with an azide.³³ Incubating sfGFP incorporated with **2** with **D1** overnight led to an intensely fluorescently labeled protein; however, the same reaction with sfGFP incorporated with **3** did not yield any fluorescently labeled final product. This result indicates that genetically incorporated **2** can be applied to site-specifically introduce biophysical and biochemical probes to proteins for a large variety of studies.

In summary, we have shown that thirteen novel NCAs were genetically incorporated into protein at the amber codon in *E. coli* using the PylRS(N346A/C348A)-tRNA^{Pyl}_{CUA} pair. This result, coupled with our previous findings, shows a surprisingly broad substrate scope for PylRS(N346A/C348A). Investigations are underway to determine aspects of the active site pocket of PylRS(N346A/C348A) that lead to this broad substrate spectrum. The current study has great implications in understanding amino acid structure tolerance of the protein translation system. The expanded genetically encoded NCAAs

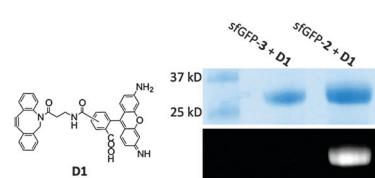


Fig. 4 Labeling of sfGFP incorporated with **2** (sfGFP-**2**) and sfGFP incorporated with **3** (sfGFP-**3**) with dye **D1**. The top panel shows the Coomassie blue stained SDS-PAGE gel and the bottom panel shows the fluorescent image of the same gel under UV irradiation before the gel was stained with Coomassie blue.

pool can also be applied to generate phage and *E. coli* displayed peptide libraries with expanded chemical moieties for drug discovery, a direction we are actively pursuing at the current stage.

This work was supported in part by the Welch Foundation (grant A-1715), the National Science Foundation (grant CHEM-1148684), and the National Institute of Health (grant 1R01CA161158).

Notes and references

- W. R. Liu, Y. S. Wang and W. Wan, *Mol. Biosyst.*, 2011, **7**, 38–47.
- A. Gautier, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2011, **133**, 2124–2127.
- M. Zhang, S. Lin, X. Song, J. Liu, Y. Fu, X. Ge, X. Fu, Z. Chang and P. R. Chen, *Nat. Chem. Biol.*, 2011, **7**, 671–677.
- Y. Tang, G. Ghirlanda, N. Vaidehi, J. Kua, D. T. Mainz, I. W. Goddard, W. F. DeGrado and D. A. Tirrell, *Biochemistry*, 2001, **40**, 2790–2796.
- C. C. Liu and P. G. Schultz, *Annu. Rev. Biochem.*, 2010, **79**, 413–444.
- L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498–500.
- J. Xie and P. G. Schultz, *Nat. Rev. Mol. Cell Biol.*, 2006, **7**, 775–782.
- J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang and P. G. Schultz, *J. Am. Chem. Soc.*, 2002, **124**, 9026–9027.
- G. Srinivasan, C. M. James and J. A. Krzycki, *Science*, 2002, **296**, 1459–1462.
- S. K. Blight, R. C. Larue, A. Mahapatra, D. G. Longstaff, E. Chang, G. Zhao, P. T. Kang, K. B. Green-Church, M. K. Chan and J. A. Krzycki, *Nature*, 2004, **431**, 333–335.
- H. Neumann, S. Y. Peak-Chew and J. W. Chin, *Nat. Chem. Biol.*, 2008, **4**, 232–234.
- W. Wan, Y. Huang, Z. Wang, W. K. Russell, P. J. Pai, D. H. Russell and W. R. Liu, *Angew. Chem., Int. Ed.*, 2010, **49**, 3211–3214.
- S. Greiss and J. W. Chin, *J. Am. Chem. Soc.*, 2011, **133**, 14196–14199.
- S. M. Hancock, R. Uprety, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2010, **132**, 14819–14824.
- T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto and S. Yokoyama, *Biochem. Biophys. Res. Commun.*, 2008, **371**, 818–822.
- T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto and S. Yokoyama, *Chem. Biol.*, 2008, **15**, 1187–1197.
- A. R. Parrish, X. She, Z. Xiang, I. Coin, Z. Shen, S. P. Briggs, A. Dillin and L. Wang, *ACS Chem. Biol.*, 2012, **7**, 1292–1302.
- P. R. Chen, D. Groff, J. Guo, W. Ou, S. Cellitti, B. H. Geierstanger and P. G. Schultz, *Angew. Chem., Int. Ed.*, 2009, **48**, 4052–4055.
- C. J. Chou, R. Uprety, L. Davis, J. W. Chin and A. Deiters, *Chem. Sci.*, 2011, **2**, 480–483.
- Y. S. Wang, B. Wu, Z. Wang, Y. Huang, W. Wan, W. K. Russell, P. J. Pai, Y. N. Moe, D. H. Russell and W. R. Liu, *Mol. Biosyst.*, 2010, **6**, 1557–1560.
- Y. J. Lee, B. Wu, J. E. Raymond, Y. Zeng, X. Fang, K. L. Wooley and W. R. Liu, *ACS Chem. Biol.*, 2013, **8**, 1664–1670.
- T. Fekner, X. Li, M. M. Lee and M. K. Chan, *Angew. Chem., Int. Ed.*, 2009, **48**, 1633–1635.
- X. Li, T. Fekner, J. J. Ottesen and M. K. Chan, *Angew. Chem., Int. Ed.*, 2009, **48**, 9184–9187.
- T. Umehara, J. Kim, S. Lee, L. T. Guo, D. Soll and H. S. Park, *FEBS Lett.*, 2012, **586**, 729–733.
- C. R. Polycarpo, S. Herring, A. Berube, J. L. Wood, D. Soll and A. Ambrogelly, *FEBS Lett.*, 2006, **580**, 6695–6700.
- T. Plass, S. Milles, C. Koehler, C. Schultz and E. A. Lemke, *Angew. Chem., Int. Ed.*, 2011, **50**, 3878–3881.
- T. Plass, S. Milles, C. Koehler, J. Szymanski, R. Mueller, M. Wiessler, C. Schultz and E. A. Lemke, *Angew. Chem., Int. Ed.*, 2012, **51**, 4166–4170.
- D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2009, **131**, 8720–8721.
- Y. S. Wang, W. K. Russell, Z. Wang, W. Wan, L. E. Dodd, P. J. Pai, D. H. Russell and W. R. Liu, *Mol. Biosyst.*, 2011, **7**, 714–717.
- Y. S. Wang, X. Fang, A. L. Wallace, B. Wu and W. R. Liu, *J. Am. Chem. Soc.*, 2012, **134**, 2950–2953.
- Y. S. Wang, X. Fang, H. Y. Chen, B. Wu, Z. U. Wang, C. Hilti and W. R. Liu, *ACS Chem. Biol.*, 2013, **8**, 405–415.
- H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004–2021.
- J. C. Jewett, E. M. Sletten and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2010, **132**, 3688–3690.

Supplementary Material

The genetic incorporation of thirteen novel non-canonical amino acids

Alfred Tuley, Yane-Shih Wang, Xinqiang Fang, Yadagiri Kurra, Yohannes Rezenom,
and Wenshe Liu

Department of Chemistry, Texas A&M University, College Station, TX 77843
E-mail: wliu@chem.tamu.edu

Table of Contents

1. Superfolder Green Fluorescent Protein Expression	S3
2. Protein Labeling	S3
3. ESI-MS Analysis of Intact Proteins	S3
4. Organic Synthesis	S3
5. sfGFP Protein Sequence.....	S12
6. NMR Data.....	S13

1. Superfolder Green Fluorescent Protein Expression

The constructs used to incorporate compounds **1-4** in this study and corresponding protein purification and characterization are identical to previous reports.^{1,2} For compounds **5-15**, BL21(DE3) *E. coli* cells containing pEVOL-PylT-pylRS-PylRSN346A/C348A (Cm^r) and pET-PylT-sfGFP-S2TAG' (Amp^r) vectors were grown in 500 mL of LB media with ampicillin (100 $\mu\text{g}/\text{mL}$) and chloramphenicol (34 $\mu\text{g}/\text{mL}$) until the OD_{600} reached 1.0-1.3, at which point the cells were pelleted at 4,000 r.p.m. for 20 min, then washed and resuspended in 30 mL H_2O . The resuspended cells were added as 5 mL aliquots to 45 mL of minimal media (33.7 mM Na_2HPO_4 , 22 mM KH_2PO_4 , 8.6 mM NaCl, 9.4 mM NH_4Cl , 1 mM MgSO_4 , 0.3 mM CaCl_2 , 1% glycerol) supplemented with 2 mM non-canonical amino acid (NAA), 1 mM IPTG, and 0.2% arabinose, at which point protein expression was allowed to occur for 12 h. The cells were then pelleted at 4,000 r.p.m. for 20 min, resuspended in lysis buffer (50 mM NaH_2PO_4 , 300 mM NaCl, pH 8) and lysed via sonication. The crude lysate was then centrifuged at 10,000 r.p.m. for 1 hour, and the supernatant was treated with imidazole to a final concentration of 10 mM. Next, the supernatant was subsequently incubated with Ni^{2+} -NTA resin for 1 hour at 4 °C. The resin was washed with lysis buffer containing 10 mM imidazole (3x column volume) and 20 mM imidazole (3x column volume), then eluted with elution buffer (50 mM NaH_2PO_4 , 300 mM NaCl, 500 mM imidazole, pH 8). The protein was dialyzed against 10 mM Tris buffer, and if necessary, the proteins were concentrated using Amicon Ultracel-10k centrifugal filter units. Purity of the proteins was confirmed via 15% SDS-PAGE and ESI-MS analysis. Yields were determined using a commercially available BCA protein assay kit (Thermo Scientific).

2. sfGFP-2 Protein Labeling

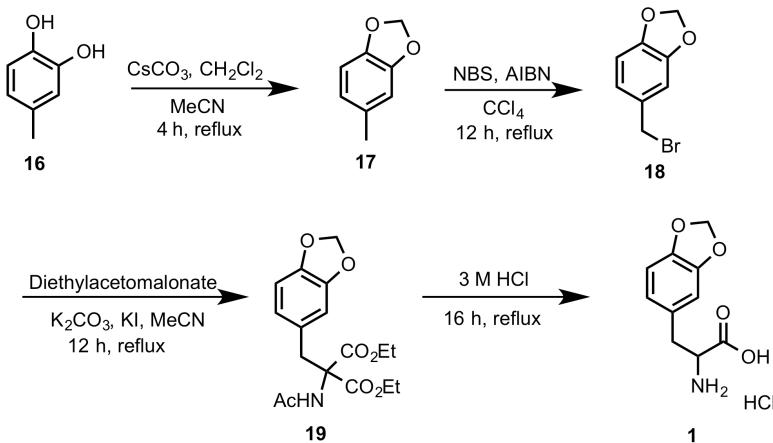
To 16 μl of 1X Phosphate buffered saline (PBS) was added 2 μl sfGFP-**2** (3 mg/mL, 1 μM) and 2 μl Dibenzylcyclooctyne (DBCO) dye (25 mM, DMSO stock solution)³ and was allowed to react at room temperature overnight. The protein was then precipitated with 180 μl of methanol and left at -20 °C for 1 h. Next, the precipitated protein was centrifuged for 5 min, 14,000 r.p.m. and washed twice with 100% methanol. Finally, the residue was resuspended in 20 μl H_2O and subjected to SDS-PAGE analysis. The control experiment was identical except 4 μl sfGFP-**3** (5 mg/mL, 0.9 μM) was added to 14 μl PBS, followed by 2 μl DBCO dye.

3. ESI-MS Analysis of Intact Proteins

Nanoelectrospray ionization in positive mode was performed using an Applied Biosystems QSTAR Pulsar (Concord, ON, Canada) equipped with a nanoelectrospray ion source. Solution was flowed at 700 nL/min through a 50 μm ID fused-silica capillary that was tapered at the tip. Electrospray needle voltage was held at 2100 V.

4. Organic Synthesis

Reactions were carried out using oven-dried glassware and under an atmosphere of argon, where appropriate. Reagents were purchased and used without further purification. NMR spectra were obtained with Inova 300 and Mercury 300 MHz instruments.



4.1 2-amino-3-(benzo[d][1,3]dioxol-5-yl)propanoic acid hydrochloride (1)

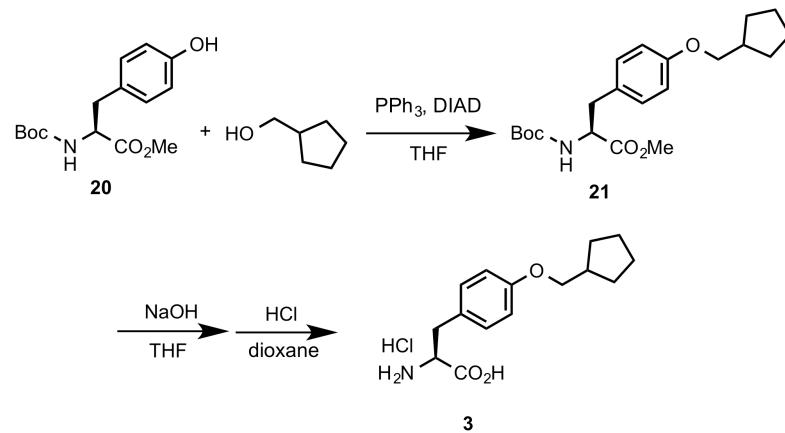
To a solution of catechol **16** (5 g, 8.06 mmol) in MeCN (25 mL) was added Cs_2CO_3 (3.93g 12.09 mmol), followed by CH_2Cl_2 (5.21 mL, 8.06 mmol). The resulting mixture was refluxed for 4 hours. After cooling to room temperature, the reaction mixture was concentrated and applied to column chromatography, yielding the acetal **17** in 40% yield (2.12 g). ^1H NMR (300 MHz, CDCl_3) δ 6.71(m, 3H), 5.93(s, 2H), 2.27(s, 3H).

The compound **17** (2.0 g, 14.7 mmol) was dissolved in CCl_4 (15 mL), followed by addition of NBS (3.12 g, 17.62 mmol) and AIBN (0.337 g, 2.05 mmol). The mixture was heated to reflux for 12 h under the protection of argon. After cooling, the precipitate was removed via filtration and the filtrate was concentrated and dried under vacuum to afford the known compound **18**,⁵ which was used directly in the next step without further purification.

To a solution of compound **18** (3.5 g, 16.27 mmol) in anhydrous acetonitrile (15 mL) was added diethyl 2-acetamidomalonate (3.88 g, 17.88 mmol), K_2CO_3 (4.49 g, 32.53 mmol) and KI (1.0 g, 15.38 mmol). The resulting mixture was heated to reflux for 12 h under the protection of argon, then cooled to room temperature. The solid was filtered, the solvent was removed under reduced pressure, and the residue was purified via column chromatograph with hexanes/ethyl acetate (3:1 v/v) as eluent to give the pure product **19** (60% yield, 3.42 g). ^1H NMR (300 MHz, CDCl_3) δ 6.69 (d, $J = 8.2\text{Hz}$, 1 H), 6.57(d, $J = 7.9\text{ Hz}$, 1 H), 6.47 (s, 1H), 5.92 (s, 2H), 4.30-4.22 (m, 4H), 3.56 (s, 2H), 2.04 (s, 3H), 1.32-1.23 (m, 6H).

A suspension of compound **19** (3.0 g, 8.54 mmol) in 3 M HCl (91.16 mL, 32 eq.) was heated to reflux for 16 h before cooling to room temperature. Water was evaporated under reduced pressure and the solid was collected. The solid was washed with Et_2O (10 mL \times 3) and then dried under vacuum to afford compound **1** as a solid (68% yield, 1.42 g). ^1H NMR (300 MHz, CD_3OD) δ 6.78-6.66 (m, 3H), 5.85 (s, 2H), 4.14 (t, $J = 6.6\text{ Hz}$, 1 H), 3.09 (ABq, $J = 15.3, 6.6\text{ Hz}$, 2 H). ^{13}C NMR (75 MHz, CD_3OD) δ 169.8, 148.1, 147.3, 127.6, 122.6, 109.1, 108.2, 101.1, 53.9, 39.5.

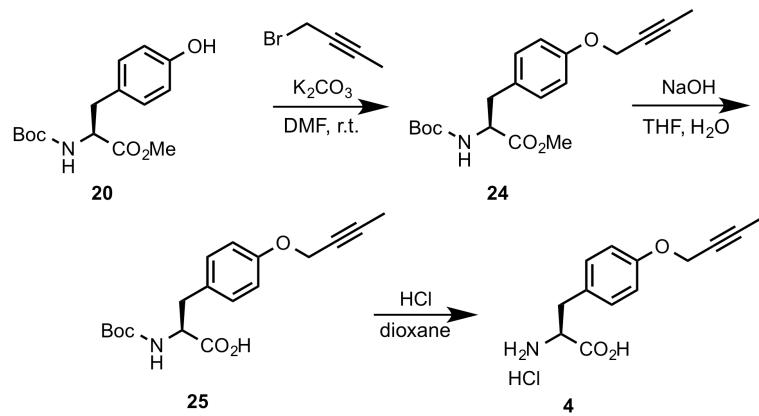
4.2 Alkylation of Tyrosine



4.2.1 (*S*)-2-amino-3-(4-(cyclopentylmethoxy)phenyl)propanoic acid hydrochloride (3)

N-Boc-*o*-methyl-L-tyrosine (**20**) (1.3 g, 4.4 mmol), cyclopentylmethanol (0.57 mL, 5.28 mmol), and triphenylphosphine (1.73 g, 6.6 mmol) were added to a round bottom flask, then the atmosphere evacuated and replaced with argon. THF (10 mL) was then added and the reaction mixture was cooled to 0 °C, at which point DEAD (1.3 mL, 6.6 mmol) was added dropwise and stirred overnight. Upon completion, the reaction mixture was concentrated and purified via column chromatography (gradient elution, 25% to 50% ethyl acetate/hexanes) to give the product **21** in 78% yield (1.3 g).

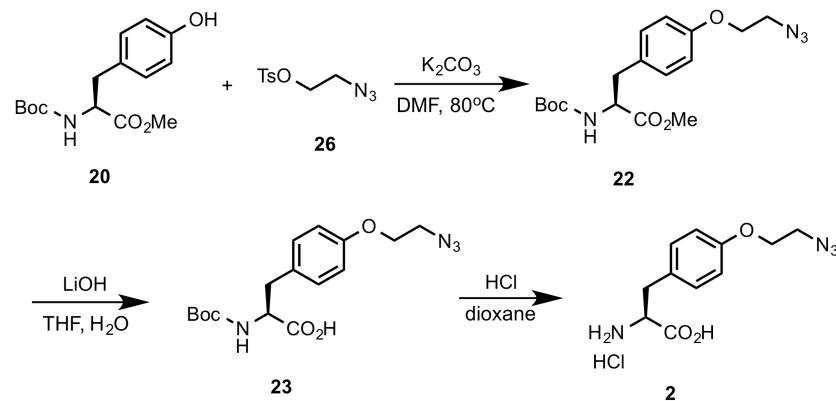
This compound was deprotected using the previously described procedure,¹ treatment with 5 mL 1 M NaOH/THF for two hours, followed by Boc deprotection in 5 mL 4 M HCl/dioxane to give the free amino **3** acid in 97% yield (898 mg), two steps. ¹H NMR (300 MHz, CD₃OD) δ 7.16 (d, *J* = 8.4 Hz, 2 H), 6.87 (d, *J* = 8.7, 6.9 Hz, 2 H), 4.15 (t, *J* = 5.4 Hz, 1 H), 3.80 (d, *J* = 6.9 Hz, 2 H), 3.14 (dd, *J* = 7.8, 5.4 Hz, 2 H), 2.31 (m, 1 H), 1.82 (m, 2 H), 1.61 (m, 4 H), 1.36 (m, 2 H). ¹³C NMR (75 MHz, CD₃OD) δ 169.9, 158.9, 130.0, 125.6, 114.6, 71.8, 53.8, 38.9, 35.0, 28.9, 24.9.



4.2.2 (*S*)-2-amino-3-(4-(but-2-yn-1-yloxy)phenyl)propanoic acid hydrochloride (4)

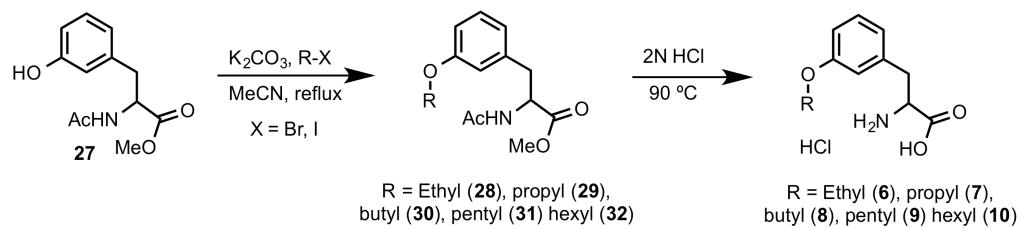
N-Boc-*o*-methyl-L-tyrosine (1 g, 3.39 mmol) was treated with 3-bromo-2-propyne (0.36 mL, 4.06 mmol) and potassium carbonate (1.4 g, 10.2 mmol) in DMF and left to

react overnight at room temperature. The reaction mixture was then diluted with ethyl acetate and washed with 3 M HCl. The organic layers were then dried, concentrated, and purified via flash chromatography (gradient elution, 25% to 50% ethyl acetate/hexanes) to afford the protected azide **24** in 30% yield (370 mg). Deprotection using the protocol listed above gave the title compound **4** in 44% yield (274 mg). ¹H NMR (300 MHz, CD₃OD) δ 7.10 (d, *J* = 8.7 Hz, 2 H), 6.83 (d, *J* = 8.7 Hz, 2 H), 4.53 (m, 2 H), 4.08 (q, *J* = 7.5 Hz, 1 H), 3.02 (dd, *J* = 14.7, 7.5 Hz, 2 H), 1.68 (s, 3 H). ¹³C NMR (75 MHz, CD₃OD) δ 168.1, 155.8, 128.4, 124.6, 113.3, 81.0, 72.0, 53.9, 52.0, 33.2.



4.2.3 (*S*)-2-amino-3-(4-(2-azidoethoxy)phenyl)propanoic acid hydrochloride (2)

N-Boc-*o*-methyl-L-tyrosine **20** (1.8 g, 6.09 mmol) was treated with compound **26** (1.8 g, 7.31 mmol) and potassium carbonate (3.3 g, 24.4 mmol) in DMF (10 mL) and the reaction was heated to 80 °C. When the reaction was complete based on TLC, the reaction mixture was worked up as usual and purified via flash chromatography (gradient elution, 25% to 50% ethyl acetate/hexanes) to yield compound **22**. Deprotection was achieved using 20 mL 1 M LiOH, followed by treatment with 5 mL 4 M HCl/Dioxane to afford free amino acid **2** (50%, two steps, 873 mg). ¹H NMR (300 MHz, D₂O) δ 7.16 (d, *J* = 8.4 Hz, 2 H), 6.88 (d, *J* = 8.4 Hz, 2 H), 4.11 (m, 3 H), 3.50 (m, 2 H), 3.11 (m, *J* = 7.2, 6.9, 5.4 Hz, 2 H). ¹³C NMR (75 MHz, D₂O) δ 169.8, 158.1, 130.2, 126.4, 114.7, 67.0, 53.8, 49.8, 35.0.



4.3 Representative Procedure for *m*-Tyrosine Alkylation

4.3.1 Methyl 2-acetamido-3-(3-ethoxyphenyl)propanoate (28)

To a solution of *m*-tyrosine⁶ **27** (1.5 g, 6.32 mmol) in DMF (12.64 mL) was added K₂CO₃ (2.84 g, 20.55 mmol), followed by ethyl iodide (0.76 mL, 9.48 mmol). The

reaction was stirred at room temperature for 18 h, then quenched with 16 mL H₂O and 79 mL ethyl acetate. The organic layer was extracted three times with 40 mL H₂O, then dried with sodium sulfate and concentrated. Column chromatography (gradient elution, 50 to 75% ethyl acetate/hexanes) afforded the compound as a yellow, crystalline solid (72% yield, 1.2 g). ¹H NMR (300 MHz, CDCl₃) δ 7.17 (t, *J* = 7.8 Hz, 1H), 6.76 (dd, *J* = 8.1, 5.7 Hz, 1H), 6.66 (m, 2H), 6.04 (d, *J* = 7.5 Hz, 1H), 4.85 (dt, *J* = 7.8, 5.7 Hz, 1H), 3.97 (q, *J* = 6.9 Hz, 2H), 3.71 (s, 3H), 3.07 (m, *J* = 8.1, 5.7 Hz, 2H), 1.97 (s, 3H), 1.38 (t, *J* = 6.9 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 172.2, 169.7, 159.2, 137.3, 129.6, 121.5, 115.6, 113.2, 63.4, 53.1, 52.4, 37.9, 23.3, 14.9.

4.3.2 Methyl 2-acetamido-3-(3-propoxymphenyl)propanoate (29)

Synthesized according to the general procedure with *m*-tyrosine (1.5 g, 6.32 mmol), DMF (12.64 mL), K₂CO₃ (2.84g, 20.55 mmol), and n-propyl bromide (0.86 mL, 9.48 mmol) to yield a yellow, crystalline solid, (68% yield, 1.2 g). ¹H NMR (300 MHz, CDCl₃) δ 7.18 (t, *J* = 8.1 Hz, 1H), 6.79 (m, *J* = 5.7 Hz, 1H), 6.66 (m, 2H), 5.88 (d, *J* = 7.8 Hz, 1H), 4.87 (dt, *J* = 7.8, 5.7 Hz, 1H), 3.88 (t, *J* = 6.6 Hz, 2H), 3.74 (s, 3H), 3.09 (m, *J* = 8.1, 5.7 Hz, 2H), 1.99 (s, 3H), 1.79 (sx, *J* = 7.2 Hz, 2H), 1.03 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 172.1, 169.7, 159.3, 137.3, 129.6, 121.4, 115.6, 113.2, 69.5, 53.1, 52.4, 37.9, 23.2, 22.6, 10.6.

4.3.3 Methyl 2-acetamido-3-(3-butoxyphenyl)propanoate (30)

Synthesized according to the general procedure with *m*-tyrosine (1.0 g, 4.21 mmol), DMF (8.42 mL), K₂CO₃ (1.89g, 13.69 mmol), and n-butyl iodide (0.72 mL, 6.32 mmol) to yield a yellow, crystalline solid (65% yield, 0.8 g). ¹H NMR (300 MHz, CDCl₃) δ 7.19 (t, *J* = 7.8 Hz, 1H), 6.77 (d, *J* = 7.2 Hz, 1H), 6.64 (m, 2H), 5.87 (d, *J* = 6.9 Hz, 1H), 4.87 (dt, *J* = 7.8, 5.4 Hz, 1H), 3.92 (t, *J* = 6.6 Hz, 2H), 3.74 (s, 3H), 3.09 (dd, *J* = 8.1, 5.7 Hz, 2H), 1.99 (s, 3H), 1.76 (p, *J* = 6.9, 6.6, 6.3 Hz, 2H), 1.46 (m, 2H), 0.97 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 172.2, 169.7, 159.4, 137.3, 129.6, 121.4, 115.6, 113.2, 67.7, 53.1, 52.4, 37.9, 31.4, 23.2, 19.3, 13.9.

4.3.4 Methyl 2-acetamido-3-(3-(pentyloxy)phenyl)propanoate (31)

Synthesized according to the general procedure with *m*-tyrosine (1.5 g, 6.32 mmol), DMF (12.64 mL), K₂CO₃ (2.84g, 20.55 mmol), and n-pentyl bromide (1.18 mL, 9.48 mmol) to yield a yellow, crystalline solid (72% yield, 1.4 g). ¹H NMR (300 MHz, CDCl₃) δ 7.19 (t, *J* = 8.1 Hz, 1H), 6.78 (m, 1H), 6.63 (m, 2H), 5.87 (d, *J* = 6.9 Hz, 1H), 4.82 (dt, *J* = 7.8, 5.7 Hz, 1H), 3.91 (t, *J* = 6.6 Hz, 2H), 3.74 (s, 3H), 3.10 (dq, *J* = 8.1, 6.0 Hz, 2H), 1.99 (s, 3H), 1.77 (p, *J* = 7.2, 6.9, 6.6 Hz, 2H), 1.39 (m, 4H), 0.93 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 172.1, 169.7, 159.2, 137.3, 129.5, 121.2, 115.5, 113.0, 67.8, 53.1, 52.3, 37.8, 28.9, 28.2, 23.0, 22.4, 14.0.

4.4.5 Methyl 2-acetamido-3-(3-(hexyloxy)phenyl)propanoate (32)

Synthesized according to the general procedure with *m*-tyrosine (1.0 g, 4.21 mmol), DMF (8.42 mL), K₂CO₃ (1.89g, 13.69 mmol), and 1-iodohexane (0.93 mL, 6.32 mmol) to yield a yellow, crystalline solid (44% yield, 0.6 g). ¹H NMR (300 MHz, CDCl₃) δ 7.18 (t, *J* = 8.1 Hz, 1H), 6.77 (dd, *J* = 8.1, 5.7 Hz, 1H), 6.63 (m, 2H), 5.88 (d, *J* = 7.2 Hz, 1H), 4.87 (dt, *J* = 7.8, 5.7 Hz, 1H), 3.91 (t, *J* = 6.6 Hz, 1H), 3.74 (s, 3H), 3.09 (m, *J*

= 8.1, 5.7 Hz, 2 H), 1.99 (s, 3H), 1.77 (p, J = 8.1, 6.6 Hz, 2 H), 1.44 (m, 4 H), 1.33 (m, 4H), 0.91 (t, J = 7.2 Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 172.2, 169.7, 159.4, 137.3, 129.6, 121.4, 115.6, 113.2, 68.0, 53.1, 52.4, 37.9, 31.7, 29.3, 25.8, 23.3, 22.7, 14.1.

4.5 Representative Procedure for *meta*-alkoxy Phenylalanine Deprotection:

4.5.1 2-amino-3-(3-ethoxyphenyl)propanoic acid hydrochloride (6)

A round bottomed flask charged with compound **28** (1.1 g, 4.15 mmol) was treated with 14.3 mL 2 M HCl, and the resulting suspension was heated to 100 °C, which was refluxed overnight until complete consumption of starting material as observed on TLC. The reaction mixture was then cooled to room temperature and concentrated *in vacuo*. If necessary, the compound was resubjected to the reaction conditions due to the persistence of methyl ester and/or N-acyl amide signals in ^1H and ^{13}C NMR. The title compound was obtained as a white solid in 45% yield (463 mg). ^1H NMR (300 MHz, D_2O) δ 7.35 (t, J = 7.8 Hz, 1 H), 6.93 (m, 3 H), 4.31 (t, J = 7.5 Hz 1 H), 4.11 (q, J = 6.9 Hz, 2 H), 3.24 (dq, J = 14.7, 5.7 Hz, 2 H), 1.37 (t, J = 6.9 Hz, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 171.4, 158.3, 135.6, 130.4, 122.1, 115.5, 114.1, 64.2, 54.0, 35.5, 13.8.

4.5.2 2-amino-3-(3-propoxypheNyl)propanoic acid hydrochloride (7)

Prepared according to the general procedure using compound **29** (1.0 g, 3.58 mmol) and 12.34 mL 2 M HCl. Obtained as a white solid in 43% yield (402 mg). ^1H NMR (300 MHz, D_2O) δ 7.34 (t, J = 7.5 Hz, 1 H), 6.90 (m, 3 H), 4.29 (t, J = 5.4 Hz, 1 H), 3.97 (t, J = 6.6 Hz), 3.21 (dd, J = 14.7, 5.7 Hz, 2 H), 1.73 (sx, J = 7.5, 7.2, 6.9, 6.6 Hz, 2 H), 0.96 (t, J = 7.5, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 171.3, 158.5, 135.5, 130.3, 122.0, 115.5, 114.1, 70.1, 53.9, 35.4, 21.7, 9.5.

4.5.3 2-amino-3-(3-butoxyphenyl)propanoic acid hydrochloride (8)

Prepared according to the general procedure using compound **30** (0.8 g, 2.73 mmol) in 9.4 mL 2 M HCl to obtain the title compound in 30% yield (222 mg). ^1H NMR (300 MHz, DMSO) δ 8.53 (s, 2 H), 7.21 (t, J = 7.8 Hz, 1 H), 6.89 (s, 1 H), 6.82 (d, J = 7.8 Hz, 2 H), 4.14 (t, J = 5.4 Hz, 1 H), 3.94 (t, J = 6.6 Hz, 2 H), 3.12 (d, J = 6.0 Hz, 2 H), 1.69 (p, J = 7.5, 6.9, 6.6 Hz, 2 H), 1.45 (sx, J = 7.5, 7.2 Hz, 2 H), 0.93 (t, J = 7.2 Hz, 3 H). ^{13}C NMR (75 MHz, DMSO) δ 170.2, 158.7, 136.4, 129.5, 121.5, 115.6, 113.1, 66.9, 53.1, 35.5, 30.8, 18.8, 13.7.

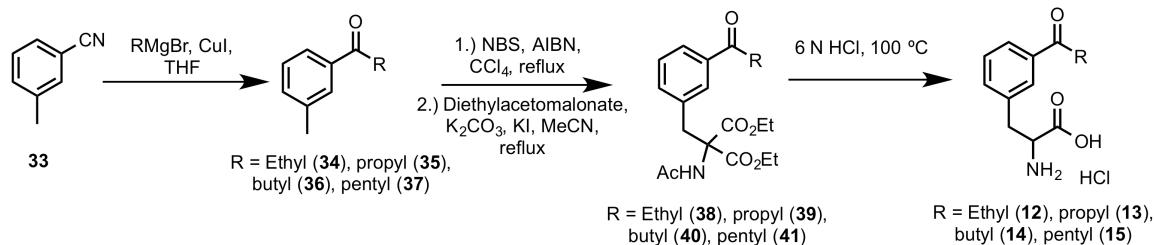
4.5.4 2-amino-3-(3-(pentyloxy)phenyl)propanoic acid hydrochloride (9)

Prepared according to the general procedure using compound **31** (1.0 g, 3.25 mmol) in 11.2 mL 2 M HCl to obtain the title compound in 61% yield (570 mg). ^1H NMR (300 MHz, DMSO) δ 8.49 (s, 2 H), 7.21 (t, J = 7.5 Hz, 1 H), 6.88 (s, 1 H), 6.82 (d, J = 8.1 Hz, 2 H), 4.14 (t, J = 6.0 Hz, 1 H), 3.94 (t, J = 6.6 Hz, 1 H), 3.11 (d, J = 6.0 Hz, 2 H), 1.71 (p, J = 6.9, 6.6, 6.3 Hz, 2 H), 1.36 (m, 4 H), 0.89 (t, J = 6.6 Hz, 3 H). ^{13}C NMR (75 MHz, DMSO) δ 170.2, 158.7, 136.4, 129.5, 121.6, 115.6, 113.1, 67.2, 53.0, 35.6, 28.4, 27.7, 21.9, 13.9.

4.5.5 2-amino-3-(3-(hexyloxy)phenyl)propanoic acid hydrochloride (10)

Prepared according to the general procedure using compound **32** (0.6 g, 1.87 mmol) in 6.44 mL 2 M HCl to yield the compound as a white solid in 65% yield (367

mg). ^1H NMR (300 MHz, DMSO) δ 8.48 (s, 3 H), 7.21 (t, J = 7.8 Hz, 1 H), 6.88 (s, 1 H), 6.82 (d, J = 7.8 Hz, 2 H), 4.13 (m, J = 5.1, 4.5 Hz, 1 H), 3.93 (t, J = 6.6 Hz, 2 H), 3.10 (d, J = 6.3 Hz, 2 H), 1.69 (p, J = 7.8, 6.6 Hz, 2 H), 1.41 (m, 2 H), 1.29 (m, 4 H), 0.88 (t, J = 6.6 Hz, 3 H). ^{13}C NMR (75 MHz, DMSO) δ 170.2, 158.7, 136.3, 129.5, 121.5, 115.6, 113.1, 67.2, 53.0, 35.6, 31.0, 28.7, 25.2, 22.1, 13.9.



4.6 Representative Procedure for Grignard Addition to *m*-Tolunitrile

4.6.1 1-(*m*-tolyl)propan-1-one (34)

Ethyl magnesium bromide (1 M/THF, 85.36 mL, 85.36 mmol) was added dropwise to a solution of *m*-tolunitrile (10.25 mL, 85.36 mmol) and copper (I) iodide (40.64 mg, 0.213 mmol) in anhydrous THF (170.72 mL) under argon. After stirring for 22 h, the reaction was quenched with approx. 5 mL 1 M HCl at 0 °C and stirred for 4 h, allowing the reaction to warm to room temperature. The resulting layers were separated and the organic layer was dried with MgSO_4 , filtered, and concentrated. The crude, yellow oil was purified via silica gel chromatography (gradient, 0 to 10% EtOAc/Hex) to afford the known compound⁷ **34** in 94% yield (9.4 g) as a yellow oil. ^1H NMR (300 MHz, CDCl_3) δ 7.76 (m, 1H), 7.38 (m, 3H), 2.99 (q, J = 7.2 Hz, 2H), 2.41 (s, 3H), 1.22 (t, J = 7.2 Hz, 3H).

4.6.2 1-(*m*-tolyl)butan-1-one (35)

Prepared according to the general procedure with Propylmagnesium bromide (2 M/THF, 42.68 mL, 85.36 mmol), *m*-tolunitrile (10.25 mL, 85.36 mmol), and CuI (40.64 mg, 0.213 mmol) to yield the known compound⁸ in 96 % yield (13.24 g). ^1H NMR (300 MHz, CDCl_3) δ 7.77 (s, 2H), 7.36 (s, 2H), 2.94 (t, J = 7.2 Hz, 2H), 2.41 (s 3H), 1.76 (sx, J = 7.2 Hz, 2H), 1.00 (t, J = 7.2 Hz, 3H).

4.6.3 1-(*m*-tolyl)pentan-1-one (36)

Prepared according to the general procedure with butylmagnesium chloride (2 M/THF, 42.68 mL, 85.36 mmol), *m*-tolunitrile (10.25 mL, 85.36 mmol), and CuI (40.64 mg, 0.213 mmol) to yield the known compound⁹ in 76% yield (11.43 g). ^1H NMR (300 MHz, CDCl_3) δ 7.75 (m, 2H), 7.34 (m, 2H), 2.95 (t, J = 7.2 Hz, 2H), 2.41 (s, 3H), 1.71 (p, J = 7.5 Hz, 2H), 1.41 (sx, J = 7.2 Hz, 2H), 0.95 (t, J = 7.5 Hz, 3H).

4.6.4 1-(*m*-tolyl)hexan-1-one (37)

Prepared according to the general procedure with Pentylmagnesium bromide (1 M/THF, 42.68 mL, 85.36 mmol), *m*-tolunitrile (10.25 mL, 85.36 mmol), and CuI (40.64 mg, 0.213 mmol) to yield the known compound¹⁰ in 94% yield (15.3 g). ^1H NMR (300 MHz, CDCl_3) δ 7.74 (m, 1H), 7.46 (s, 1H), 7.37 (m, 2H), 2.94 (t, J = 6.9 Hz, 2H), 2.41 (s, 3H), 1.73 (p, J = 7.2 Hz, 2H), 1.35 (m, 4H), 0.91 (t, J = 6.6 Hz, 3H).

4.7 Representative Procedure for the Synthesis of Protected Ketones

4.7.1 Diethyl 2-acetamido-2-(3-propionylbenzyl)malonate (38)

To a solution of *m*-Keto toluene **34** (8.0 g, 53.98 mmol) in CCl_4 (134.95 mL) was added NBS (10.57 g, 59.38 mmol) and AIBN (2.67 g, 16.19 mmol), and the resulting suspension was refluxed overnight. Upon completion, the reaction was filtered and concentrated, and the resulting crude material was subjected to the next step without further purification.

A round-bottom flask was charged with brominated **34** (4.24 g, 19.09 mmol), diethyl 2-acetamidomalonate (3.73 g, 17.18 mmol), K_2CO_3 (5.28 g, 38.18 mmol), and KI (3.17 g, 19.09 mmol), followed by 119.3 mL of MeCN, and the resulting suspension was heated to reflux and stirred overnight. Upon completion, the reaction was cooled to room temperature, filtered with celite, and concentrated. Purification via silica gel chromatography (gradient elution, 0 to 30% EtOAc/Hex) afforded **38** as a yellow solid in 58% yield (3.6 g). ^1H NMR (300 MHz, CDCl_3) δ 7.83 (d, $J = 7.8$ Hz, 1 H), 7.64 (s, 1 H), 7.36 (t, $J = 7.8, 7.5$ Hz, 1 H), 7.20 (d, $J = 8.4$ Hz, 1 H), 6.52 (s, 1 H), 4.28 (q, $J = 7.2, 6.9$ Hz, 4 H), 2.96 (q, $J = 7.2$ Hz, 2 H), 2.05 (s, 3 H), 1.31 (t, $J = 7.2$ Hz, 6 H), 1.21 (t, $J = 7.2$ Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 200.5, 169.3, 167.4, 137.0, 136.0, 134.5, 129.3, 128.6, 127.1, 67.2, 63.0, 37.7, 31.9, 23.1, 14.1, 8.3.

4.7.2 Diethyl 2-acetamido-2-(3-butyrylbenzyl)malonate (39)

Synthesized according to the general procedure with ketone **35** (8.0 g, 49.31 mmol), NBS (9.65 g, 54.24 mmol), AIBN (2.43 g, 14.79 mmol), and 123.28 mL CCl_4 . The crude product was subjected to the next step without further purification.

Malonate synthesis was performed according to the general procedure using brominated **35** (5.73 g, 23.75 mmol), diethyl 2-acetamidomalonate (4.64 g, 21.37 mmol), K_2CO_3 (6.56 g, 47.5 mmol), KI (3.94 g, 23.75 mmol), and 148 mL MeCN. 37% yield, 2.9 g. ^1H NMR (300 MHz, CDCl_3) δ 7.82 (d, $J = 7.8$ Hz, 1 H), 7.63 (s, 1 H), 7.34 (t, $J = 7.5$ Hz, 1 H), 7.20 (d, $J = 7.2$ Hz, 1 H), 6.52 (s, 1 H), 4.28 (q, $J = 7.2, 6.9$ Hz, 4 H), 2.89 (t, $J = 7.2$ Hz, 2 H), 2.05 (s, 3 H), 1.75 (sx, $J = 7.5, 7.2$ Hz, 2 H), 1.31 (t, $J = 7.2$ Hz, 6 H), 0.99 (t, $J = 7.5$ Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 200.1, 169.3, 167.4, 137.2, 136.0, 134.5, 129.4, 128.6, 127.1, 67.2, 63.0, 40.6, 37.7, 23.1, 17.8, 14.1, 14.0.

4.7.3 Diethyl 2-acetamido-2-(3-pentanoylbenzyl)malonate (40)

Synthesized according to the general procedure with ketone **36** (9.5 g, 53.9 mmol), NBS (10.55 g, 59.29 mmol), AIBN (2.66 g, 16.17 mmol), and 134.75 mL CCl_4 . The crude product was subjected to the next step without further purification.

Malonate synthesis was performed according to the general procedure using brominated **36** (6.26 g, 24.53 mmol), diethyl 2-acetamidomalonate (4.79 g, 22.08 mmol), K_2CO_3 (6.78 g, 49.07 mmol), KI (4.07 g, 24.53 mmol), and 153.3 mL MeCN. 51% yield as a yellow solid (4.4 g). ^1H NMR (300 MHz, CDCl_3) δ 7.82 (d, $J = 8.1$ Hz, 1 H), 7.62 (s, 1 H), 7.36 (t, $J = 7.8$ Hz, 1 H), 6.52 (s, 1 H), 4.28 (q, $J = 7.2$ Hz, 4 H), 3.7 (s, 2 H), 2.92

(t, $J = 7.5$ Hz, 2 H), 2.05 (s, 3 H), 1.70 (p, $J = 7.8, 7.5, 7.2$ Hz, 2 H), 1.38 (p, $J = 7.8, 7.5, 7.2$ Hz, 2 H), 1.31 (t, $J = 6.9, 7.2$ Hz, 6 H), 0.95 (t, $J = 7.2$ Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 200.2, 169.3, 167.4, 137.2, 135.9, 134.5, 129.3, 128.6, 127.1, 67.2, 62.9, 38.4, 37.7, 26.4, 23.1, 22.5, 14.0.

4.7.4 Diethyl 2-acetamido-2-(3-hexanoylbenzyl)malonate (41)

Synthesized according to the general procedure with ketone **37** (10.0 g, 52.55 mmol), NBS (10.3 g, 57.81 mmol), AIBN (2.59 g, 15.77 mmol), and 131.38 mL CCl_4 . The crude product was subjected to the next step without further purification.

Malonate synthesis was performed according to the general procedure using brominated **37** (5.73 g, 23.75 mmol), diethyl 2-acetamidomalonate (4.52 g, 20.79 mmol), K_2CO_3 (6.39 g, 46.21 mmol), KI (3.84 g, 23.11 mmol), and 144.44 mL MeCN. 37% yield as a yellow solid (3.1 g). ^1H NMR (300 MHz, CDCl_3) δ 7.82 (d, $J = 6.6$ Hz, 1 H), 7.62 (s, 1 H), 7.36 (t, $J = 7.5$ Hz, 1 H), 7.20 (d, $J = 7.5$ Hz, 1 H), 6.52 (s, 1 H), 4.28 (q, $J = 7.2$ Hz, 4 H), 3.71 (s, 2 H), 2.91 (t, $J = 7.5$ Hz, 2 H), 2.05 (s, 3 H), 1.72 (m, 2 H), 1.34 (m, 4 H), 1.31 (m, 6 H), 0.91 (t, $J = 6.6$ Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 200.2, 169.4, 167.4, 137.2, 135.9, 134.5, 129.3, 128.6, 127.1, 67.2, 62.9, 38.7, 37.7, 31.6, 24.0, 23.1, 22.6, 14.1.

4.8 Representative Procedure for Malonate Deprotection

4.8.1 2-amino-3-(3-propionylphenyl)propanoic acid hydrochloride (12)

A suspension of **38** (1.0 g, 2.75 mmol) in 6 M HCl was refluxed overnight, until disappearance of protecting groups was verified via ^1H NMR. The resulting solution was concentrated *in vacuo* to yield **12** as a yellow solid (36% yield, 256 mg). ^1H NMR (300 MHz, D_2O) δ 7.93 (d, $J = 6.9$ Hz, 1 H), 7.87 (s, 1 H), 7.55 (m, 2 H), 4.16 (t, $J = 7.2$ Hz, 1 H), 3.30 (dq, $J = 14.1, 5.7$ Hz, 2 H), 3.09 (q, $J = 7.2$ Hz, 2 H), 1.14 (t, $J = 6.9$ Hz, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 206.4, 172.4, 136.8, 135.1, 134.4, 129.3, 128.7, 127.5, 54.8, 35.7, 31.9, 7.5.

4.8.2 2-amino-3-(3-butyrylphenyl)propanoic acid hydrochloride (13)

Synthesized according to the representative procedure using **39** (1.0 g, 2.65 mmol) in 9.14 mL 6 M HCl. 80% yield, 576 mg. ^1H NMR (300 MHz, D_2O) δ 7.95 (d, $J = 7.2$ Hz, 1 H), 7.85 (s, 1 H), 7.53 (m, 2 H), 4.28 (t, $J = 7.2$ Hz, 1 H), 3.31 (dq, $J = 14.7, 5.7$ Hz, 2 H), 3.02 (t, $J = 7.2$ Hz, 2 H), 1.66 (q, $J = 7.2$ Hz, 2 H), 0.92 (t, $J = 7.2$ Hz, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 206.1, 171.5, 136.9, 134.8, 134.5, 129.4, 128.8, 127.8, 54.2, 40.4, 35.4, 17.6, 12.8.

4.8.3 2-amino-3-(3-pentanoylphenyl)propanoic acid hydrochloride (14)

Synthesized according to the representative procedure using **40** (1.08 g, 2.75 mmol) in 9.48 mL 6 M HCl. 29% yield, 200 mg. ^1H NMR 7.96 (dt, $J = 7.2$ Hz, 1 H), 7.87 (s, 1 H), 7.58 (m, 2 H), 4.31 (t, $J = 7.2$ Hz, 1 H), 3.35 (dq, $J = 14.7, 5.7$ Hz, 2 H), 3.07 (t, $J = 7.2$ Hz, 2 H), 1.65 (p, $J = 7.2$ Hz, 2 H), 1.34 (sx, $J = 7.5, 7.2$ Hz, 2 H), 0.91 (t, $J = 7.2$ Hz, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 206.2, 171.5, 136.9, 134.7, 134.5, 129.4, 128.8, 127.8, 54.2, 38.3, 35.4, 26.2, 21.6, 13.0.

4.8.4 2-amino-3-(3-hexanoylphenyl)propanoic acid hydrochloride (15)

Synthesized according to the representative procedure using **41** (1.03 g, 2.55 mmol) in 8.79 mL 6 M HCl. 34% yield, 229 mg. ¹H NMR (300 MHz, D₂O) δ 7.95 (d, *J* = 7.2 Hz, 1 H), 7.87 (s, 1 H), 7.57 (m, 2 H), 4.22 (t, *J* = 7.2 Hz, 1 H), 3.33 (dq, *J* = 14.7, 7.8 Hz, 2 H), 3.08 (t, *J* = 7.2 Hz, 2 H), 1.68 (m, 2 H), 1.33 (m, 4 H), 0.86 (m, 3 H). ¹³C NMR (75 MHz, D₂O) δ 205.2, 181.6, 139.1 136.3, 134.6, 128.8, 128.7, 126.3, 57.3, 40.9, 30.7, 23.8, 21.8, 13.3.

5. sfGFPs2TAG' Protein Sequence

MAXKGEELFTGVVPILVELGDVNGHKFSVRGEGERGDATNGKLTLKFICTTGKL
PVPWPTLVTTLYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYK
TRAEVKFEGLDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIK
ANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRD
HMVLLEFVTAAGITHGMDELYKGSHHHHH

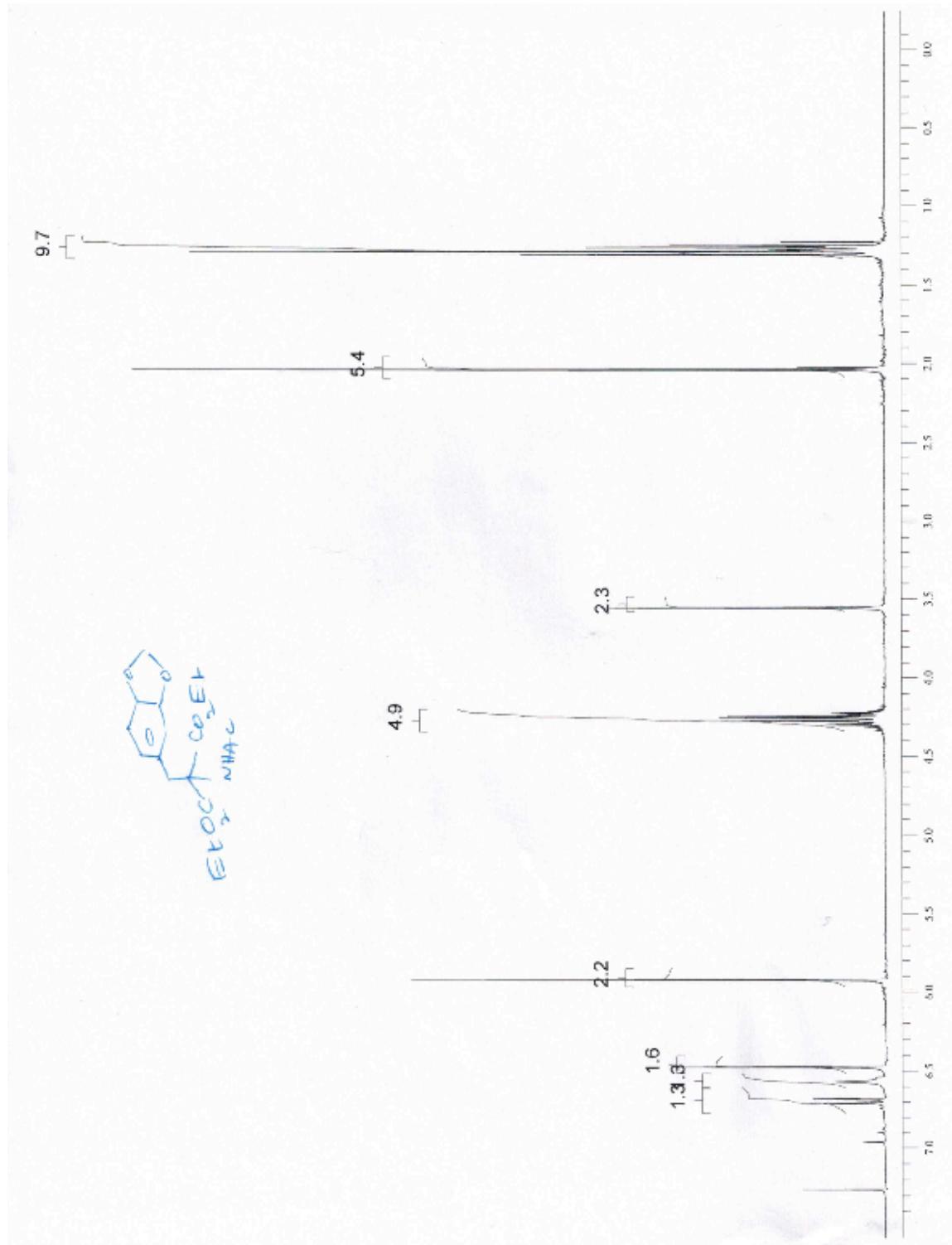
X denotes an amber stop codon for NAA incorporation in this study.

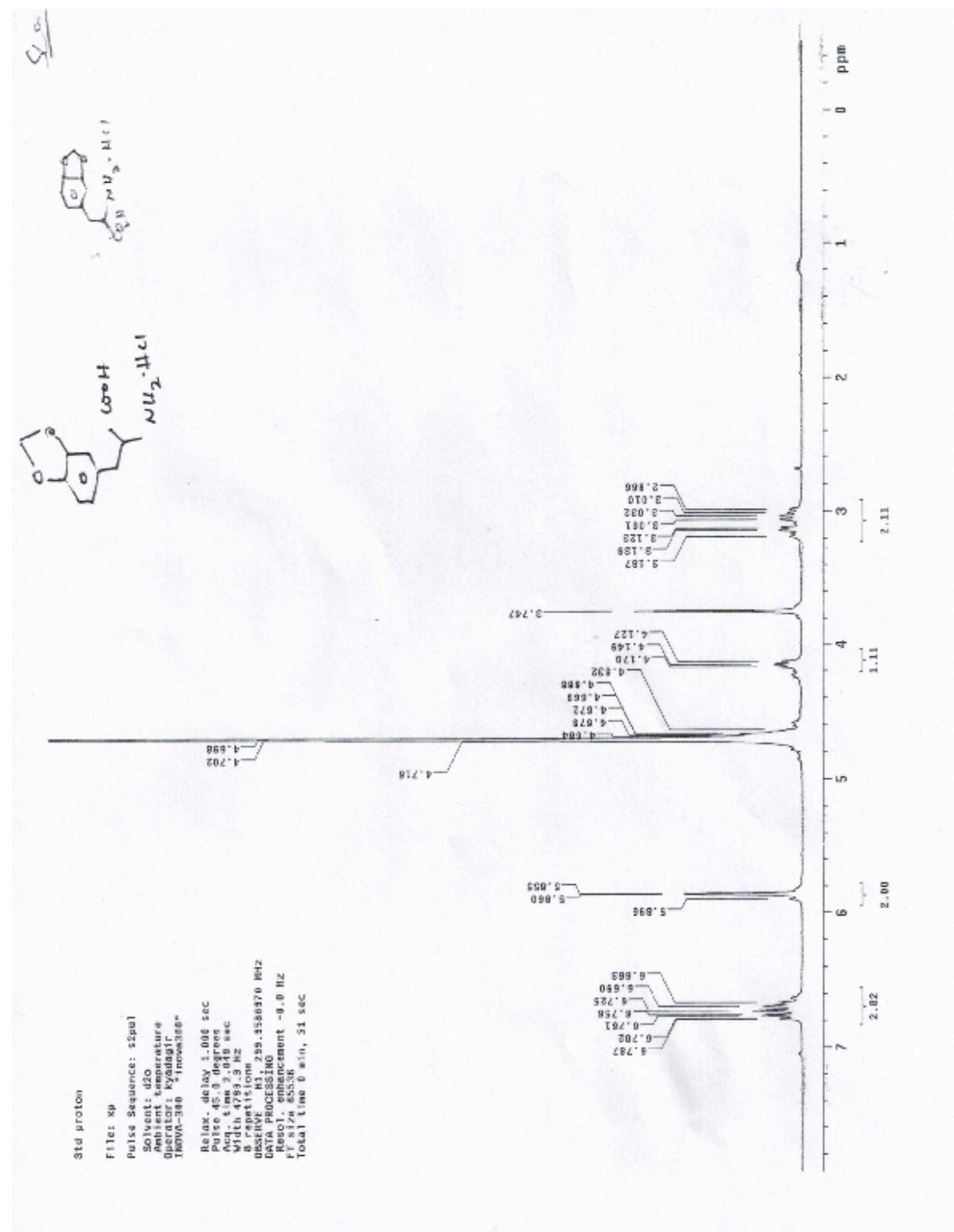
Compounds **5-15** were expressed using the sequence provided above. sfGFP expression of **1-4** used a sequence described previously.^{1,2}

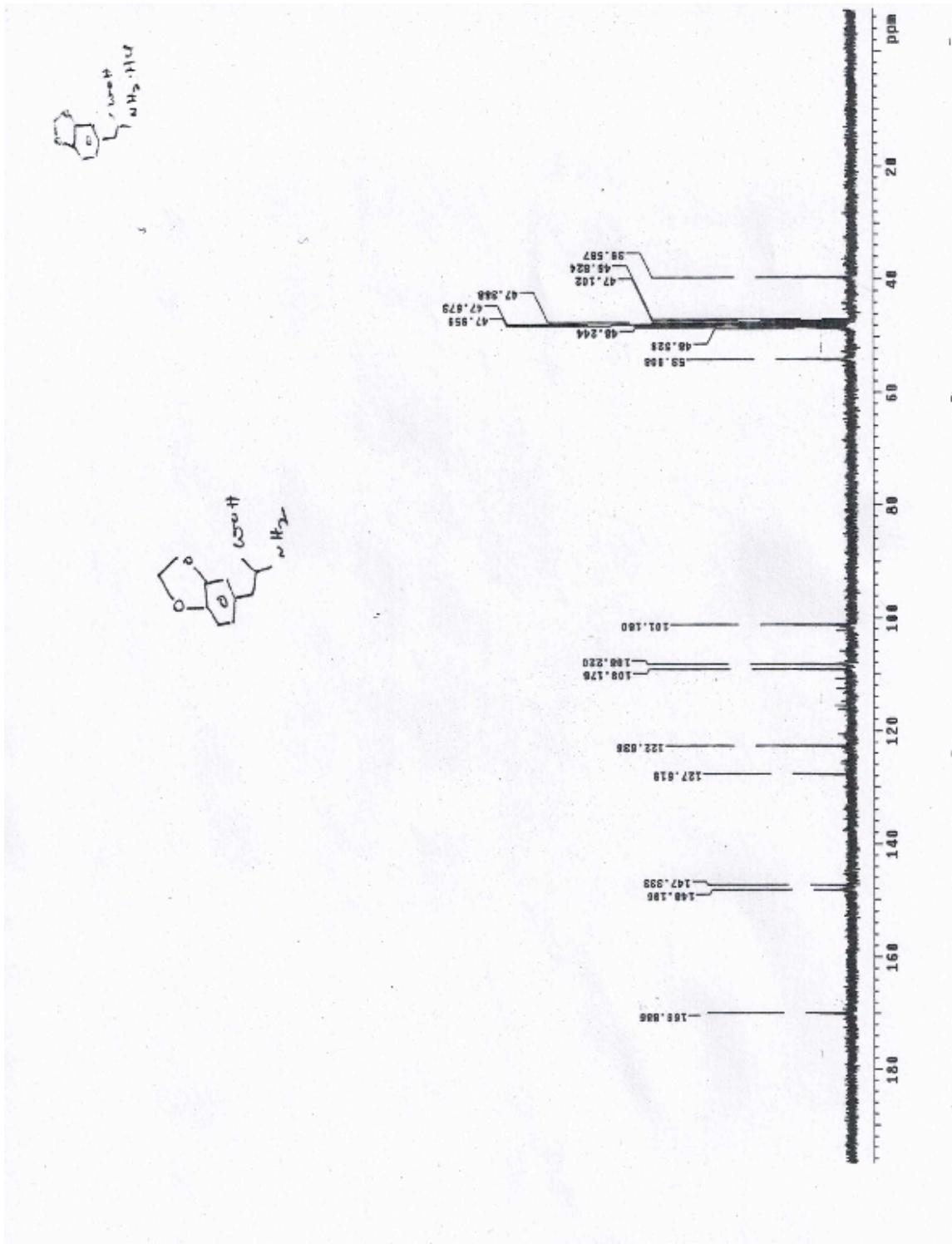
References

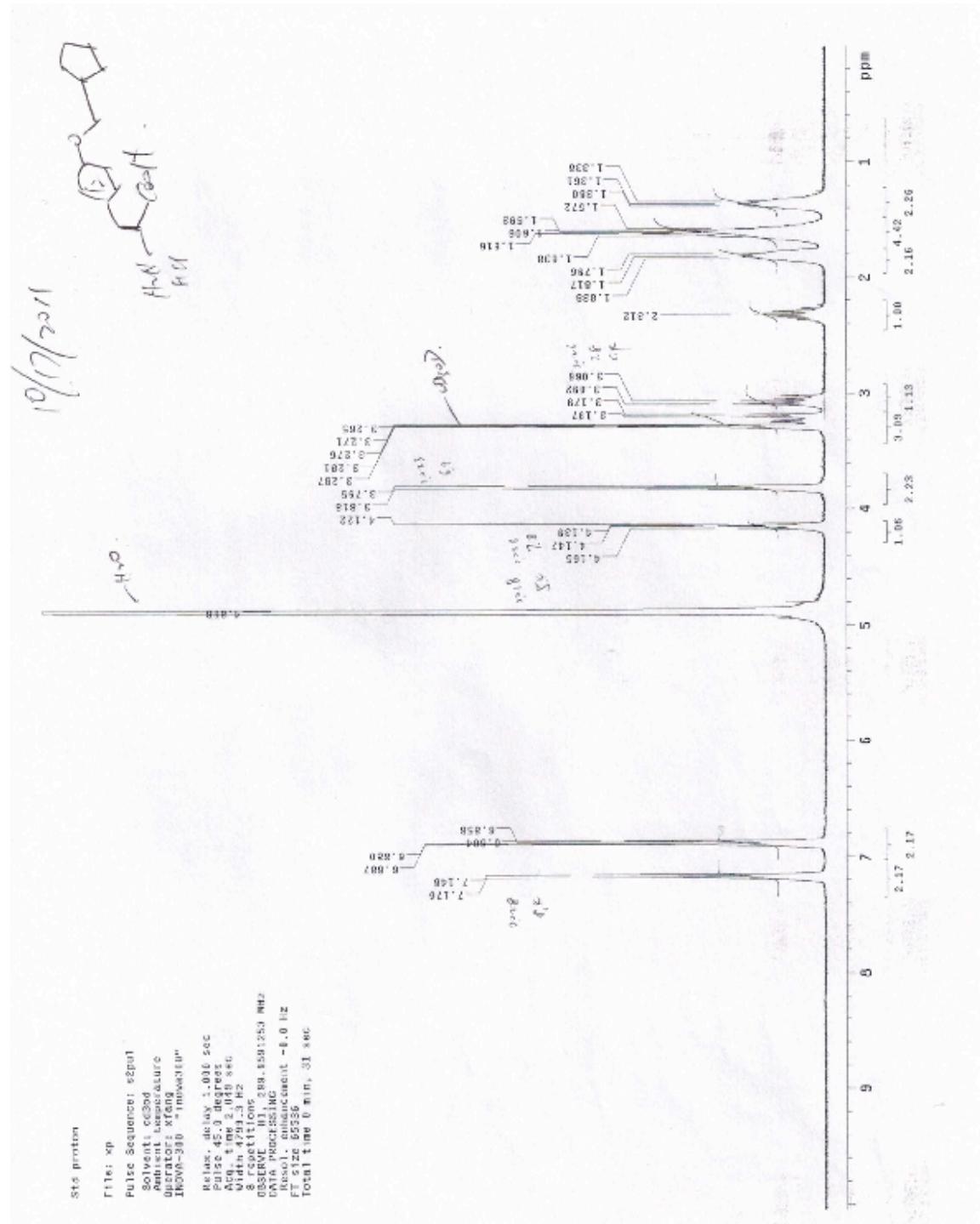
1. Wang, Y.-S.; Fang, X.; Wallace, A. L.; Wu, B.; Liu, W. R. *J. Am. Chem. Soc.* **2012**, *134*, 2950-2953.
2. Wang, Y.-S.; Fang, X.; Chen, H.-Y.; Wu, B.; Wang, Z. U.; Hilty, C.; Liu, W. R. *ACS Chem. Biol.* **2013**, *8*, 405-415.
3. Wu, B.; Wang, Z.; Huang, Y.; Liu, W. R. *ChemBiochem.* **2012**, *13*, 1405-1408.
4. a.) Cabiddu, M. G.; Cadoni, E.; Montis, S. D.; Fattouni, C.; Melis, S.; Usai, M. *Tetrahedron* **2003**, *59*, 4383-4387. b.) Archer, A. W.; Claret, P.A.; Hayman, D. F. *J. Chem. Soc. (B)* **1971**, 1231-1240.
5. Drew, S. L.; Lawrence, A. L.; Sherburn, M. S. *Ang. Chem. Int Ed.* **2013**, *52*, 4221-4224.
6. Humphrey, C. E.; Furegati, M.; Laumen, K.; Vecchia, L. L.; Leutert, T.; Muller-Hartwig, C. D.; Vogtle, M. *Org. Process Res. Dev.* **2007**, *11*, 1069-1075.
7. Carroll, F. I.; Blough, B. E.; Abraham, P.; Mills, A. C.; Holleman, J. A.; Wolkenhauer, S. A.; Decker, A. M.; Landavazo, A.; McElroy, K. T.; Navarro, H. A.; Gatch, M. B.; Forster, M. J. *J. Med. Chem.* **2009**, *52*, 6768-6781.
8. Liu, Y.; Yao, B.; Deng, C.-L.; Tang, R.-Y.; Zhang, X.-G.; Li, J.-H. *Org. Lett.* **2011**, *13*, 2184-2187.
9. Meltzer, P. Z.; Butler, D.; Deschamps, J. R.; Madras, B. K. *J. Med. Chem.* **2006**, *49*, 1420-1432.
10. Ruan, J.; Saidi, O.; Iggo, J. A.; Xiao, J. *J. Am. Chem. Soc.* **2008**, *130*, 10510-10511.

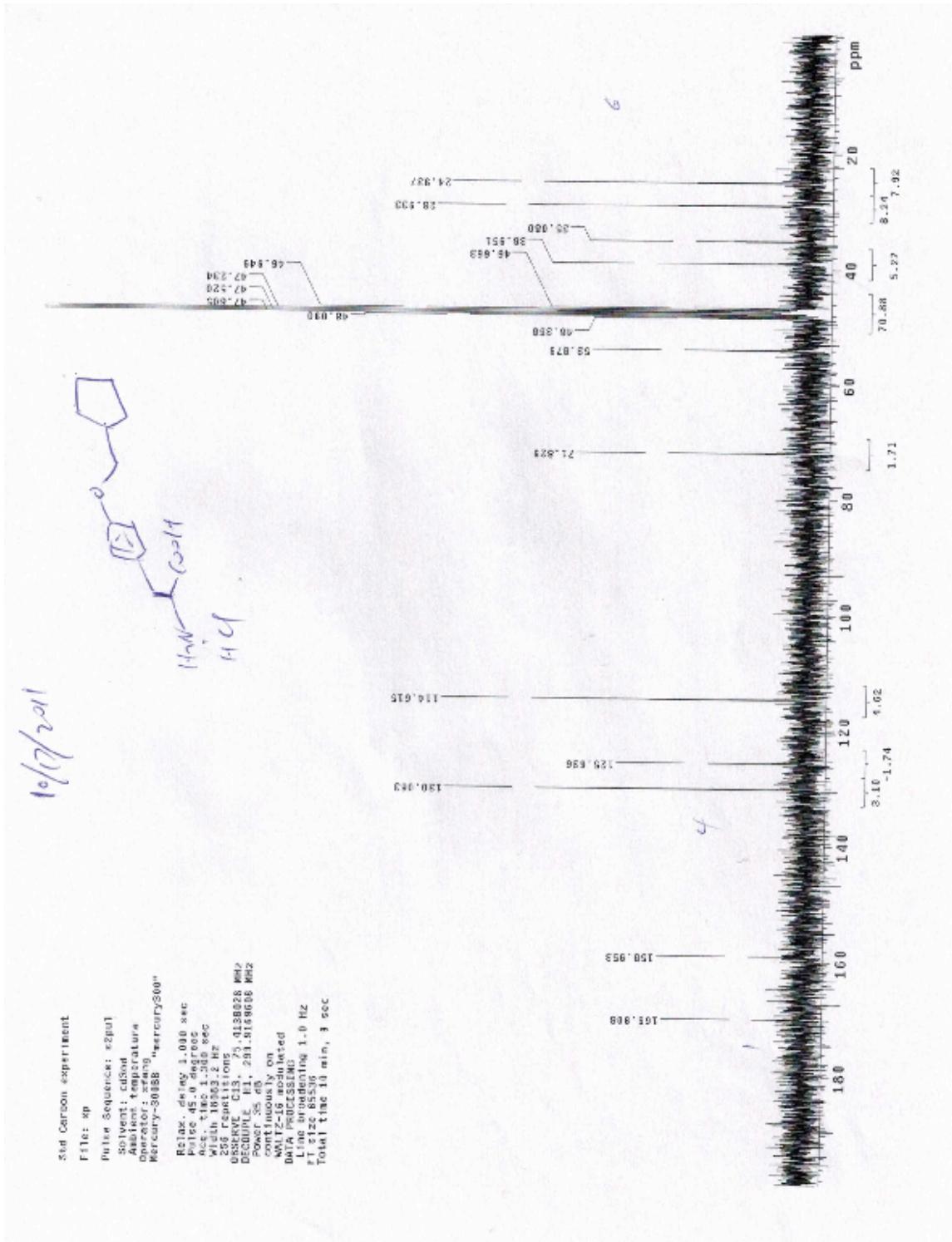
6. NMR Data

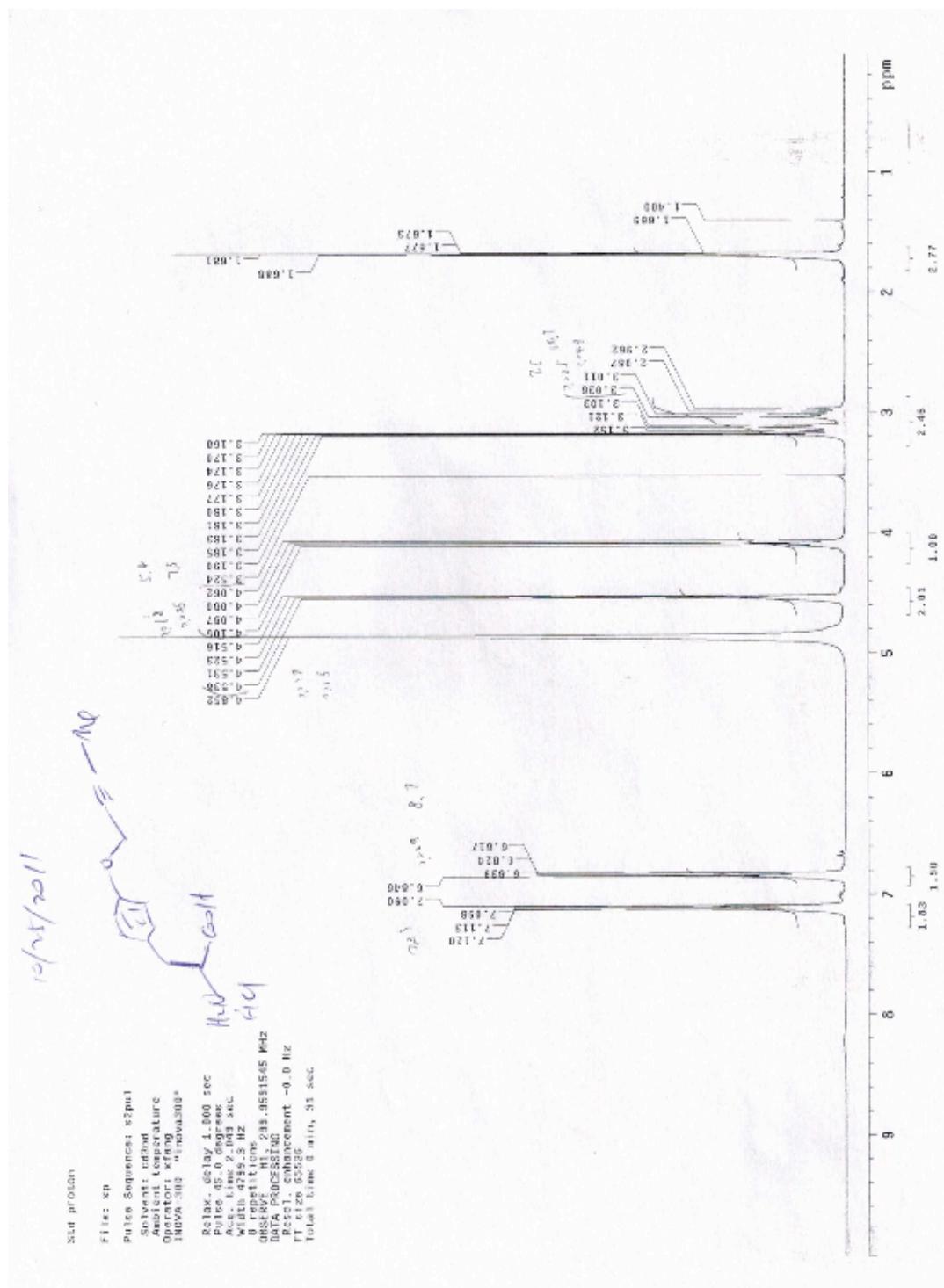


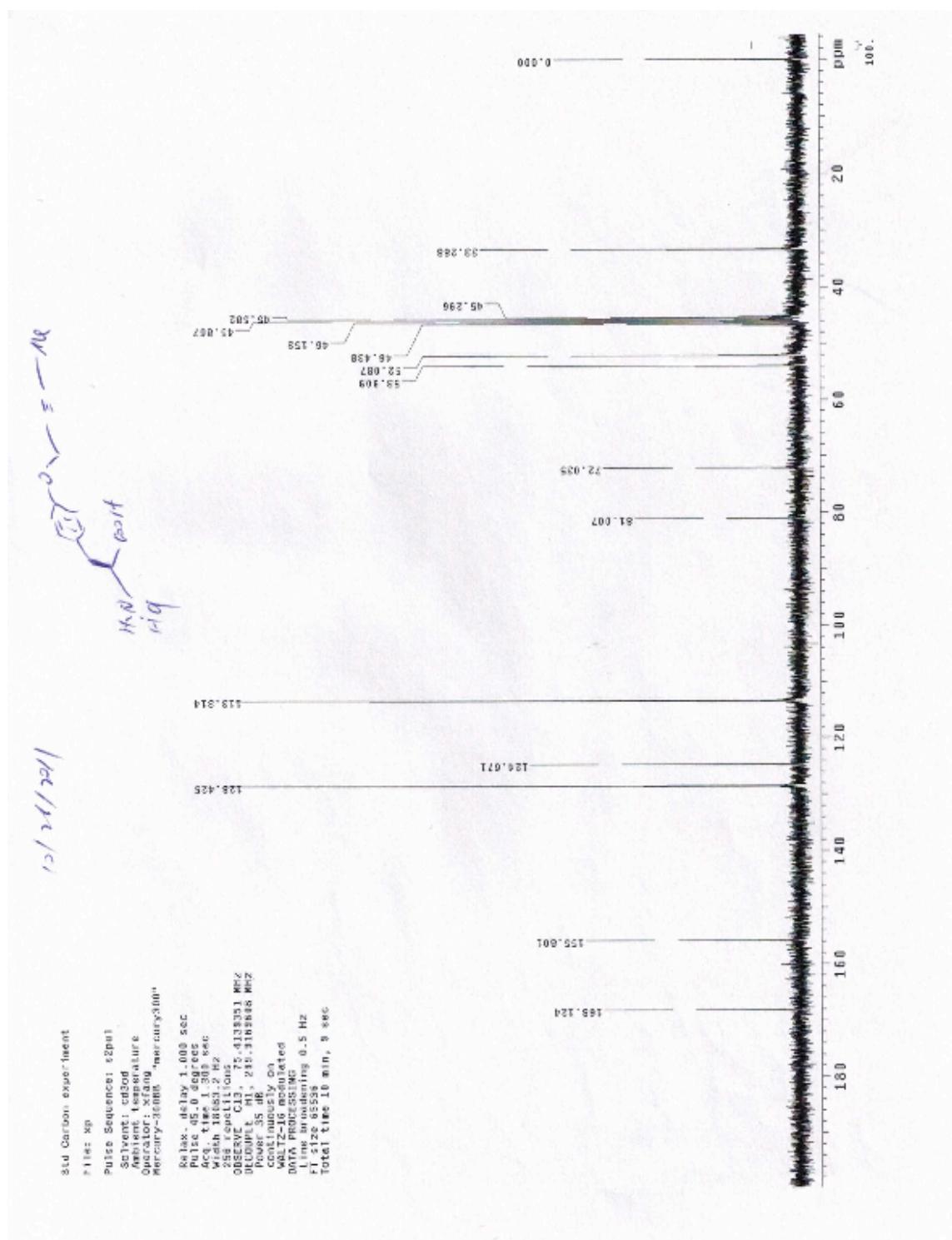


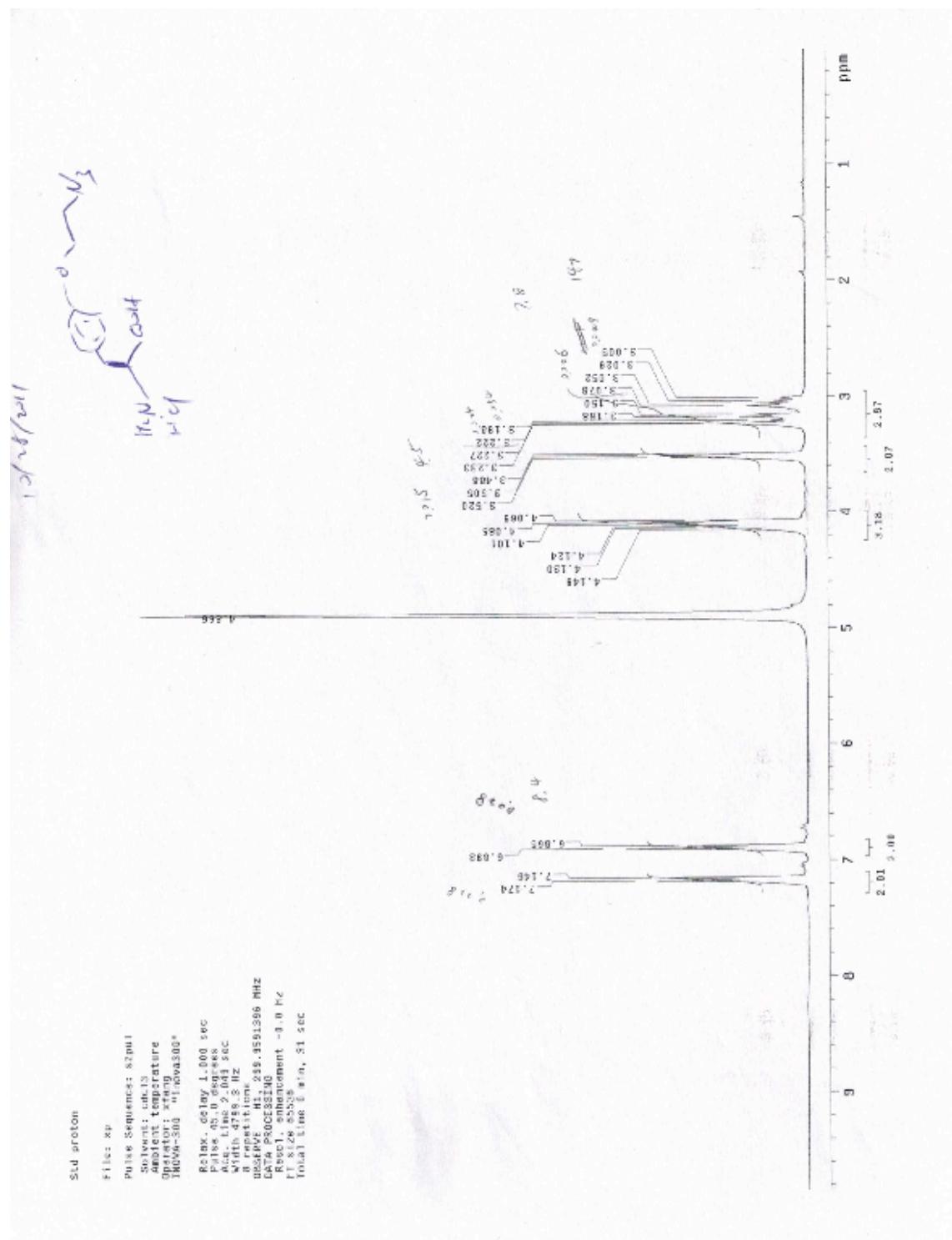


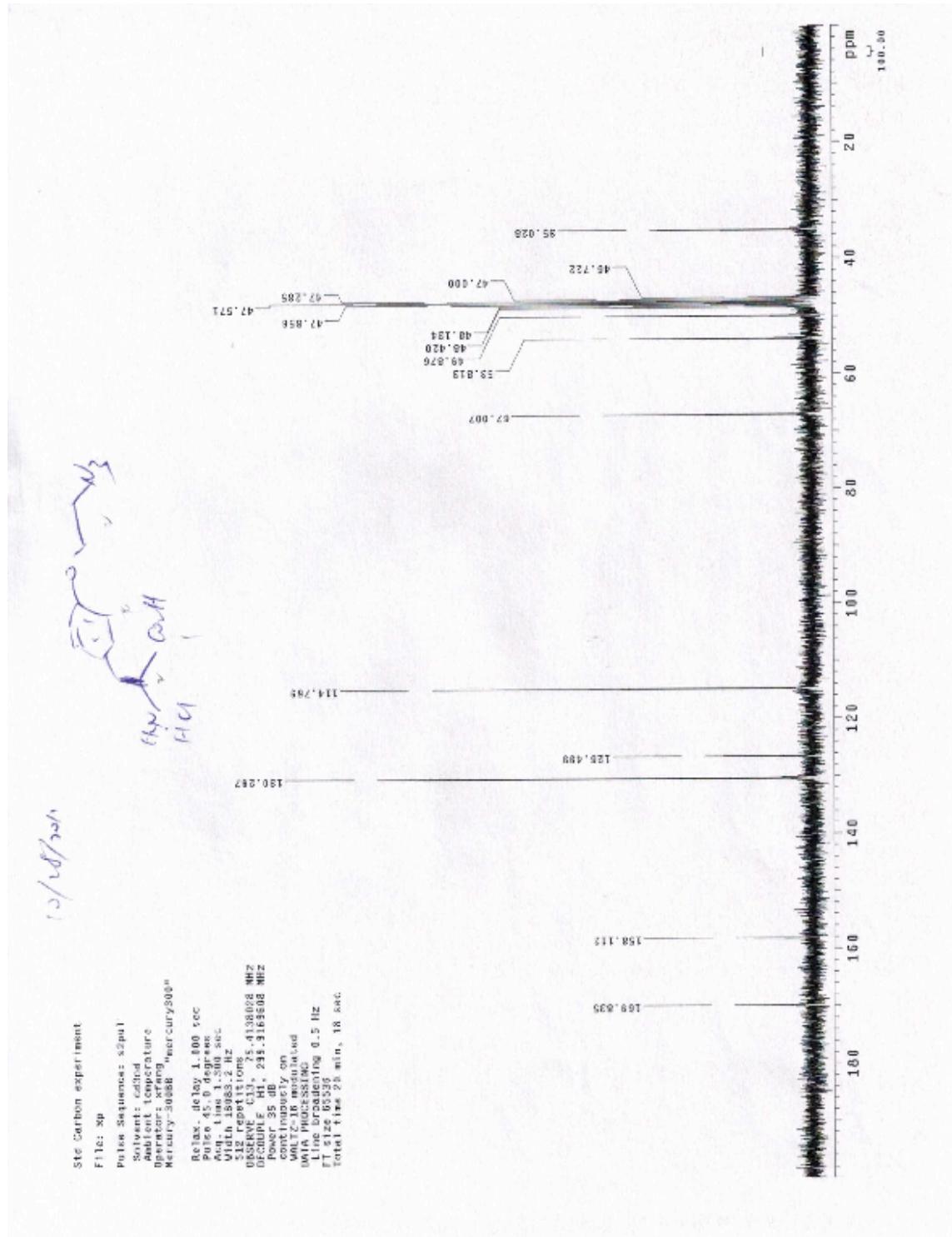


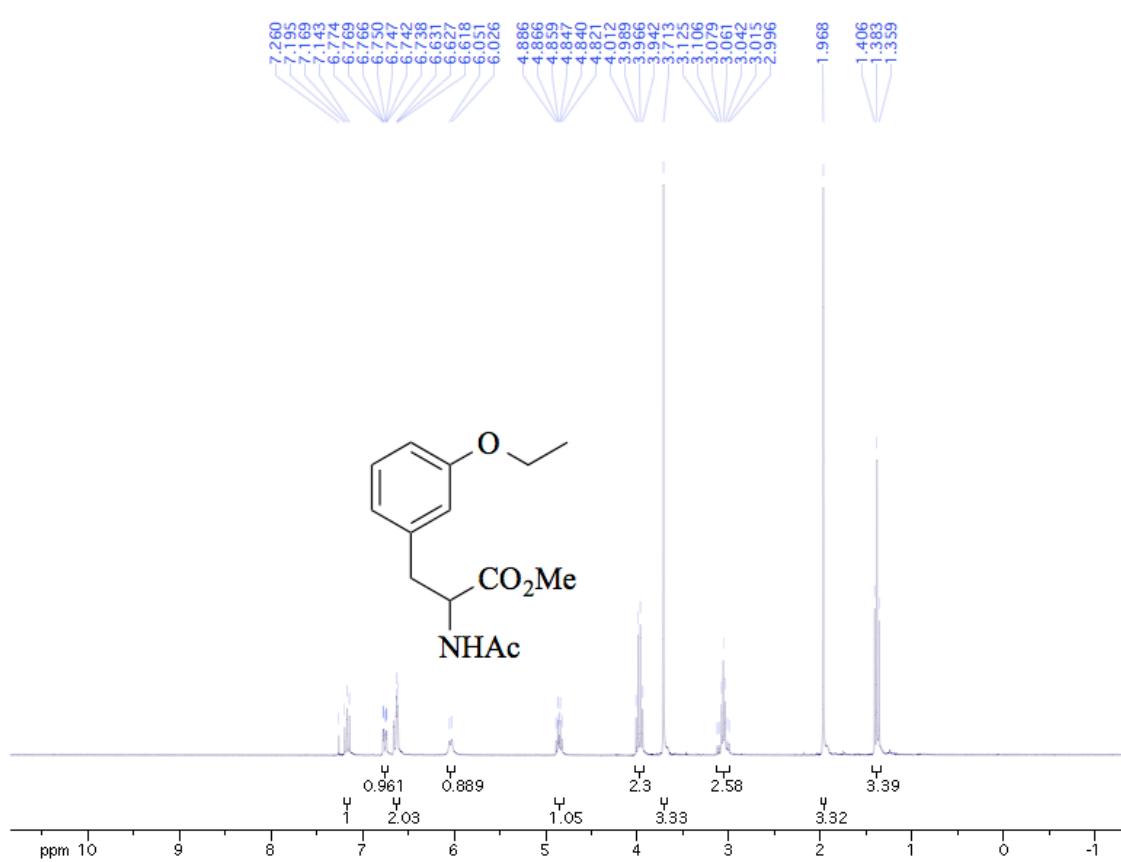


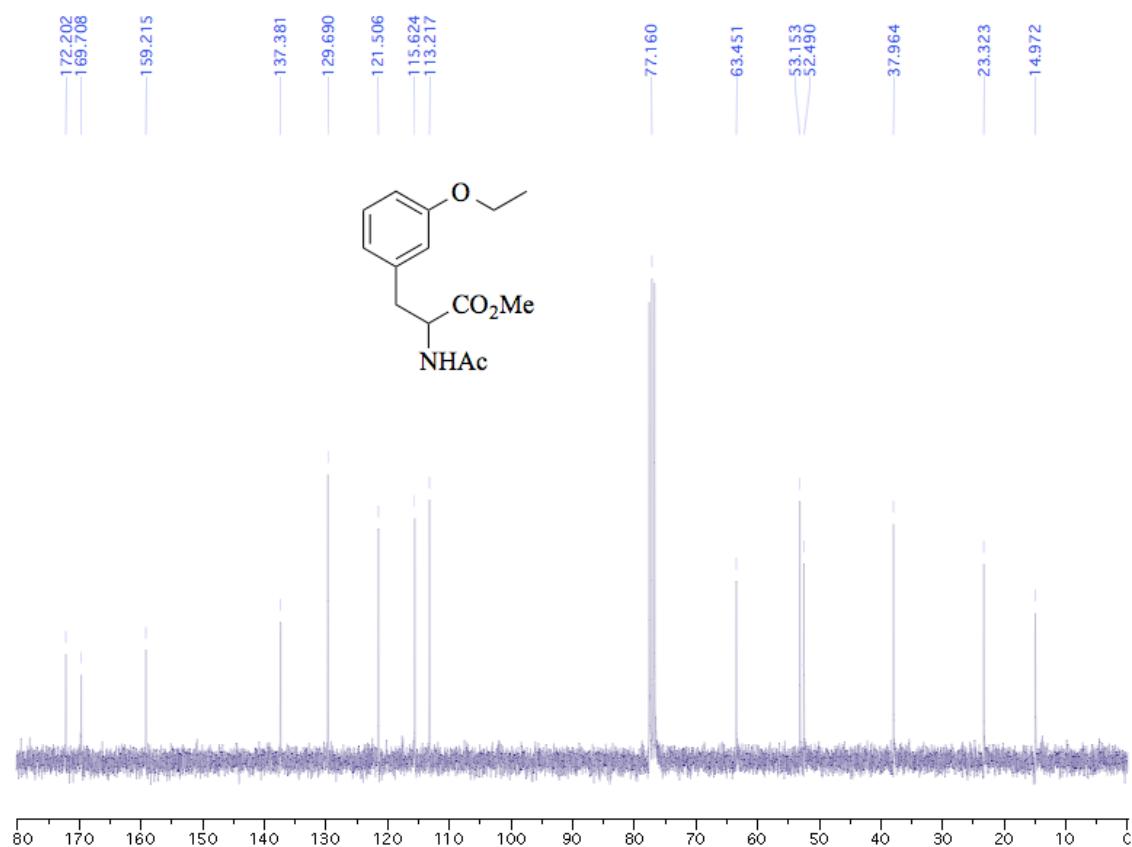


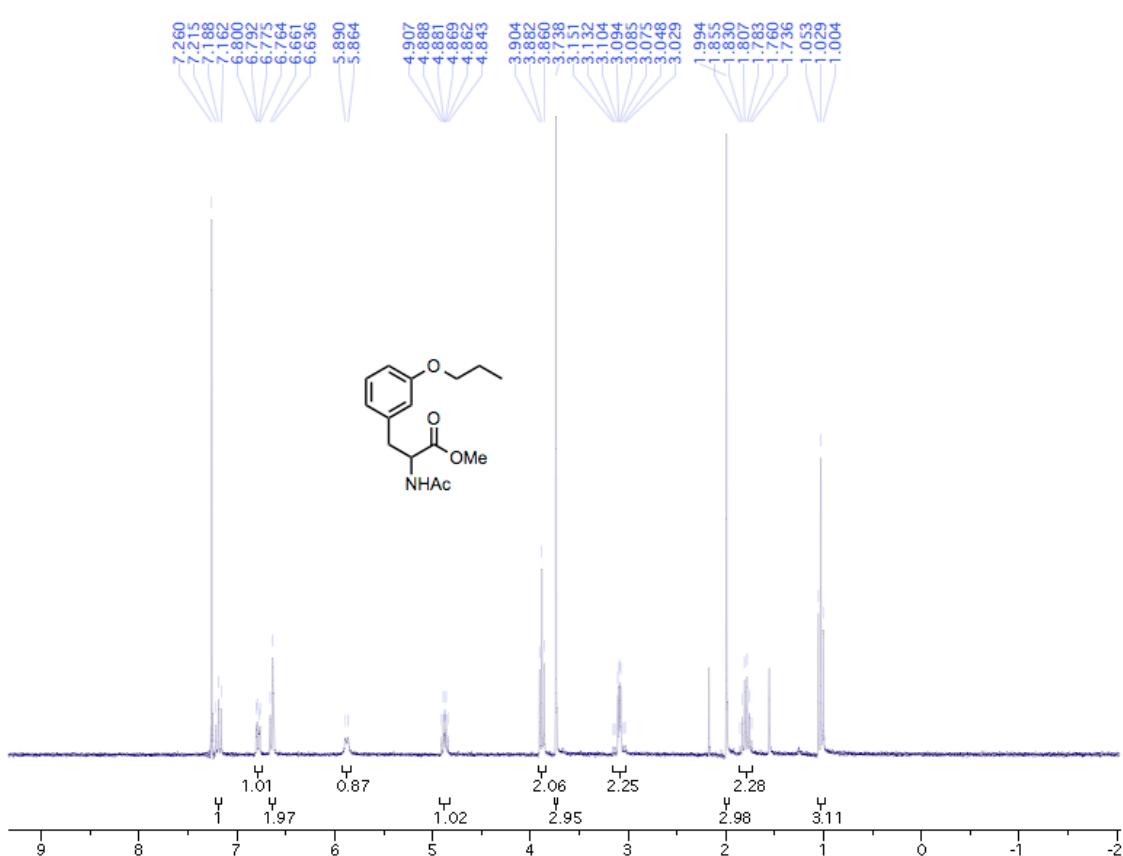


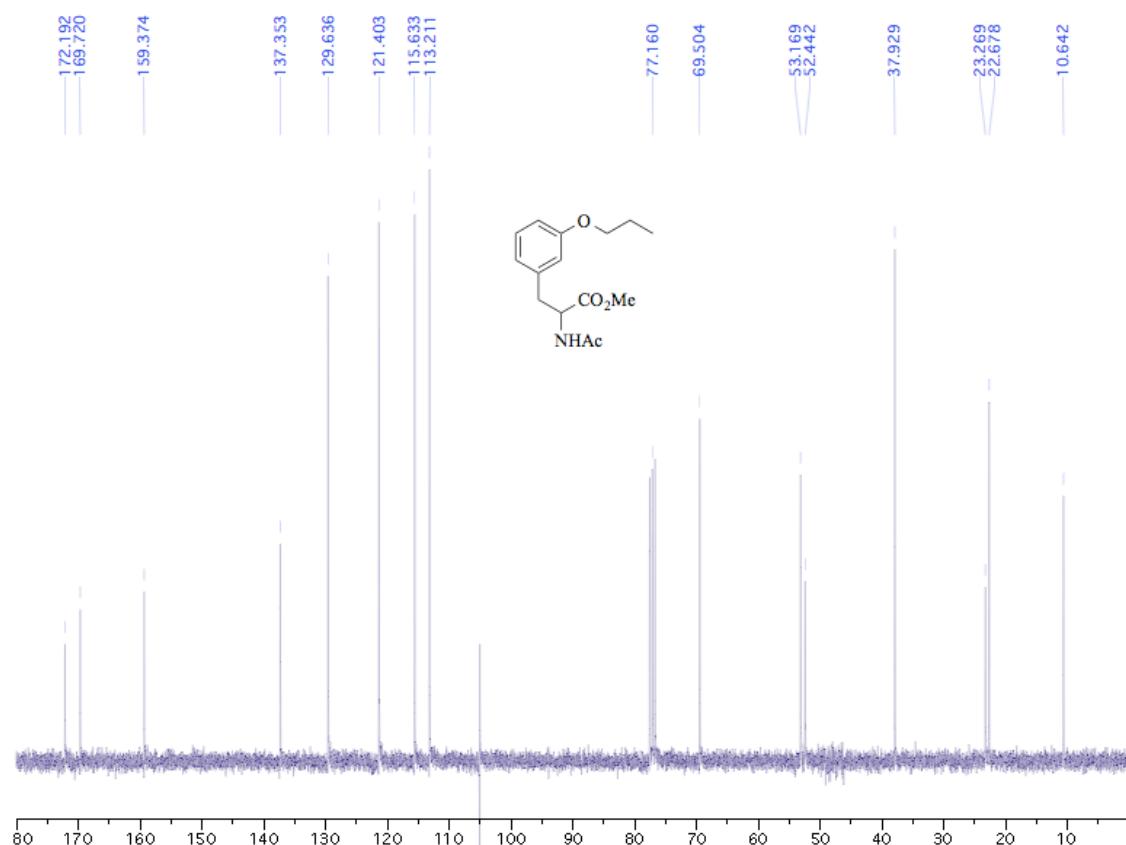


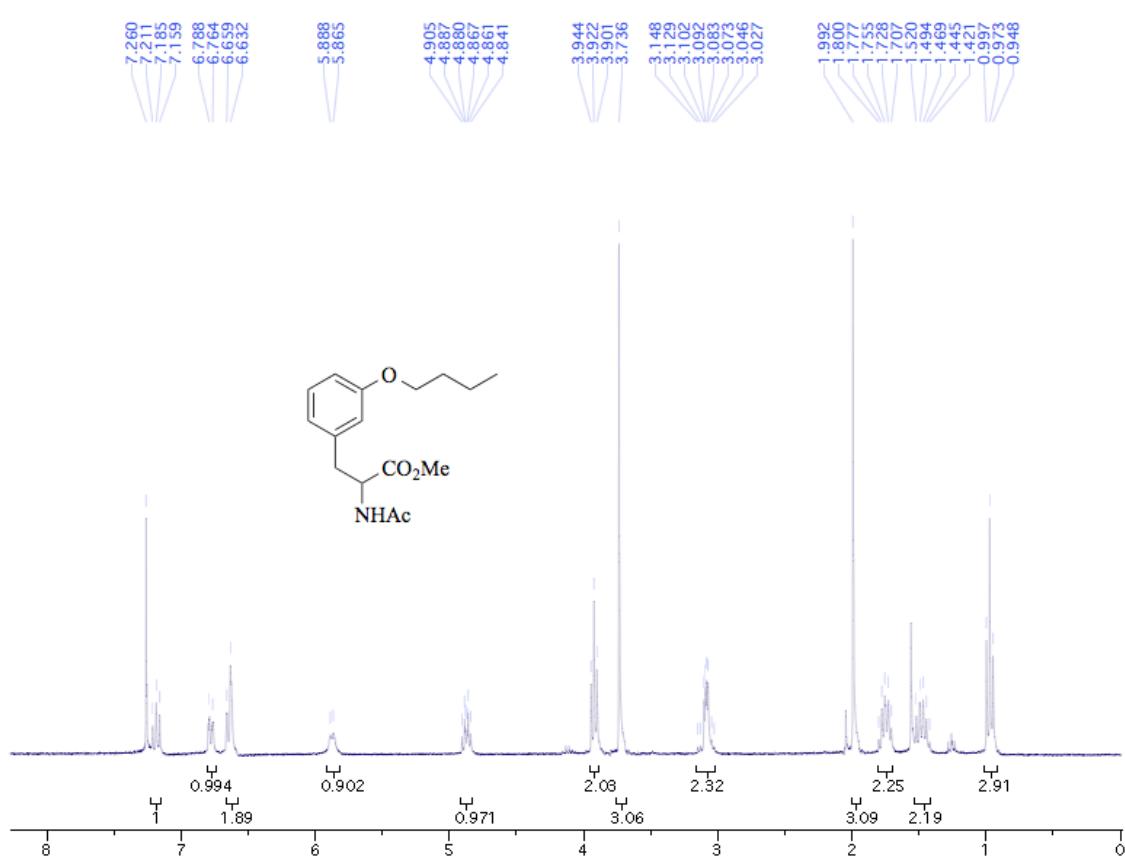


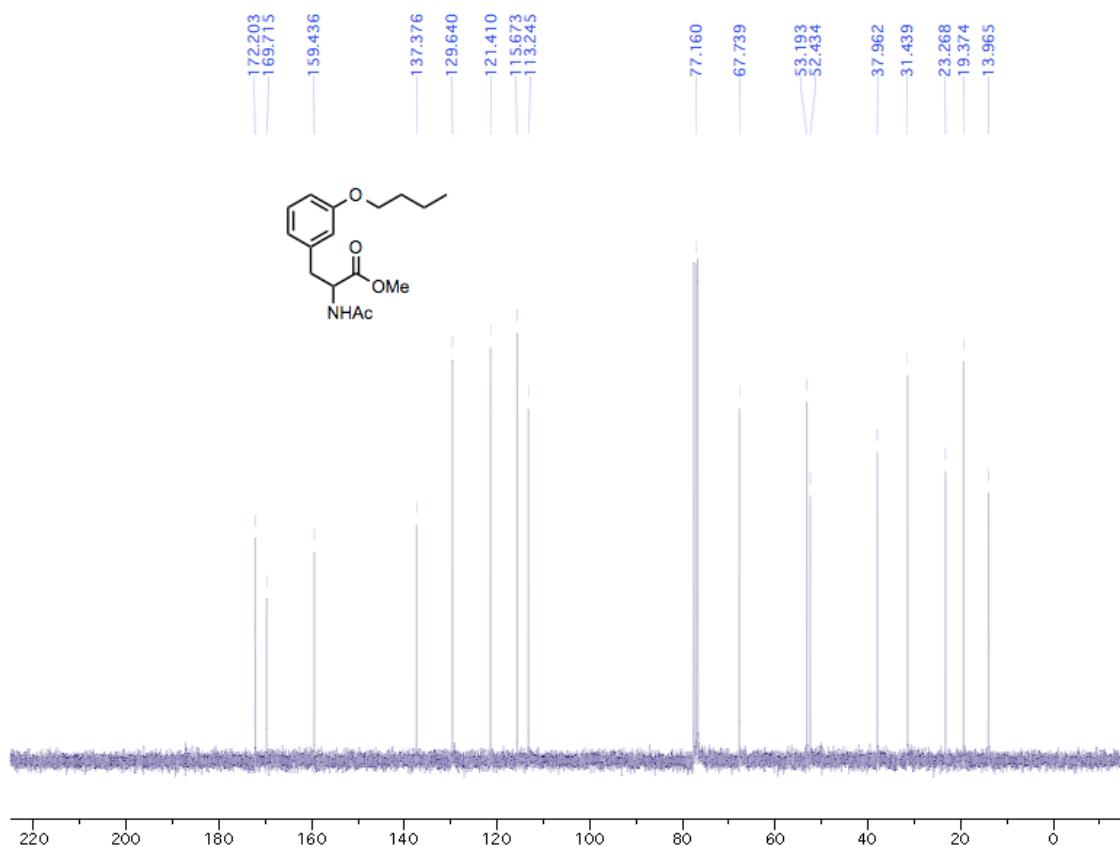


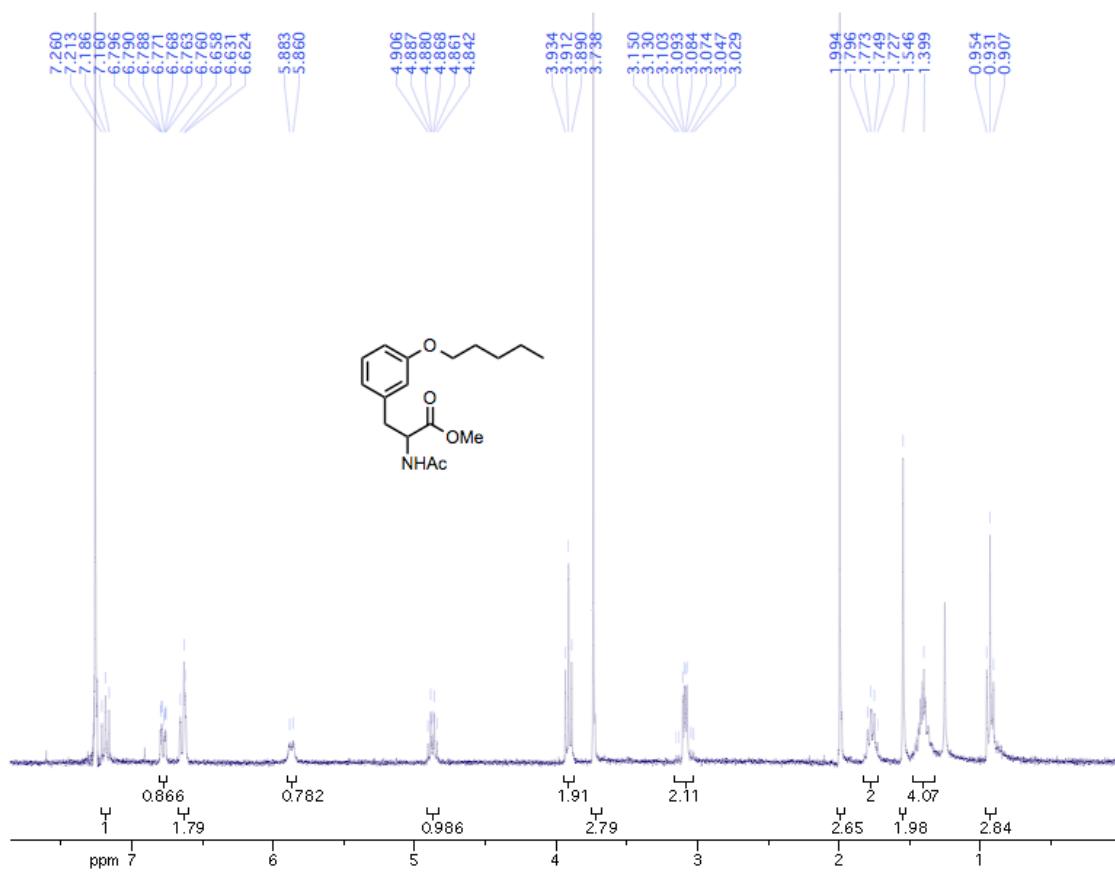


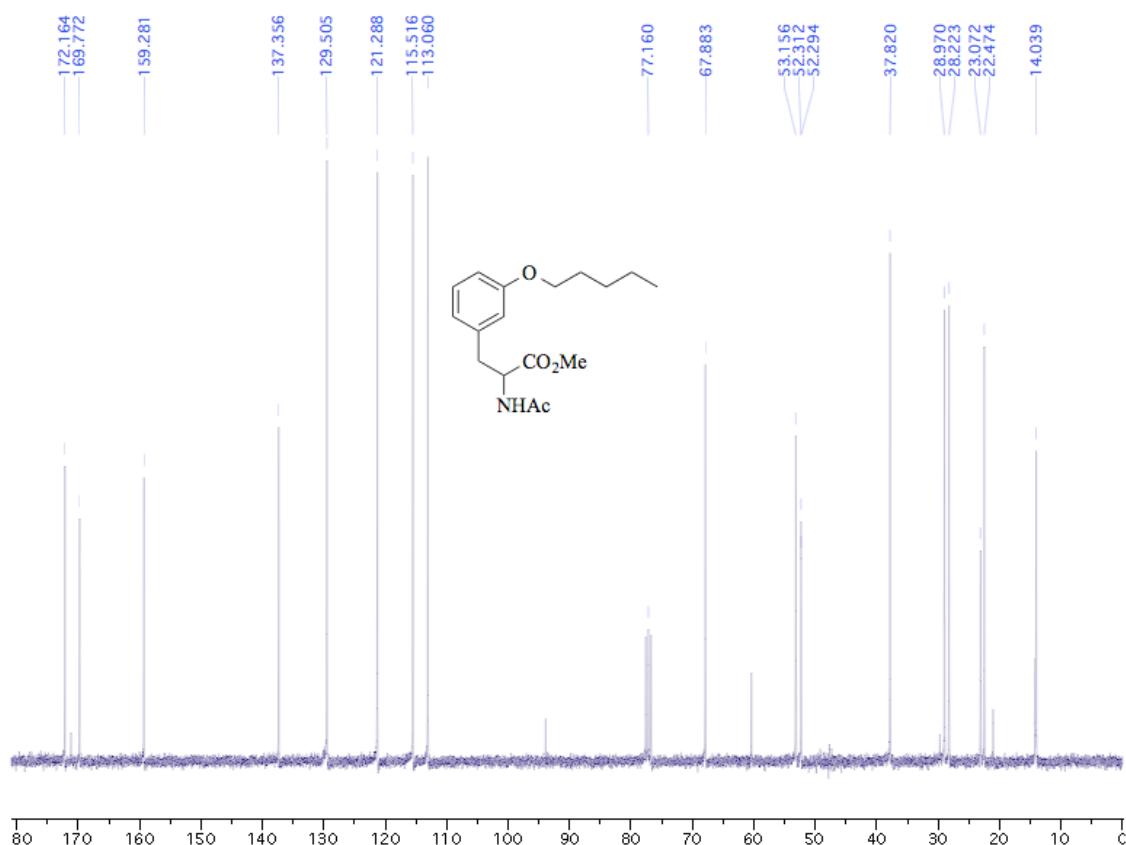


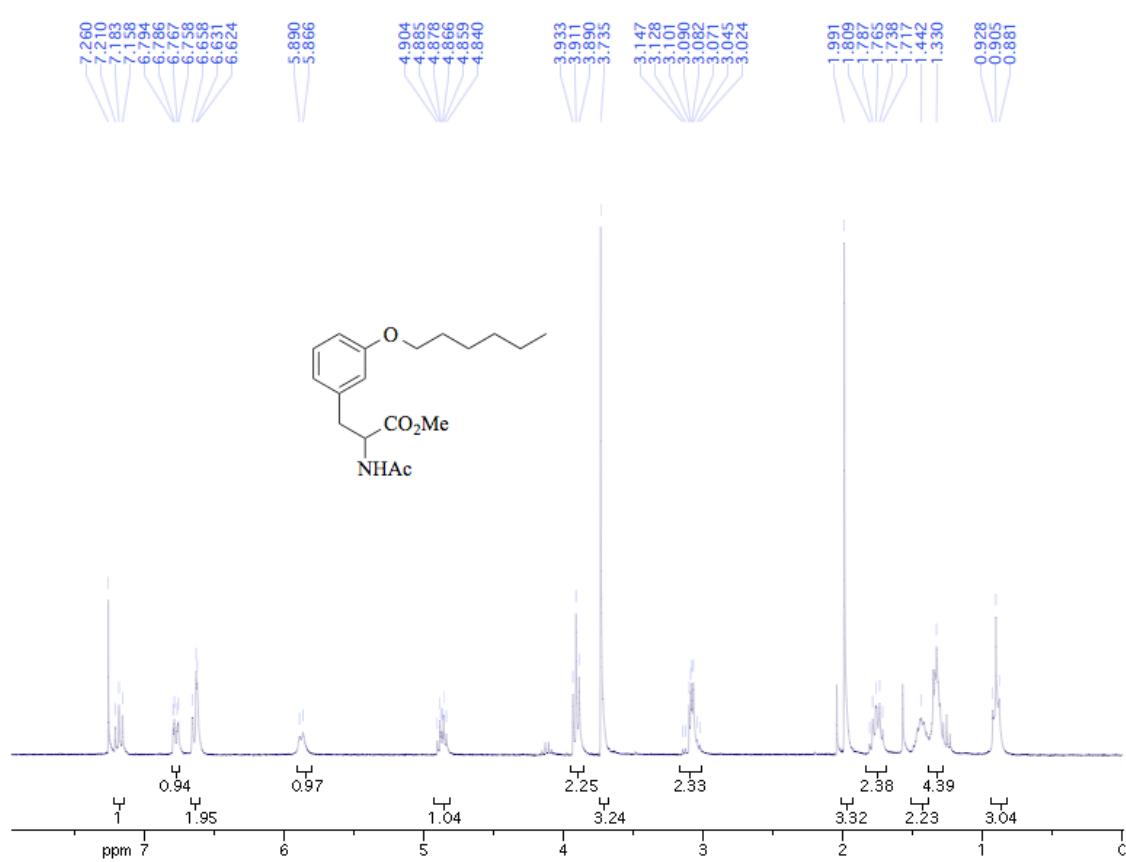


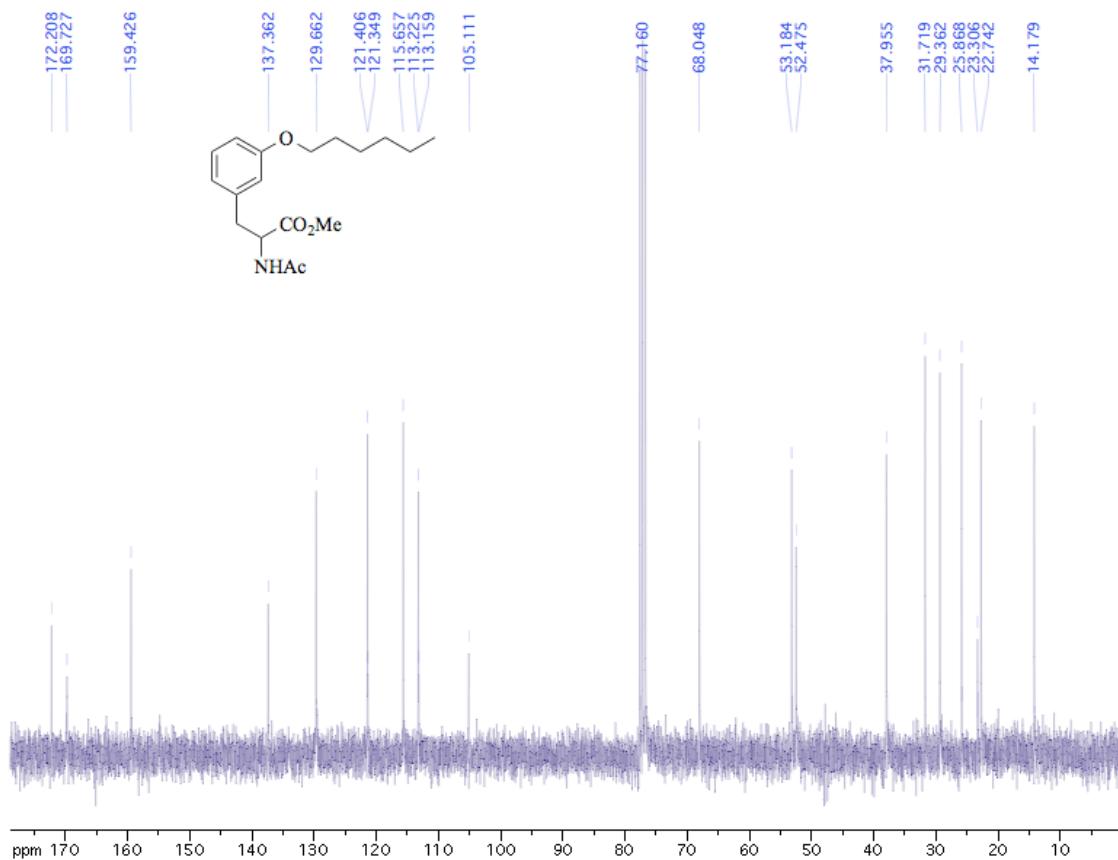


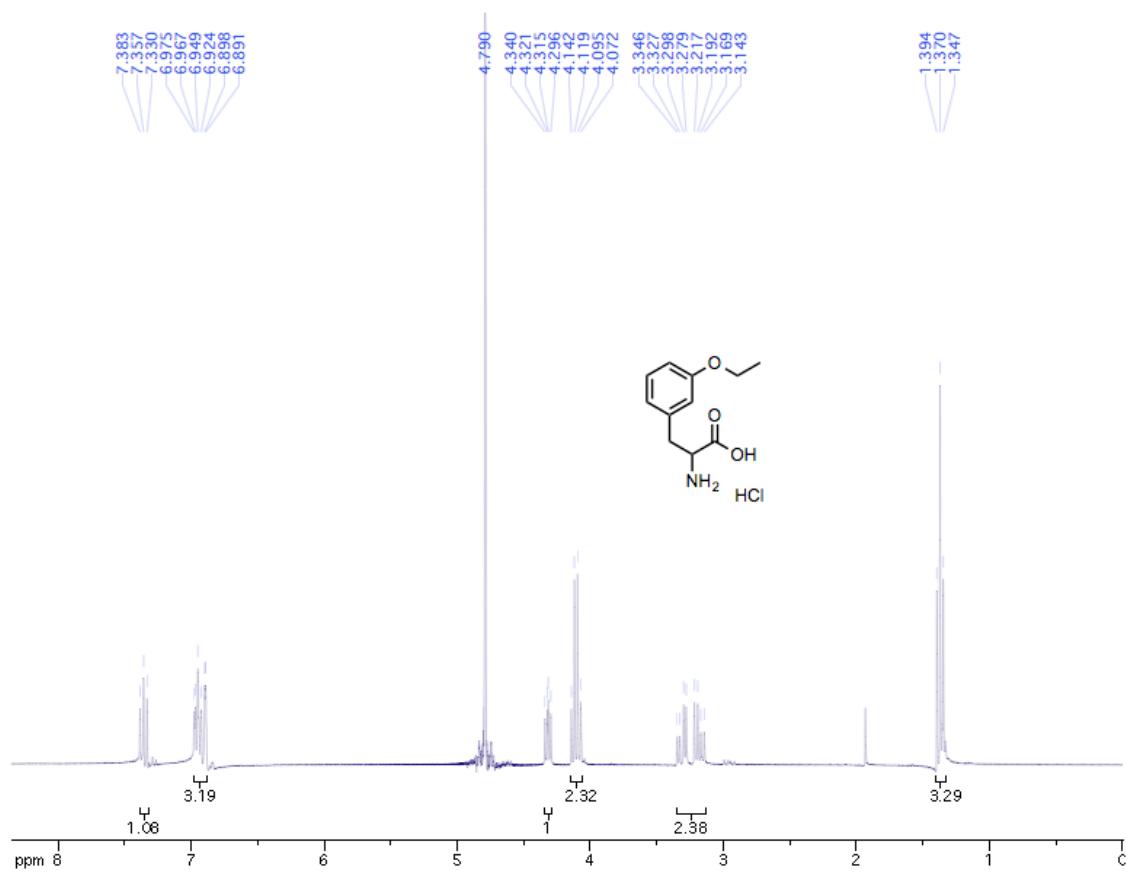


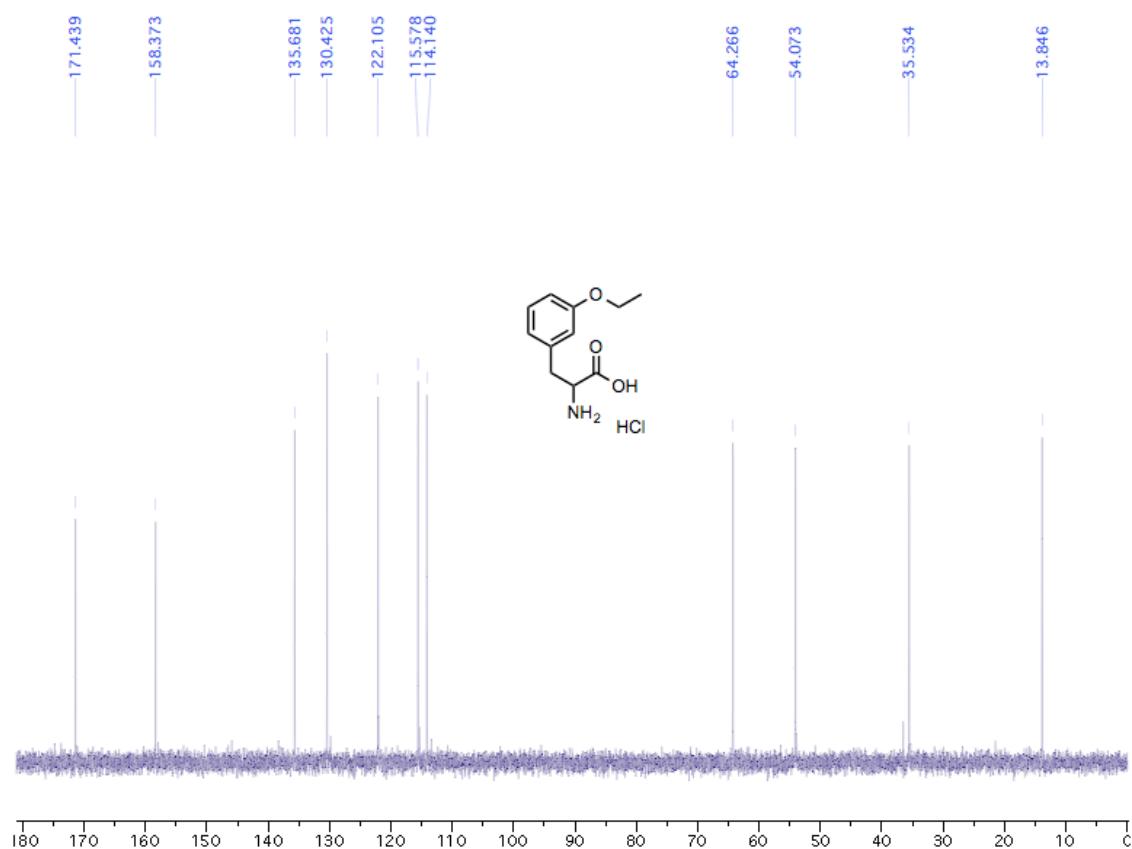


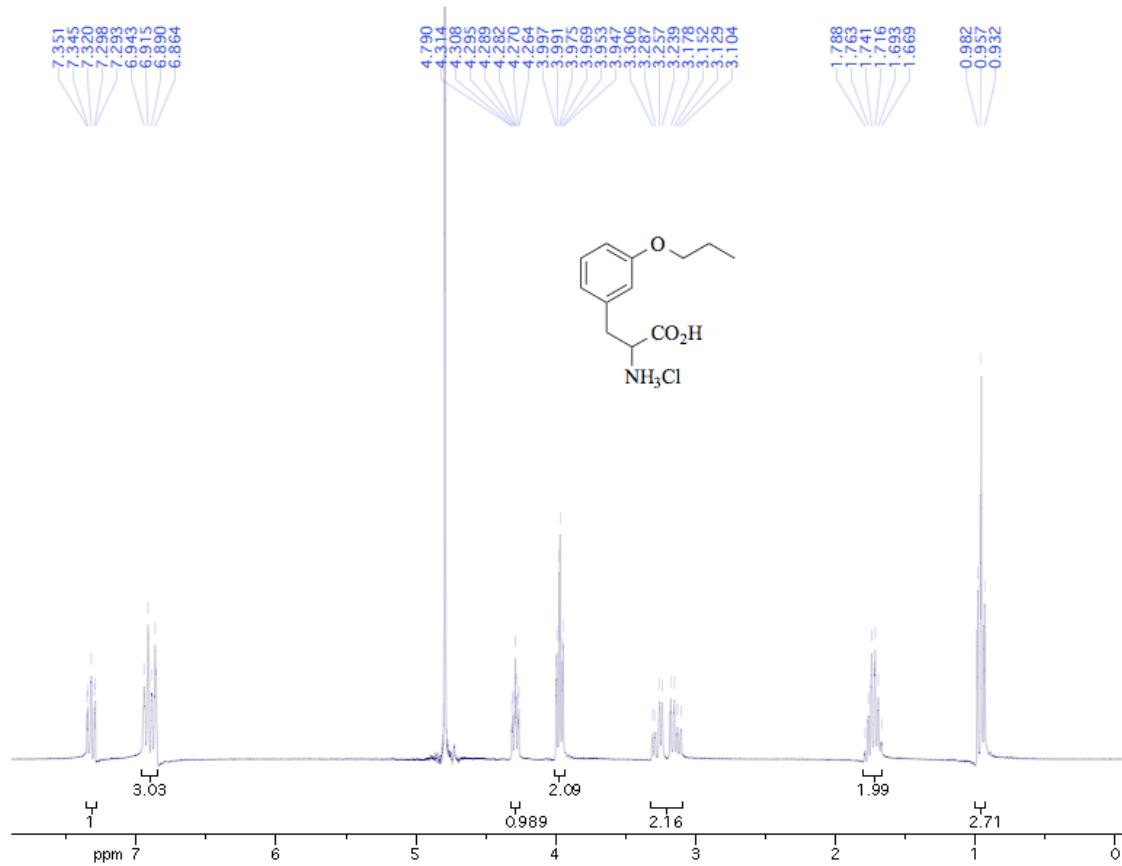


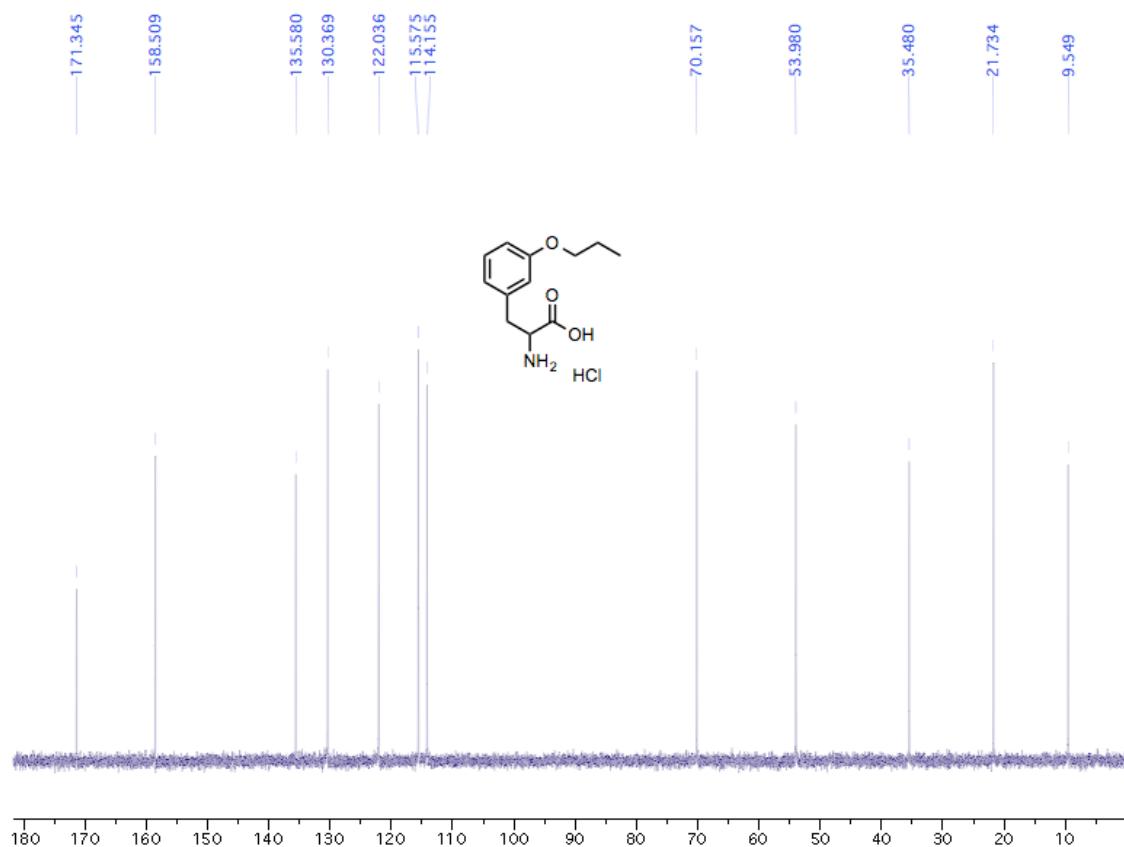


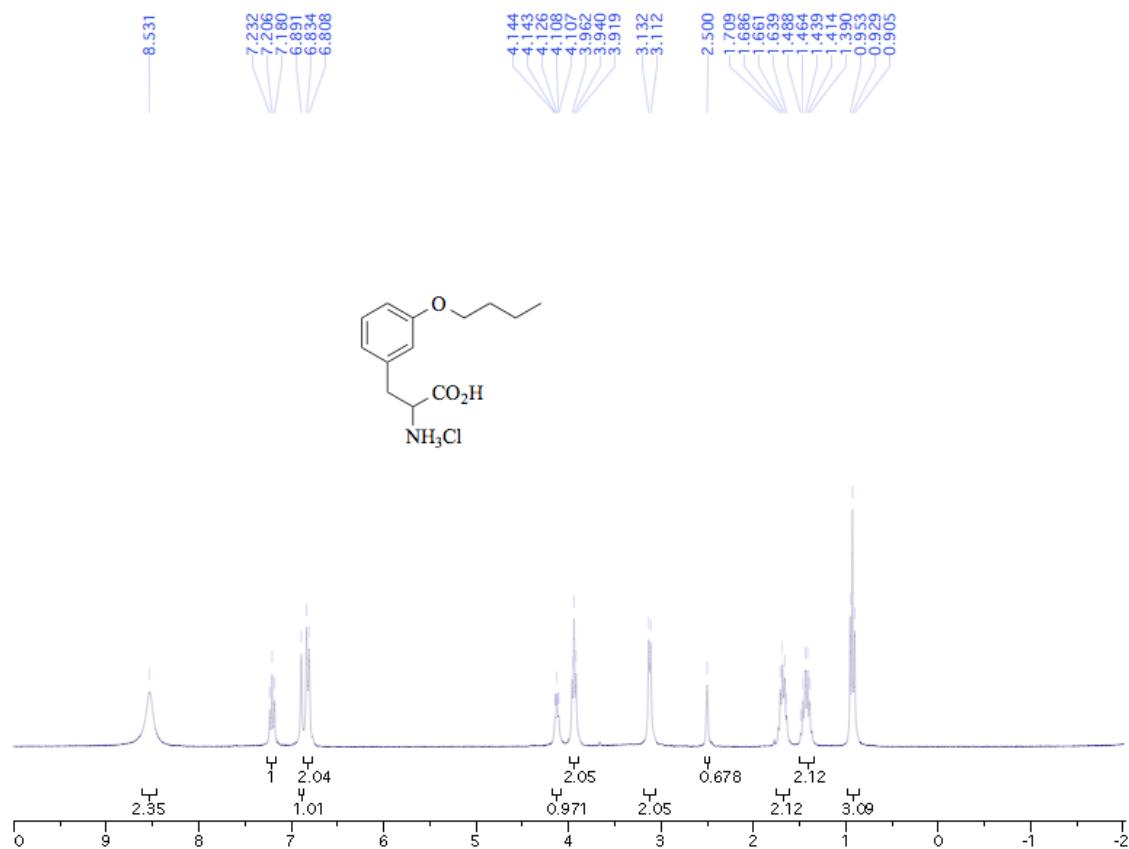


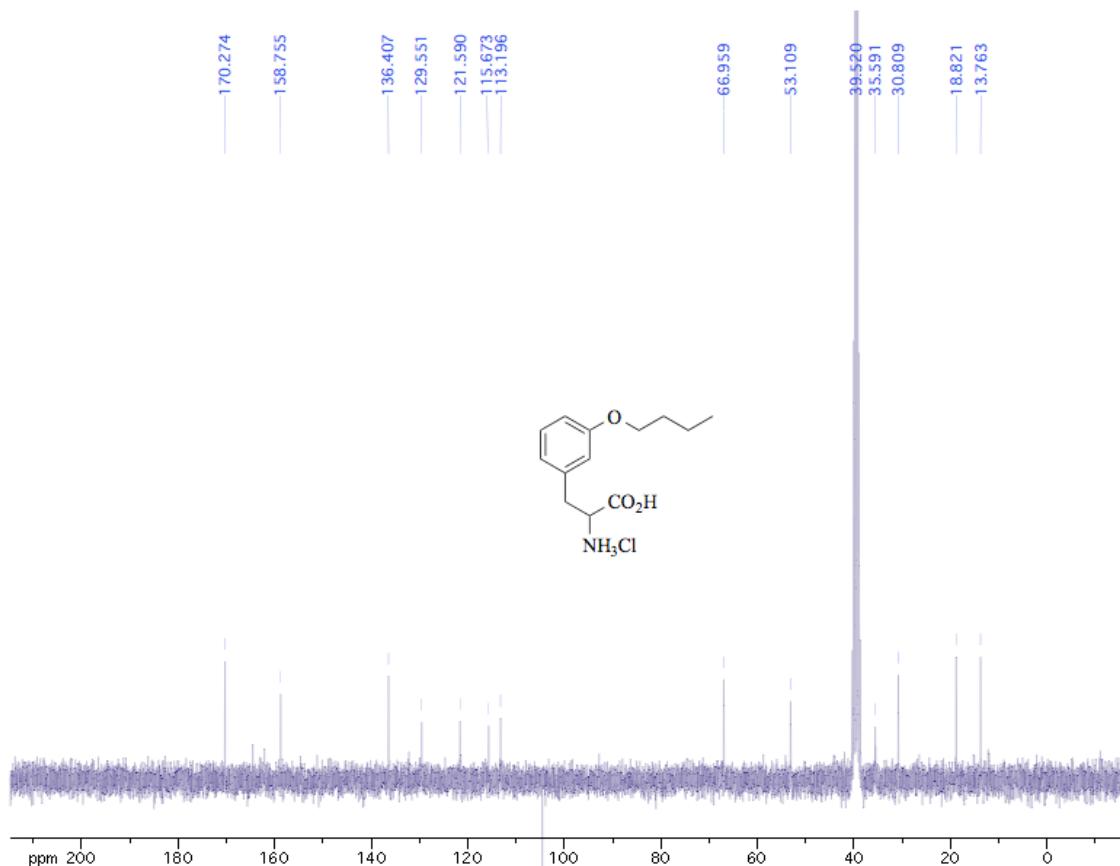


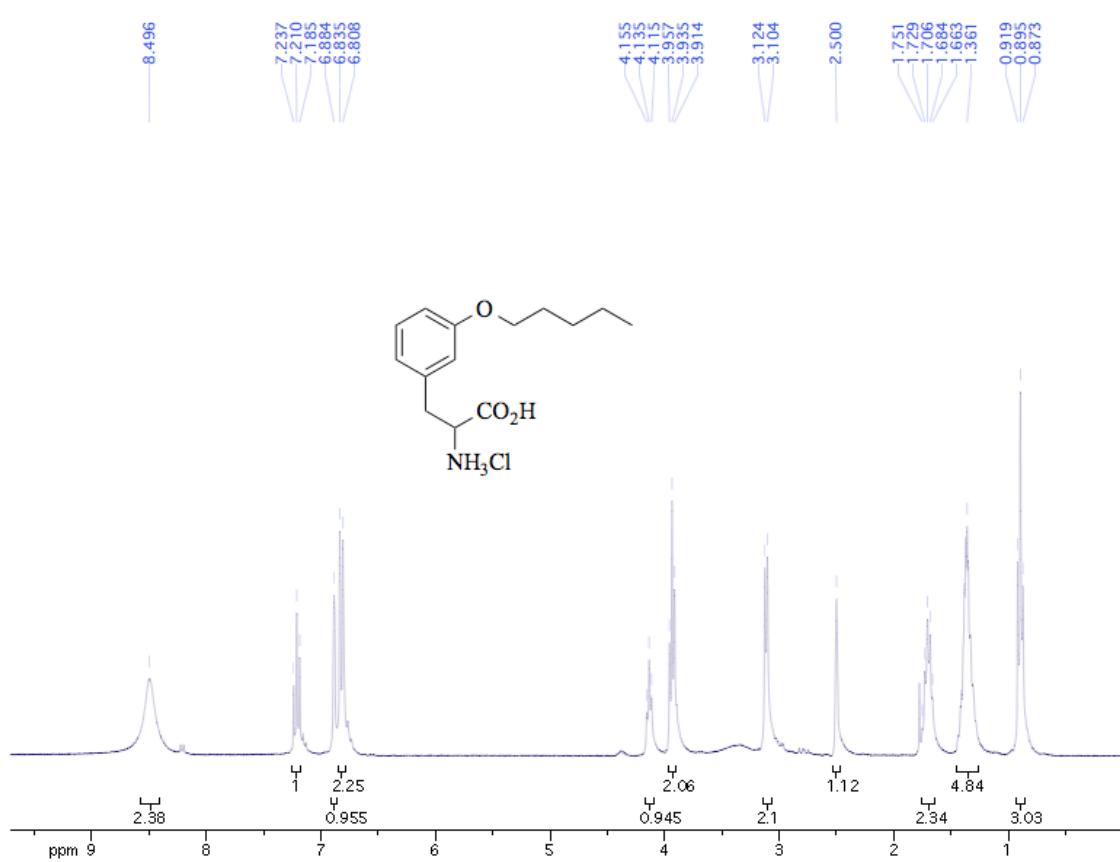


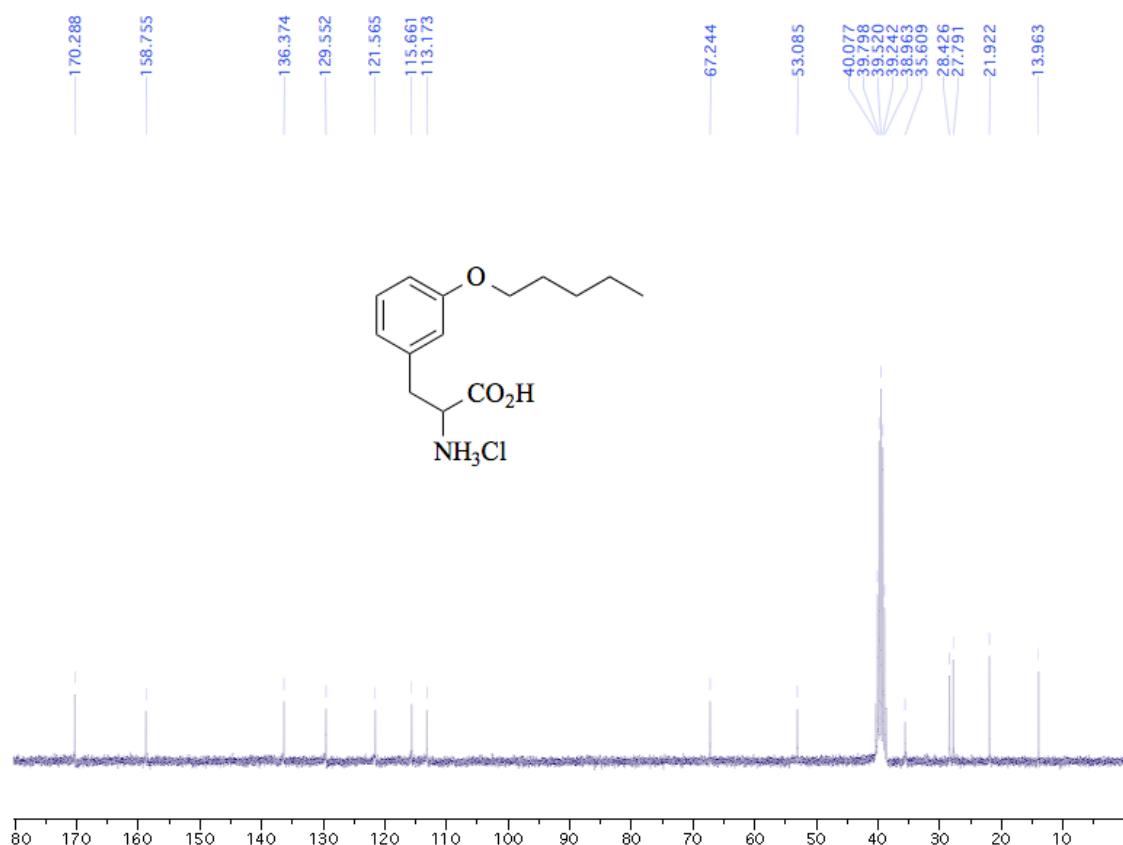


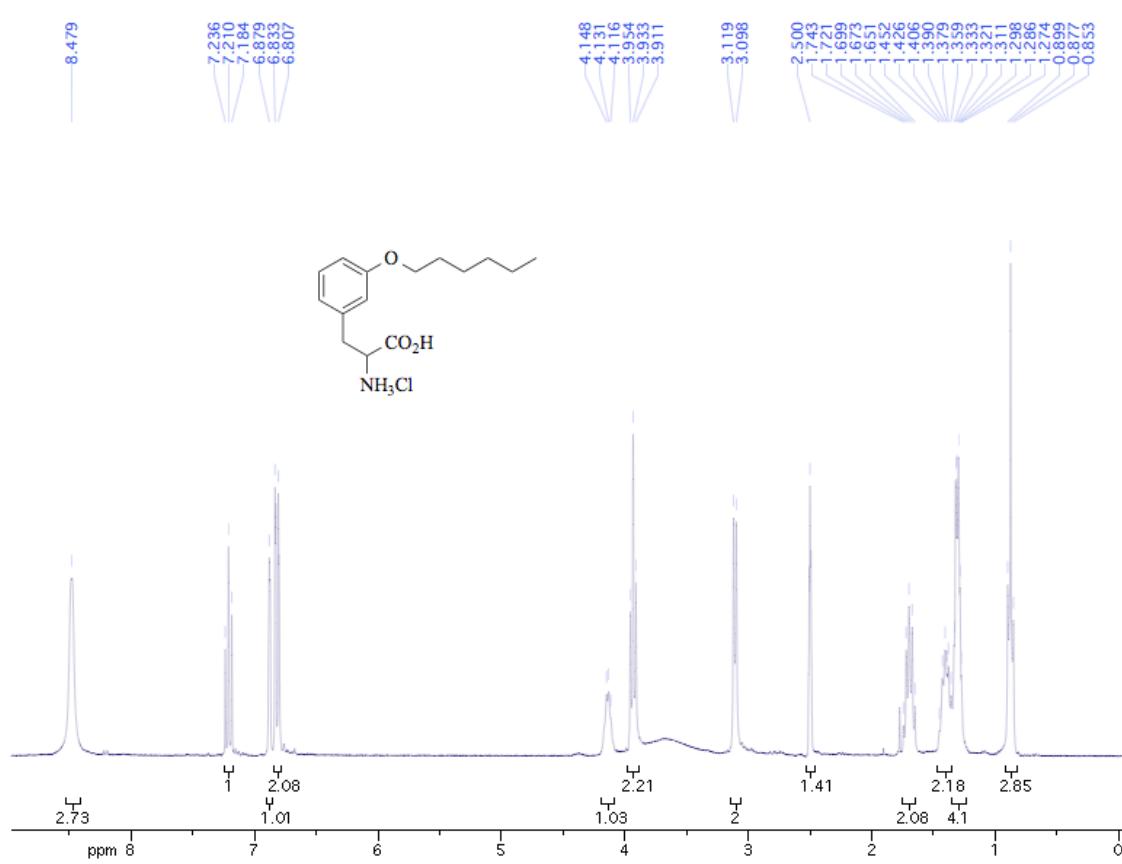


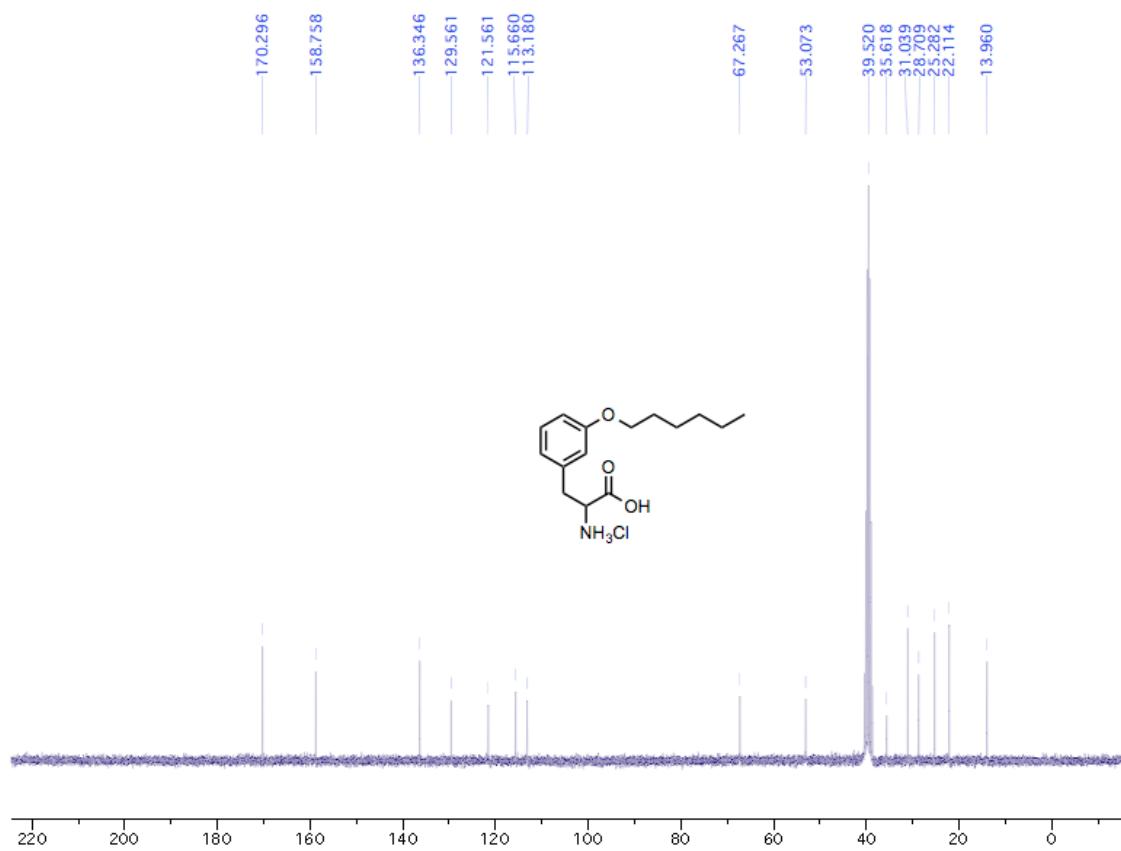


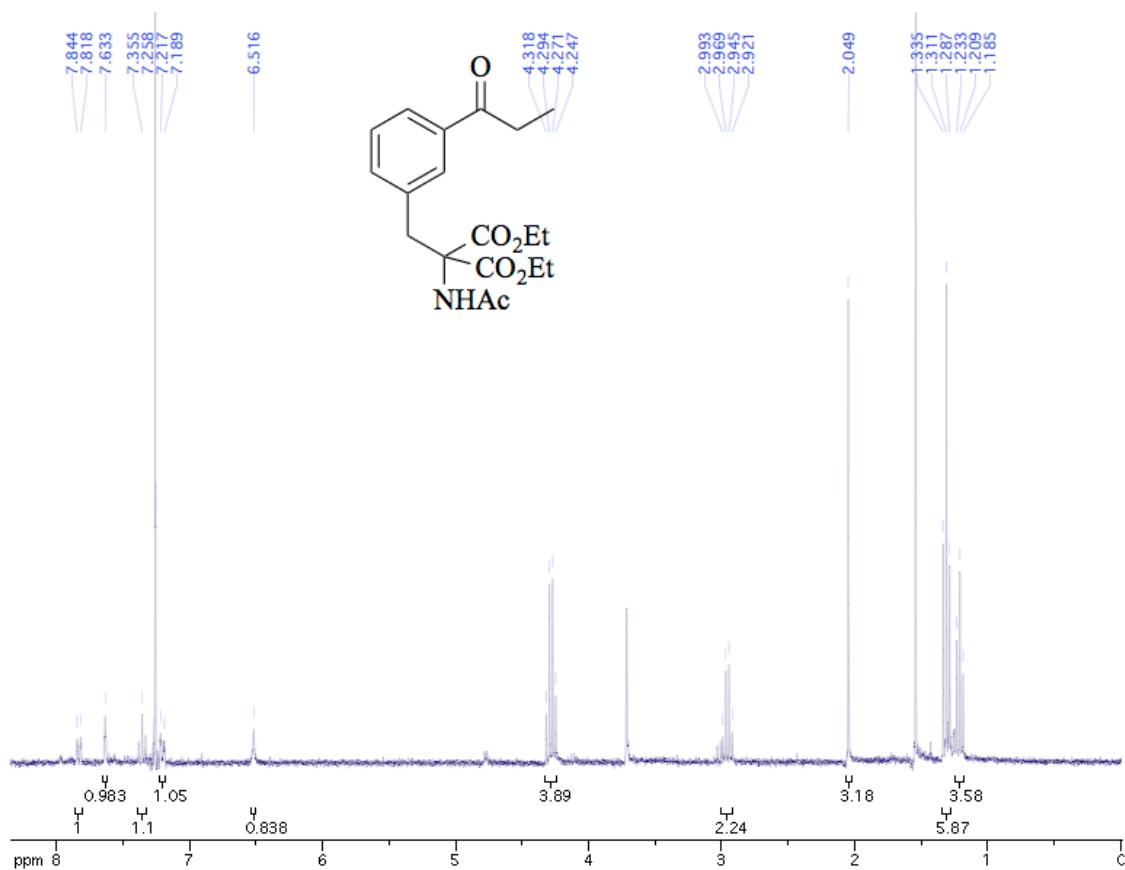


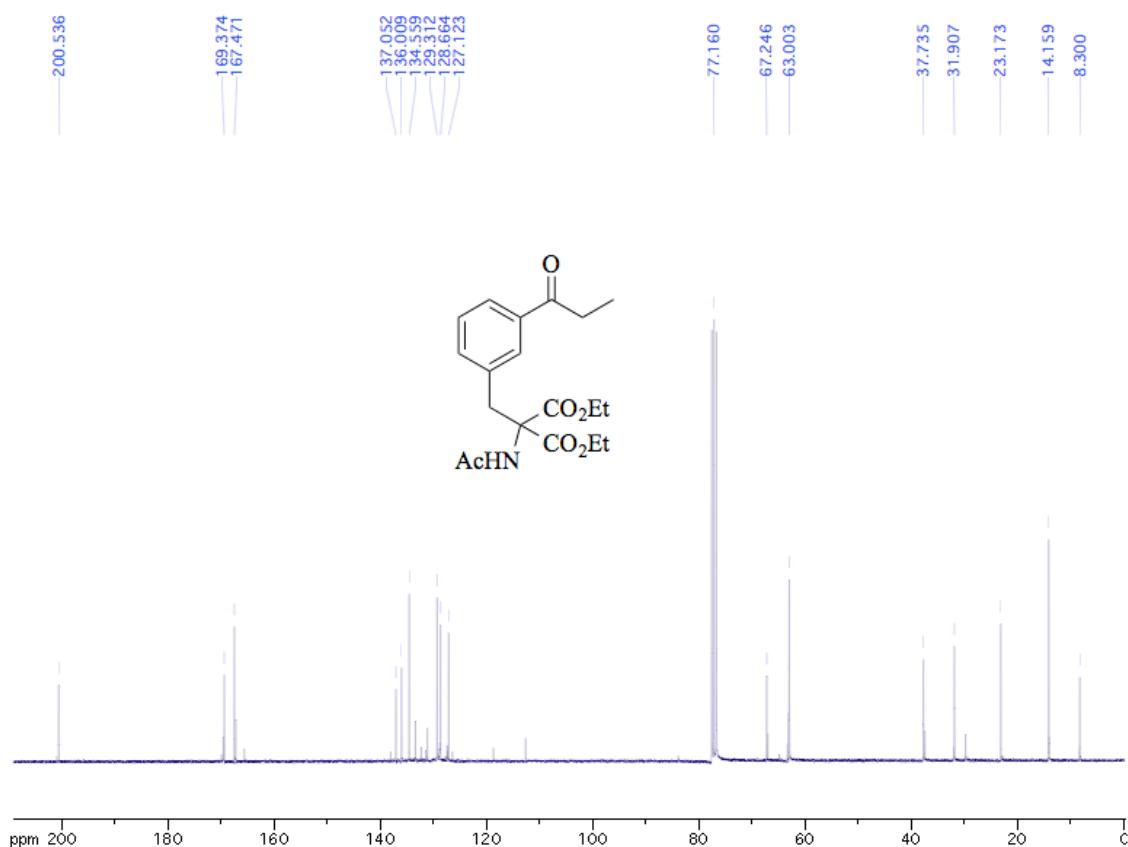


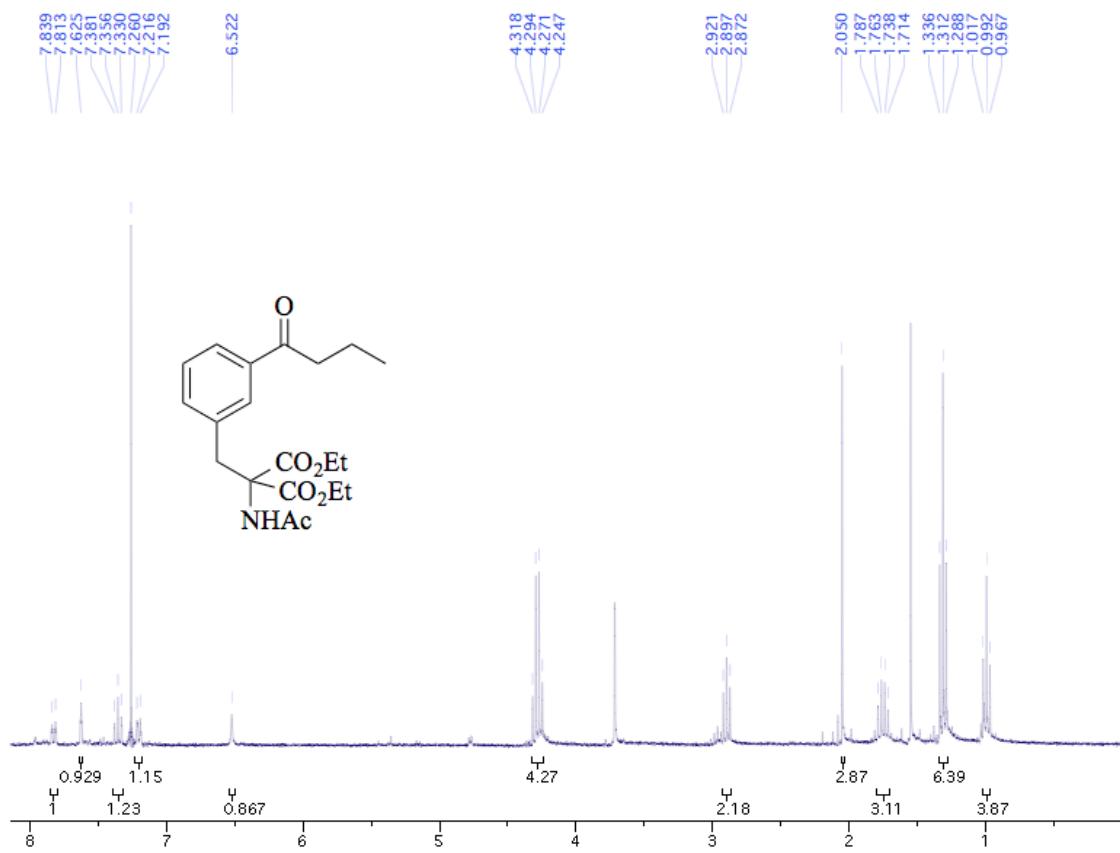


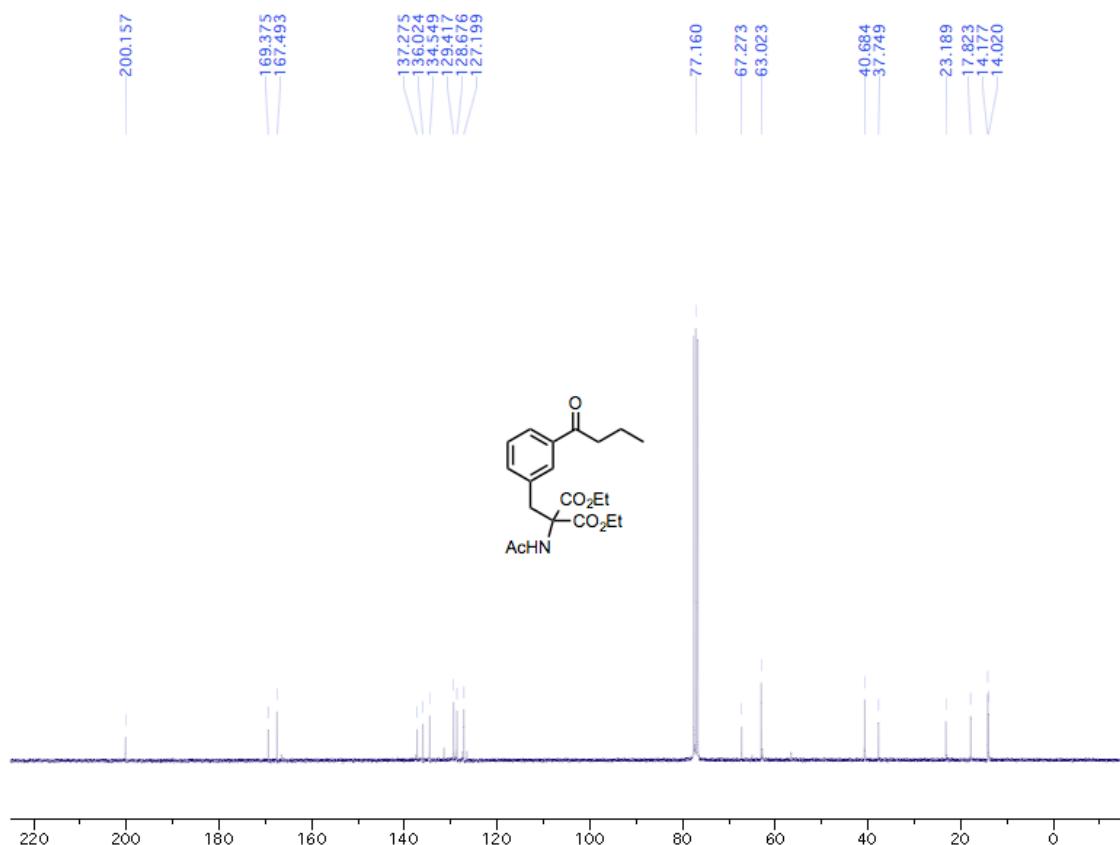


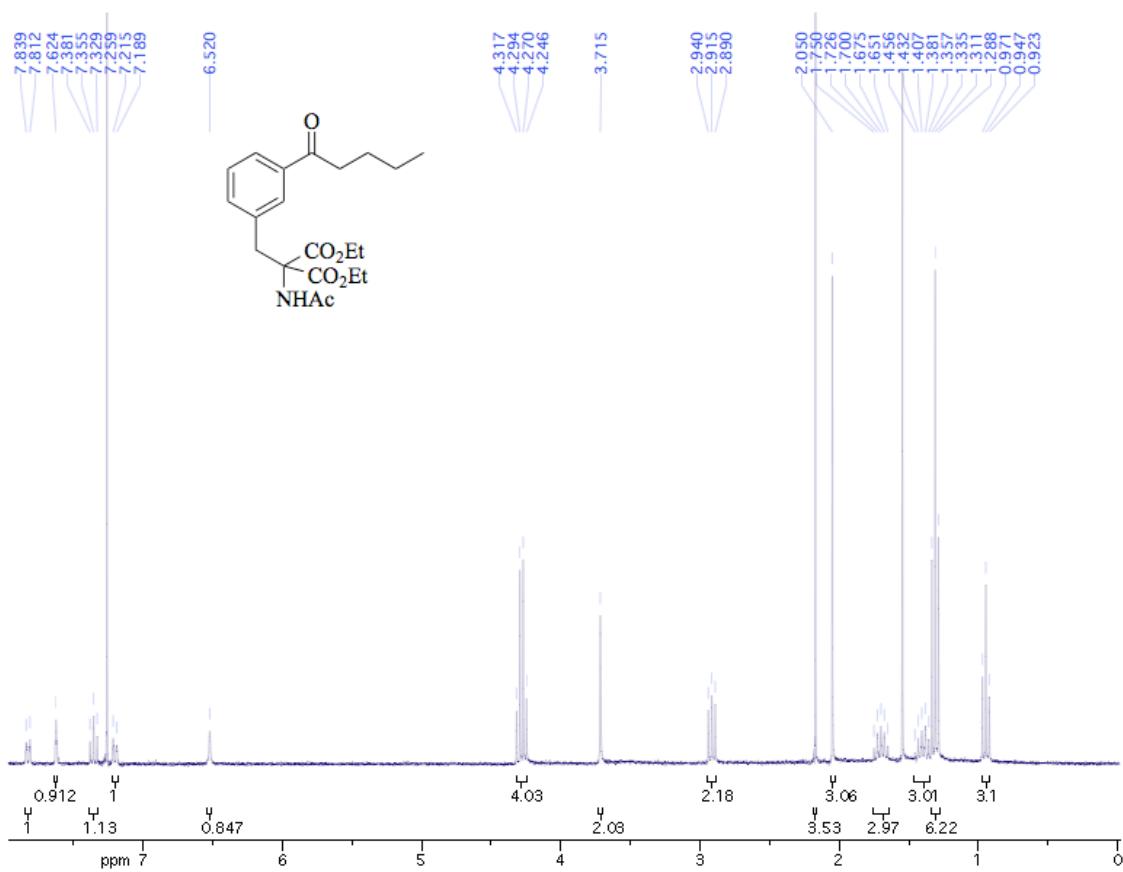


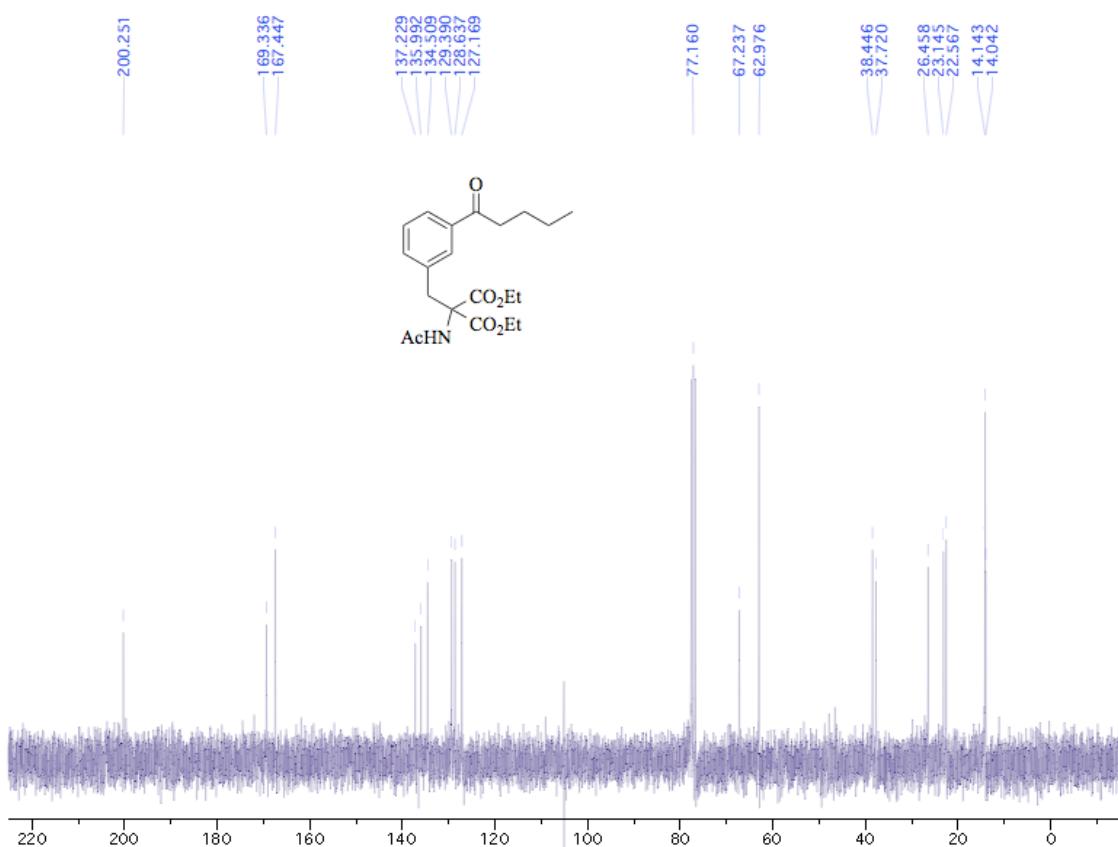


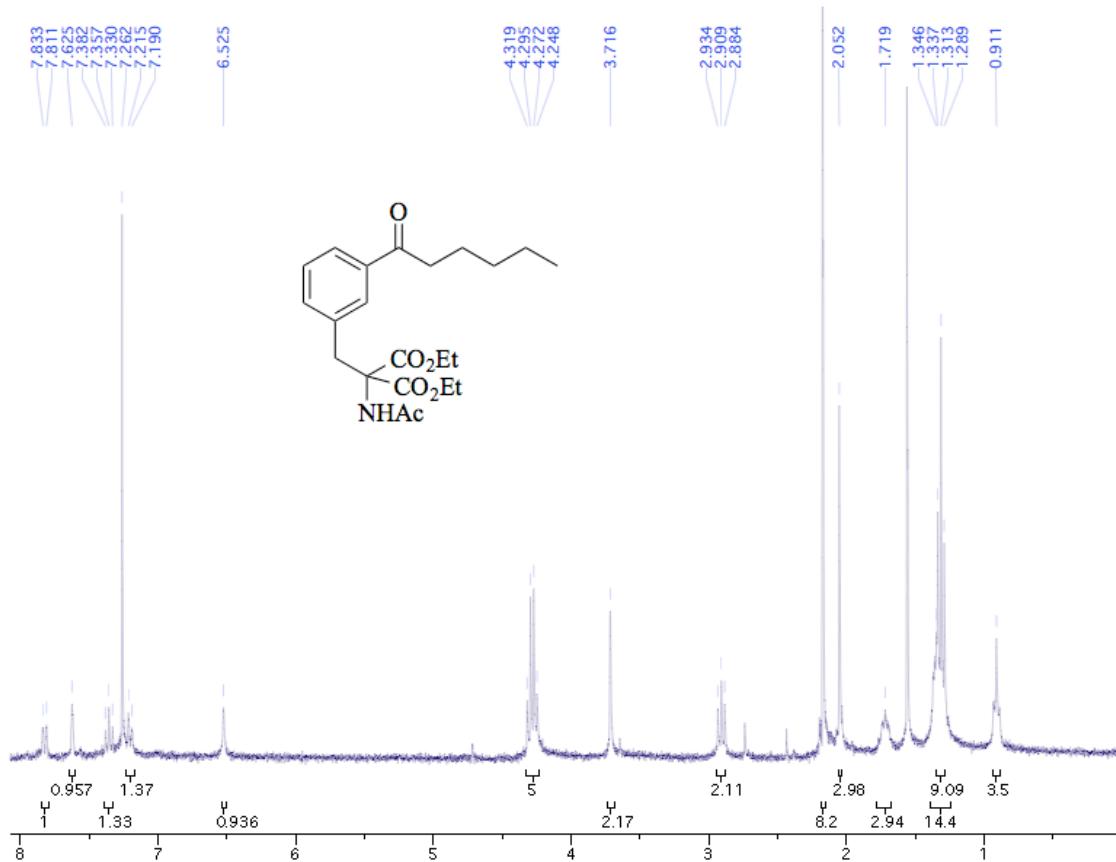


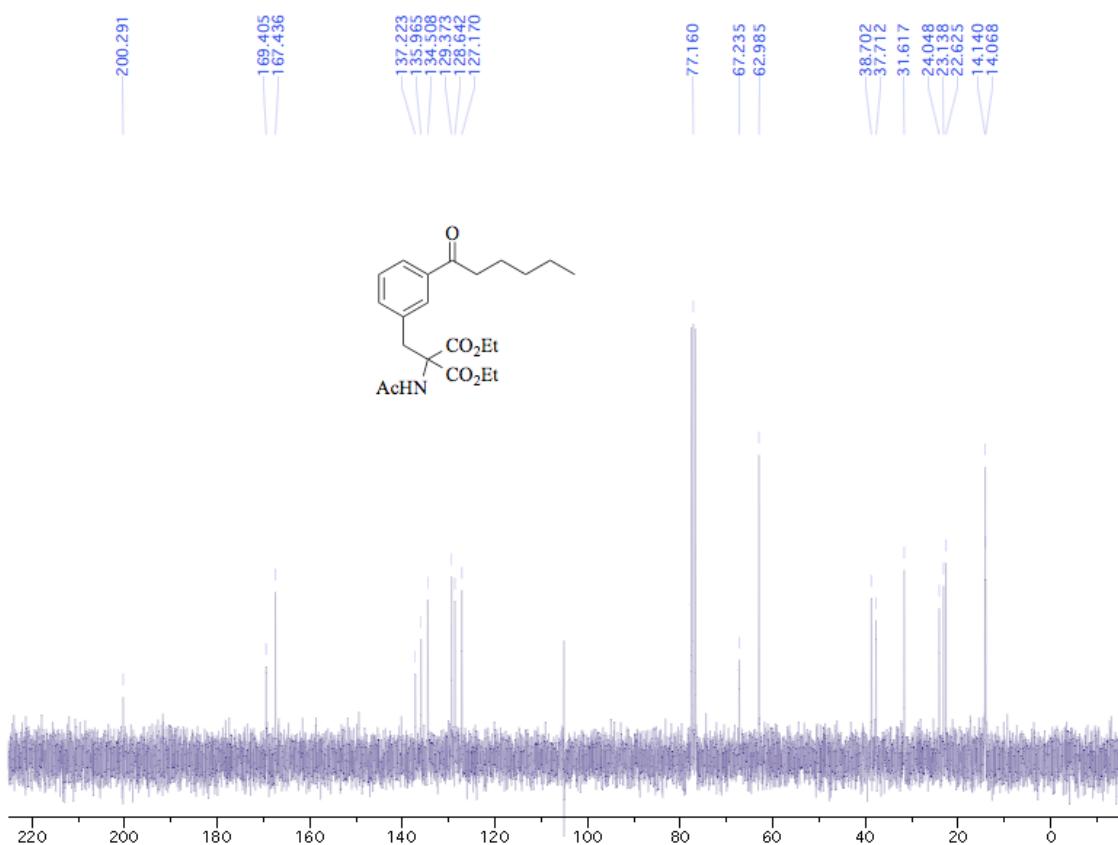


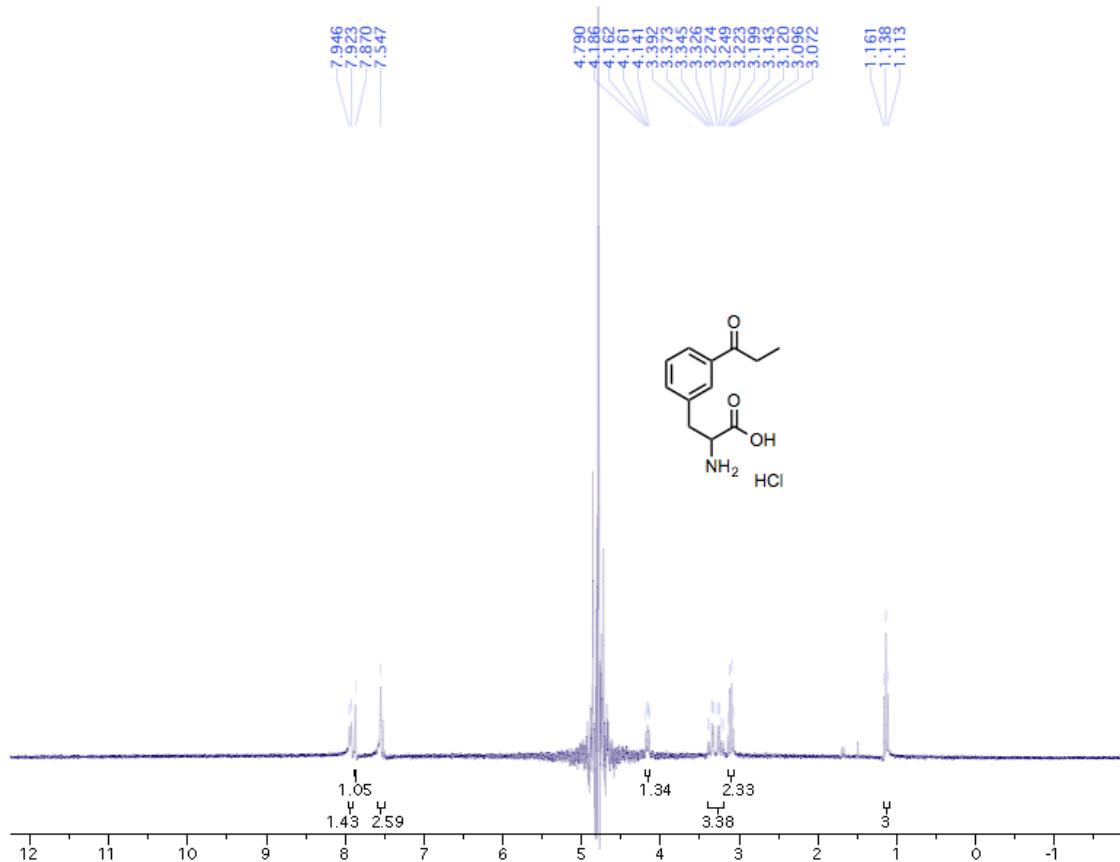


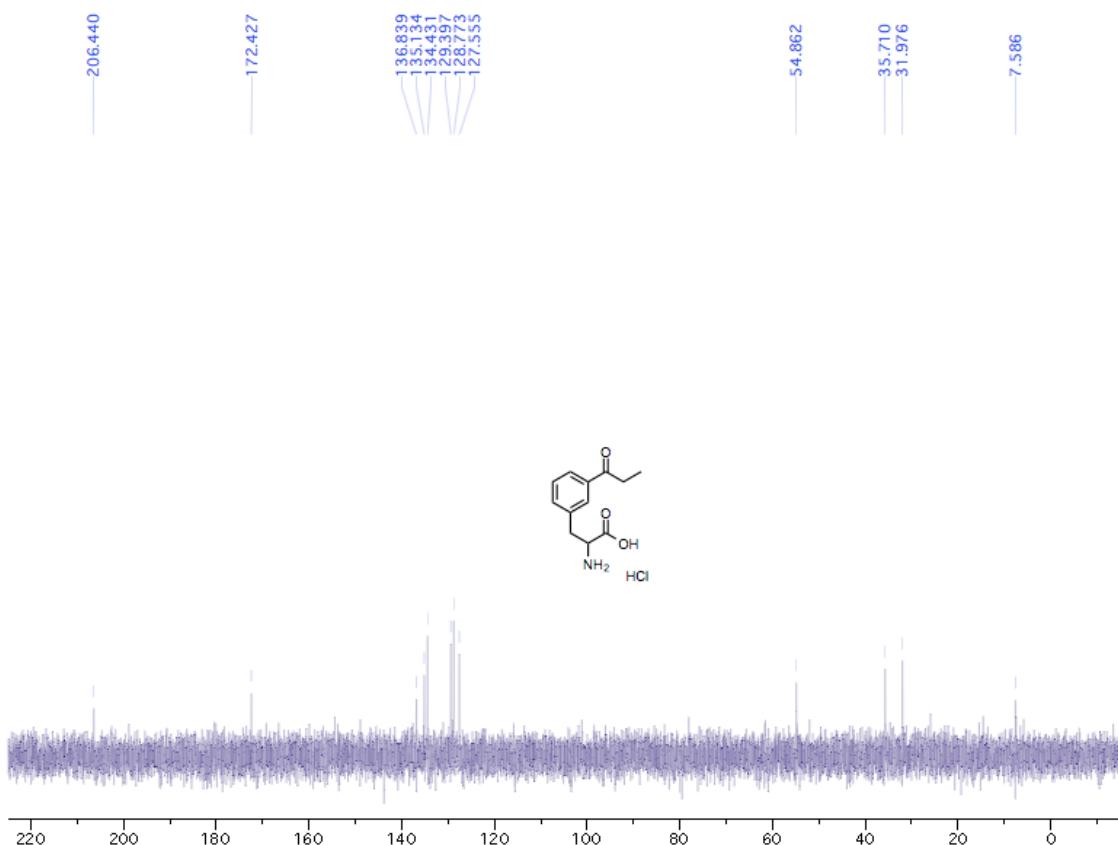


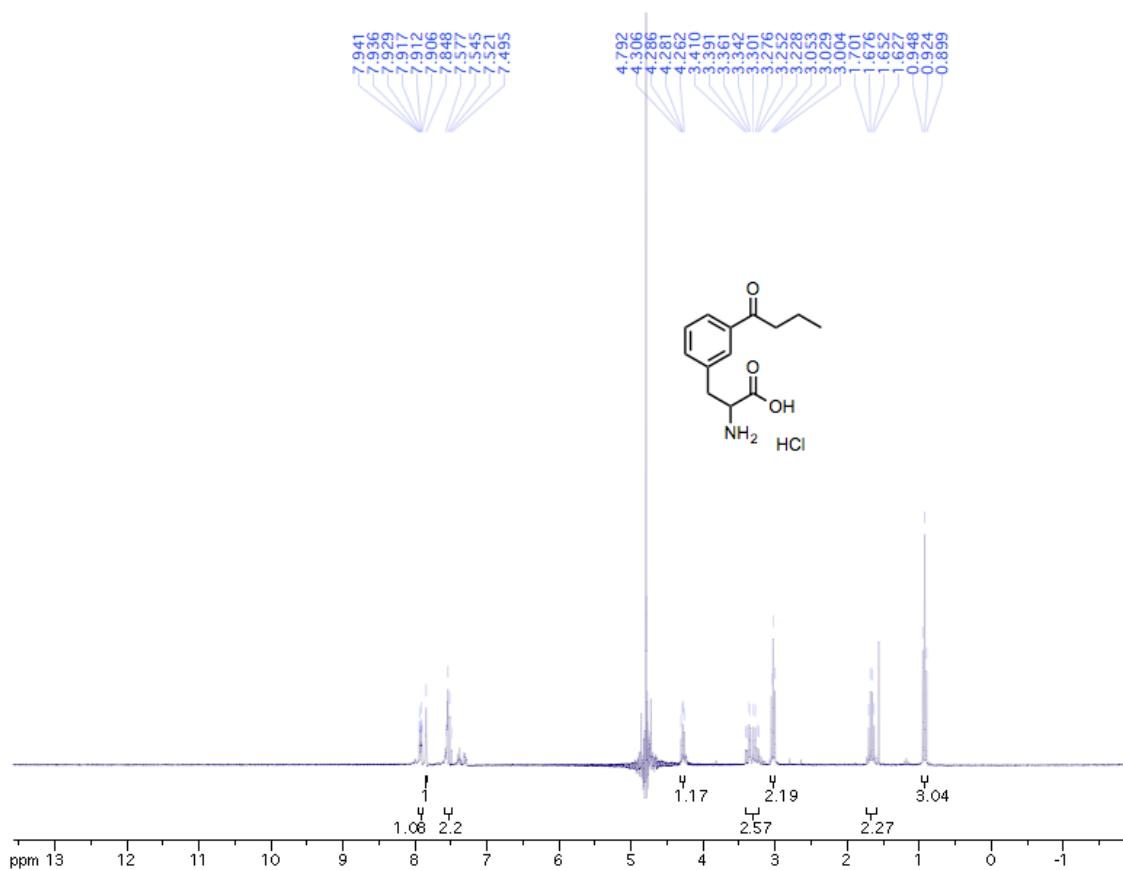


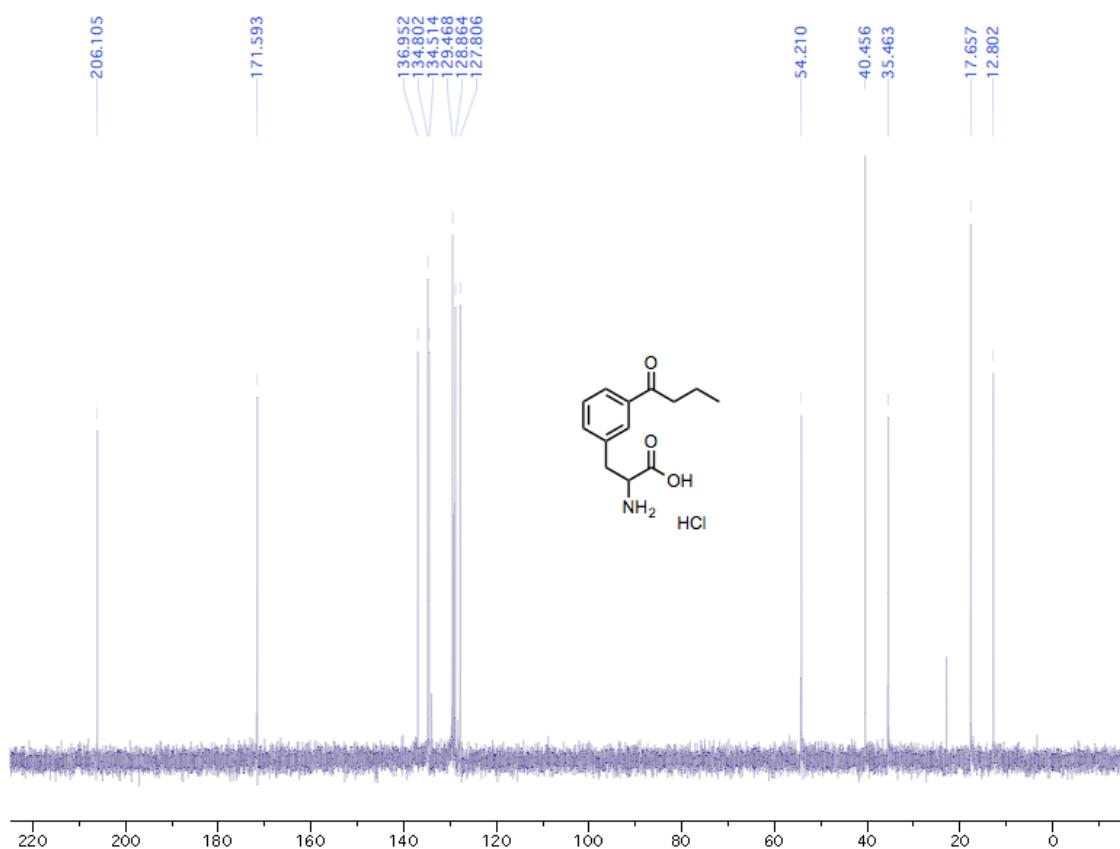


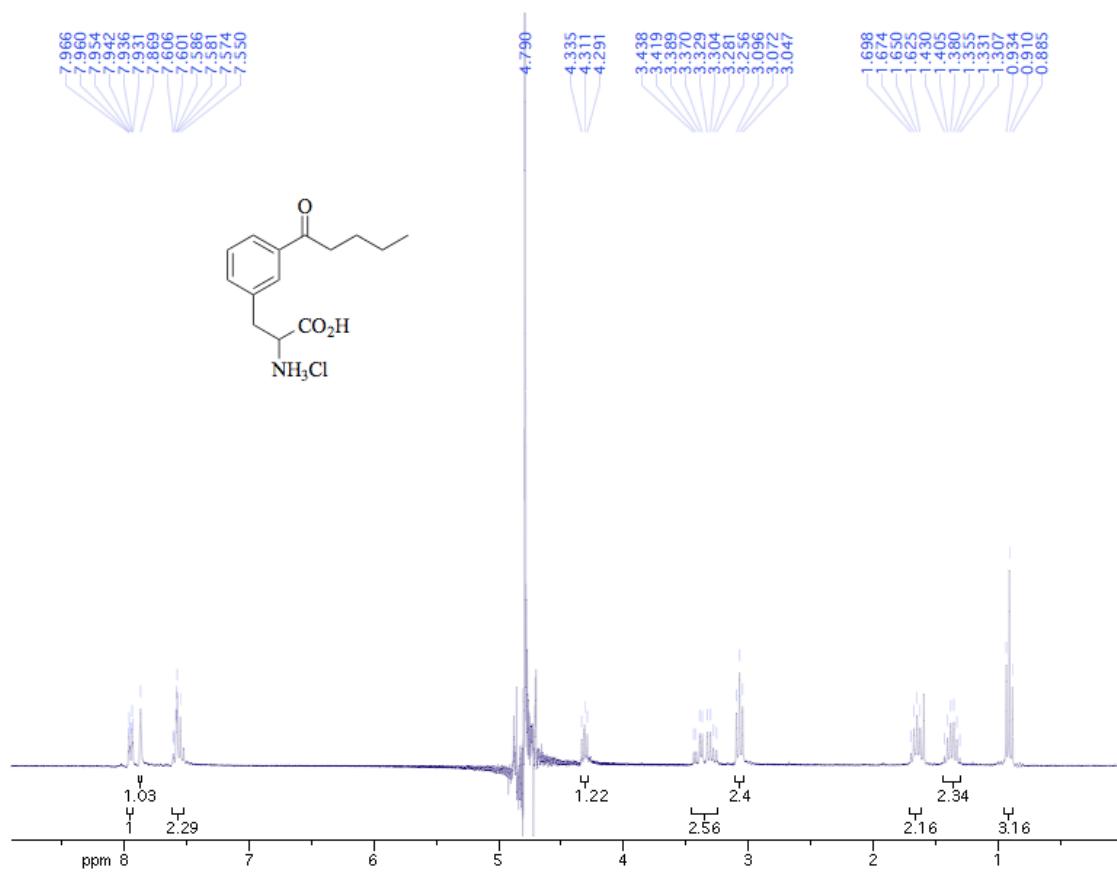


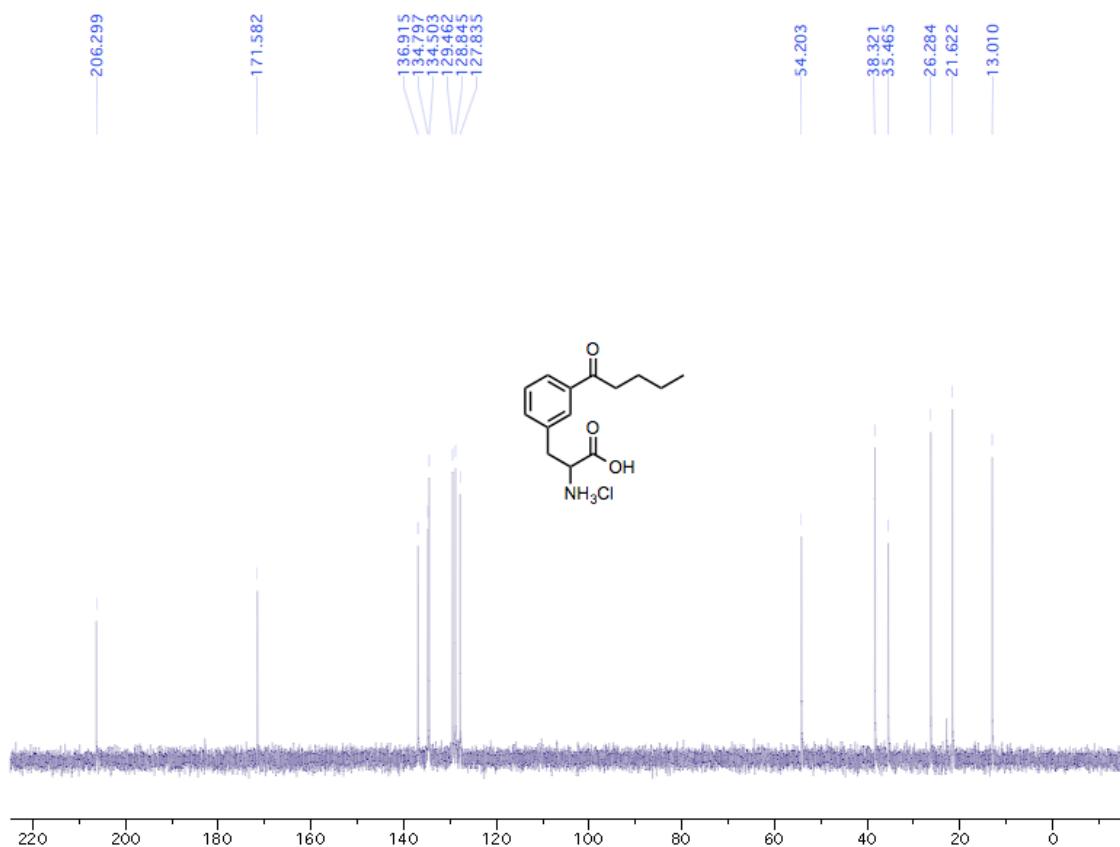


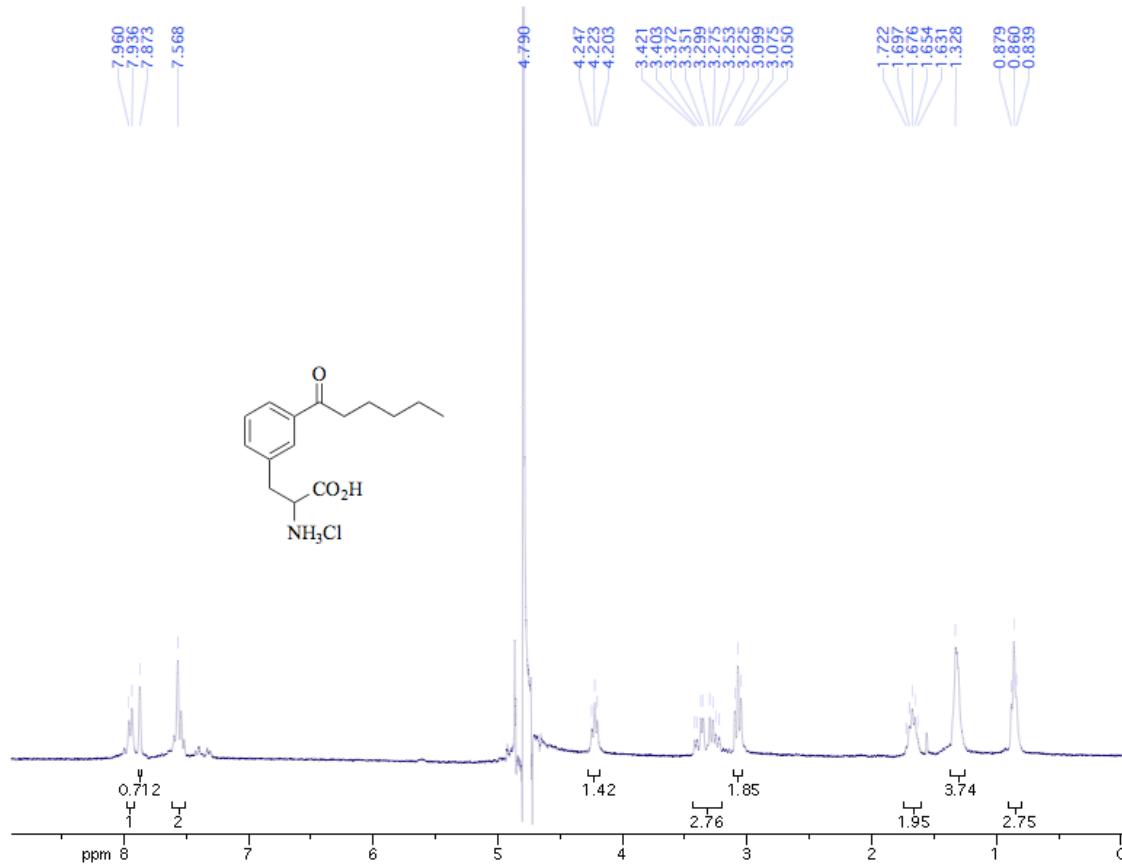


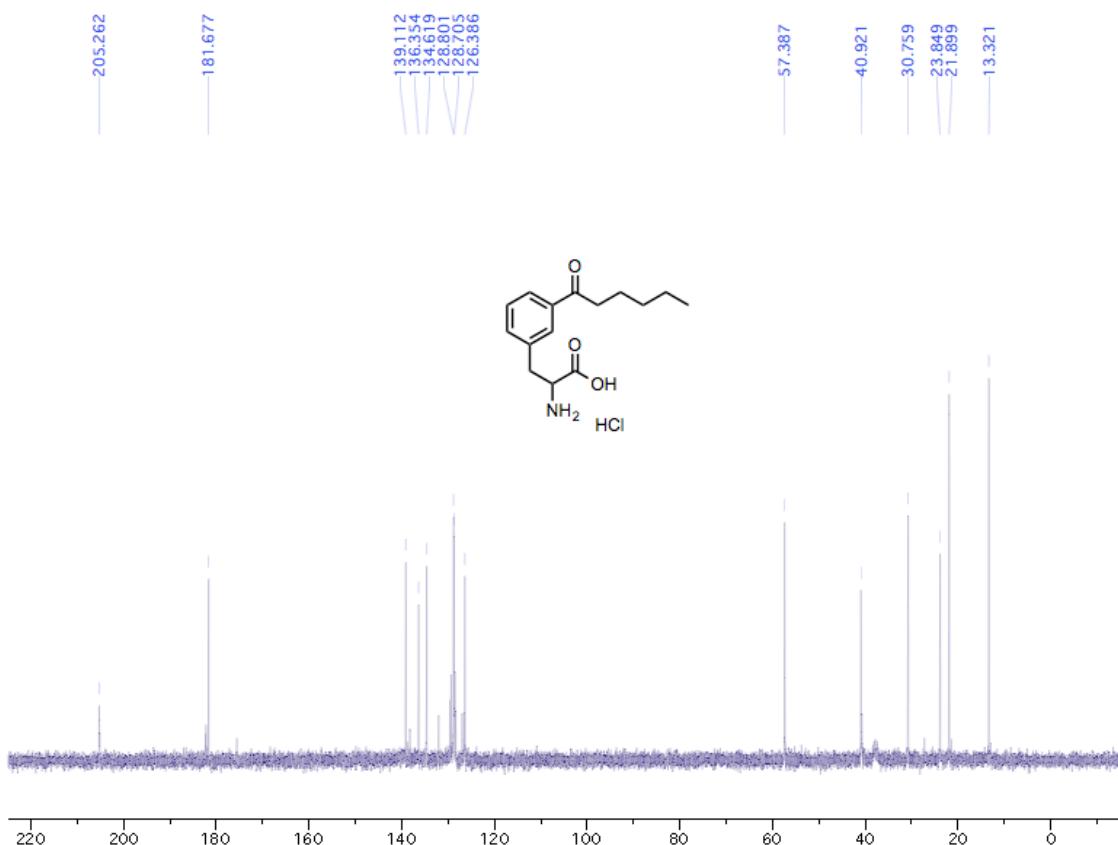












Supplementary Material

The genetic incorporation of thirteen novel non-canonical amino acids

Alfred Tuley, Yane-Shih Wang, Xinqiang Fang, Yadagiri Kurra, Yohannes Rezenom,
and Wenshe Liu

Department of Chemistry, Texas A&M University, College Station, TX 77843
E-mail: wliu@chem.tamu.edu

Table of Contents

1. Superfolder Green Fluorescent Protein Expression	S3
2. Protein Labeling	S3
3. ESI-MS Analysis of Intact Proteins	S3
4. Organic Synthesis	S3
5. sfGFP Protein Sequence.....	S12
6. NMR Data.....	S13

1. Superfolder Green Fluorescent Protein Expression

The constructs used to incorporate compounds **1-4** in this study and corresponding protein purification and characterization are identical to previous reports.^{1,2} For compounds **5-15**, BL21(DE3) *E. coli* cells containing pEVOL-PylT-pylRS-PylRSN346A/C348A (Cm^r) and pET-PylT-sfGFP-S2TAG' (Amp^r) vectors were grown in 500 mL of LB media with ampicillin (100 $\mu\text{g}/\text{mL}$) and chloramphenicol (34 $\mu\text{g}/\text{mL}$) until the OD_{600} reached 1.0-1.3, at which point the cells were pelleted at 4,000 r.p.m. for 20 min, then washed and resuspended in 30 mL H_2O . The resuspended cells were added as 5 mL aliquots to 45 mL of minimal media (33.7 mM Na_2HPO_4 , 22 mM KH_2PO_4 , 8.6 mM NaCl, 9.4 mM NH_4Cl , 1 mM MgSO_4 , 0.3 mM CaCl_2 , 1% glycerol) supplemented with 2 mM non-canonical amino acid (NAA), 1 mM IPTG, and 0.2% arabinose, at which point protein expression was allowed to occur for 12 h. The cells were then pelleted at 4,000 r.p.m. for 20 min, resuspended in lysis buffer (50 mM NaH_2PO_4 , 300 mM NaCl, pH 8) and lysed via sonication. The crude lysate was then centrifuged at 10,000 r.p.m. for 1 hour, and the supernatant was treated with imidazole to a final concentration of 10 mM. Next, the supernatant was subsequently incubated with Ni^{2+} -NTA resin for 1 hour at 4 °C. The resin was washed with lysis buffer containing 10 mM imidazole (3x column volume) and 20 mM imidazole (3x column volume), then eluted with elution buffer (50 mM NaH_2PO_4 , 300 mM NaCl, 500 mM imidazole, pH 8). The protein was dialyzed against 10 mM Tris buffer, and if necessary, the proteins were concentrated using Amicon Ultracel-10k centrifugal filter units. Purity of the proteins was confirmed via 15% SDS-PAGE and ESI-MS analysis. Yields were determined using a commercially available BCA protein assay kit (Thermo Scientific).

2. sfGFP-2 Protein Labeling

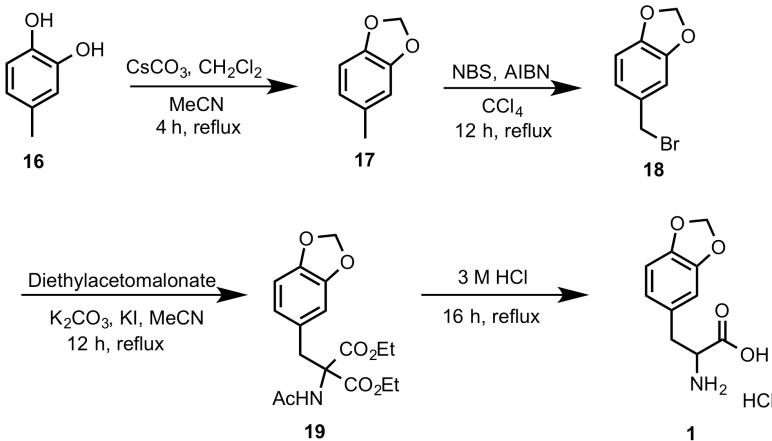
To 16 μl of 1X Phosphate buffered saline (PBS) was added 2 μl sfGFP-**2** (3 mg/mL, 1 μM) and 2 μl Dibenzylcyclooctyne (DBCO) dye (25 mM, DMSO stock solution)³ and was allowed to react at room temperature overnight. The protein was then precipitated with 180 μl of methanol and left at -20 °C for 1 h. Next, the precipitated protein was centrifuged for 5 min, 14,000 r.p.m. and washed twice with 100% methanol. Finally, the residue was resuspended in 20 μl H_2O and subjected to SDS-PAGE analysis. The control experiment was identical except 4 μl sfGFP-**3** (5 mg/mL, 0.9 μM) was added to 14 μl PBS, followed by 2 μl DBCO dye.

3. ESI-MS Analysis of Intact Proteins

Nanoelectrospray ionization in positive mode was performed using an Applied Biosystems QSTAR Pulsar (Concord, ON, Canada) equipped with a nanoelectrospray ion source. Solution was flowed at 700 nL/min through a 50 μm ID fused-silica capillary that was tapered at the tip. Electrospray needle voltage was held at 2100 V.

4. Organic Synthesis

Reactions were carried out using oven-dried glassware and under an atmosphere of argon, where appropriate. Reagents were purchased and used without further purification. NMR spectra were obtained with Inova 300 and Mercury 300 MHz instruments.



4.1 2-amino-3-(benzo[d][1,3]dioxol-5-yl)propanoic acid hydrochloride (1)

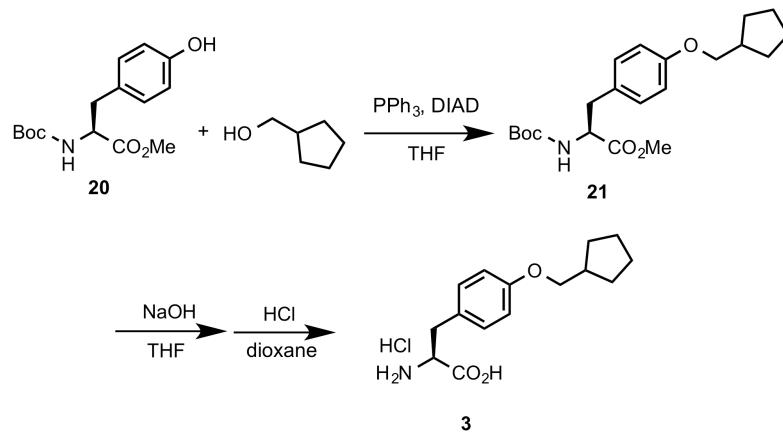
To a solution of catechol **16** (5 g, 8.06 mmol) in MeCN (25 mL) was added Cs_2CO_3 (3.93g 12.09 mmol), followed by CH_2Cl_2 (5.21 mL, 8.06 mmol). The resulting mixture was refluxed for 4 hours. After cooling to room temperature, the reaction mixture was concentrated and applied to column chromatography, yielding the acetal **17** in 40% yield (2.12 g). ^1H NMR (300 MHz, CDCl_3) δ 6.71(m, 3H), 5.93(s, 2H), 2.27(s, 3H).

The compound **17** (2.0 g, 14.7 mmol) was dissolved in CCl_4 (15 mL), followed by addition of NBS (3.12 g, 17.62 mmol) and AIBN (0.337 g, 2.05 mmol). The mixture was heated to reflux for 12 h under the protection of argon. After cooling, the precipitate was removed via filtration and the filtrate was concentrated and dried under vacuum to afford the known compound **18**,⁵ which was used directly in the next step without further purification.

To a solution of compound **18** (3.5 g, 16.27 mmol) in anhydrous acetonitrile (15 mL) was added diethyl 2-acetamidomalonate (3.88 g, 17.88 mmol), K_2CO_3 (4.49 g, 32.53 mmol) and KI (1.0 g, 15.38 mmol). The resulting mixture was heated to reflux for 12 h under the protection of argon, then cooled to room temperature. The solid was filtered, the solvent was removed under reduced pressure, and the residue was purified via column chromatograph with hexanes/ethyl acetate (3:1 v/v) as eluent to give the pure product **19** (60% yield, 3.42 g). ^1H NMR (300 MHz, CDCl_3) δ 6.69 (d, $J = 8.2\text{Hz}$, 1 H), 6.57(d, $J = 7.9\text{ Hz}$, 1 H), 6.47 (s, 1H), 5.92 (s, 2H), 4.30-4.22 (m, 4H), 3.56 (s, 2H), 2.04 (s, 3H), 1.32-1.23 (m, 6H).

A suspension of compound **19** (3.0 g, 8.54 mmol) in 3 M HCl (91.16 mL, 32 eq.) was heated to reflux for 16 h before cooling to room temperature. Water was evaporated under reduced pressure and the solid was collected. The solid was washed with Et_2O (10 mL \times 3) and then dried under vacuum to afford compound **1** as a solid (68% yield, 1.42 g). ^1H NMR (300 MHz, CD_3OD) δ 6.78-6.66 (m, 3H), 5.85 (s, 2H), 4.14 (t, $J = 6.6\text{ Hz}$, 1 H), 3.09 (ABq, $J = 15.3, 6.6\text{ Hz}$, 2 H). ^{13}C NMR (75 MHz, CD_3OD) δ 169.8, 148.1, 147.3, 127.6, 122.6, 109.1, 108.2, 101.1, 53.9, 39.5.

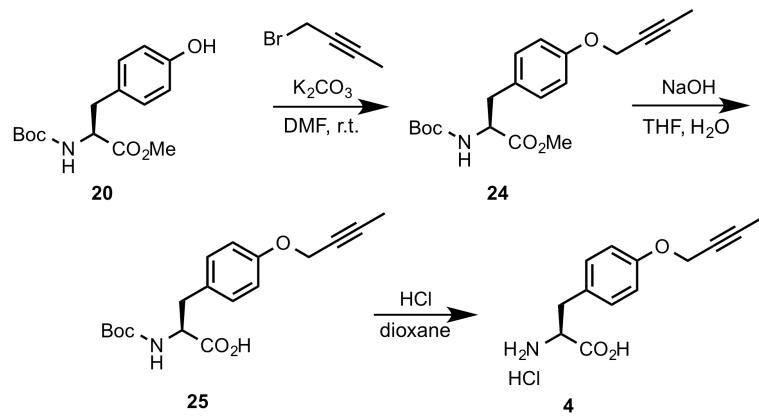
4.2 Alkylation of Tyrosine



4.2.1 (*S*)-2-amino-3-(4-(cyclopentylmethoxy)phenyl)propanoic acid hydrochloride (3)

N-Boc-*o*-methyl-L-tyrosine (**20**) (1.3 g, 4.4 mmol), cyclopentylmethanol (0.57 mL, 5.28 mmol), and triphenylphosphine (1.73 g, 6.6 mmol) were added to a round bottom flask, then the atmosphere evacuated and replaced with argon. THF (10 mL) was then added and the reaction mixture was cooled to 0 °C, at which point DEAD (1.3 mL, 6.6 mmol) was added dropwise and stirred overnight. Upon completion, the reaction mixture was concentrated and purified via column chromatography (gradient elution, 25% to 50% ethyl acetate/hexanes) to give the product **21** in 78% yield (1.3 g).

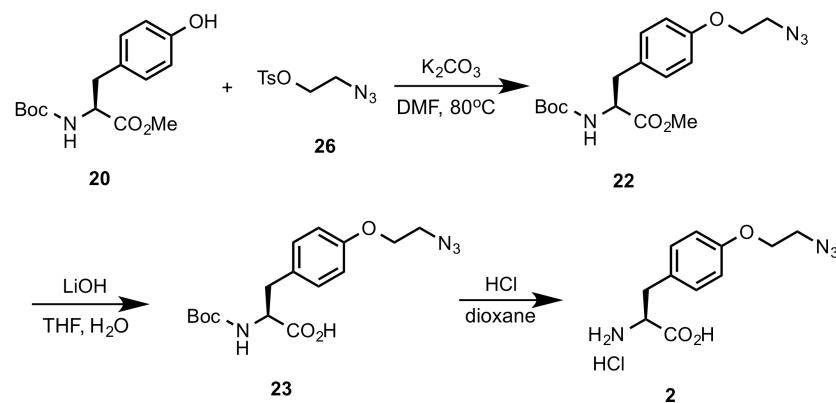
This compound was deprotected using the previously described procedure,¹ treatment with 5 mL 1 M NaOH/THF for two hours, followed by Boc deprotection in 5 mL 4 M HCl/dioxane to give the free amino **3** acid in 97% yield (898 mg), two steps. ¹H NMR (300 MHz, CD₃OD) δ 7.16 (d, *J* = 8.4 Hz, 2 H), 6.87 (d, *J* = 8.7, 6.9 Hz, 2 H), 4.15 (t, *J* = 5.4 Hz, 1 H), 3.80 (d, *J* = 6.9 Hz, 2 H), 3.14 (dd, *J* = 7.8, 5.4 Hz, 2 H), 2.31 (m, 1 H), 1.82 (m, 2 H), 1.61 (m, 4 H), 1.36 (m, 2 H). ¹³C NMR (75 MHz, CD₃OD) δ 169.9, 158.9, 130.0, 125.6, 114.6, 71.8, 53.8, 38.9, 35.0, 28.9, 24.9.



4.2.2 (*S*)-2-amino-3-(4-(but-2-yn-1-yloxy)phenyl)propanoic acid hydrochloride (4)

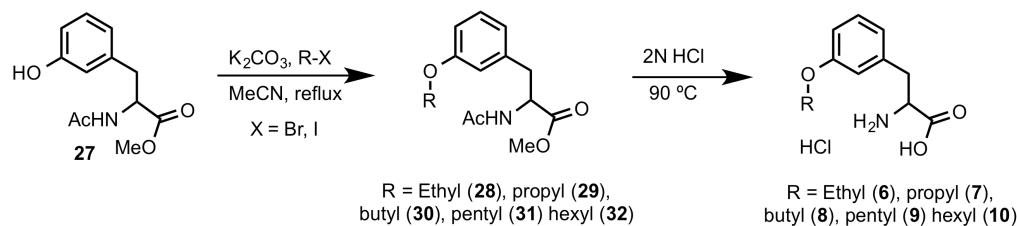
N-Boc-*o*-methyl-L-tyrosine (1 g, 3.39 mmol) was treated with 3-bromo-2-propyne (0.36 mL, 4.06 mmol) and potassium carbonate (1.4 g, 10.2 mmol) in DMF and left to

react overnight at room temperature. The reaction mixture was then diluted with ethyl acetate and washed with 3 M HCl. The organic layers were then dried, concentrated, and purified via flash chromatography (gradient elution, 25% to 50% ethyl acetate/hexanes) to afford the protected azide **24** in 30% yield (370 mg). Deprotection using the protocol listed above gave the title compound **4** in 44% yield (274 mg). ¹H NMR (300 MHz, CD₃OD) δ 7.10 (d, *J* = 8.7 Hz, 2 H), 6.83 (d, *J* = 8.7 Hz, 2 H), 4.53 (m, 2 H), 4.08 (q, *J* = 7.5 Hz, 1 H), 3.02 (dd, *J* = 14.7, 7.5 Hz, 2 H), 1.68 (s, 3 H). ¹³C NMR (75 MHz, CD₃OD) δ 168.1, 155.8, 128.4, 124.6, 113.3, 81.0, 72.0, 53.9, 52.0, 33.2.



4.2.3 (*S*)-2-amino-3-(4-(2-azidoethoxy)phenyl)propanoic acid hydrochloride (2)

N-Boc-*o*-methyl-L-tyrosine **20** (1.8 g, 6.09 mmol) was treated with compound **26** (1.8 g, 7.31 mmol) and potassium carbonate (3.3 g, 24.4 mmol) in DMF (10 mL) and the reaction was heated to 80 °C. When the reaction was complete based on TLC, the reaction mixture was worked up as usual and purified via flash chromatography (gradient elution, 25% to 50% ethyl acetate/hexanes) to yield compound **22**. Deprotection was achieved using 20 mL 1 M LiOH, followed by treatment with 5 mL 4 M HCl/Dioxane to afford free amino acid **2** (50%, two steps, 873 mg). ¹H NMR (300 MHz, D₂O) δ 7.16 (d, *J* = 8.4 Hz, 2 H), 6.88 (d, *J* = 8.4 Hz, 2 H), 4.11 (m, 3 H), 3.50 (m, 2 H), 3.11 (m, *J* = 7.2, 6.9, 5.4 Hz, 2 H). ¹³C NMR (75 MHz, D₂O) δ 169.8, 158.1, 130.2, 126.4, 114.7, 67.0, 53.8, 49.8, 35.0.



4.3 Representative Procedure for *m*-Tyrosine Alkylation

4.3.1 Methyl 2-acetamido-3-(3-ethoxyphenyl)propanoate (28)

To a solution of *m*-tyrosine⁶ **27** (1.5 g, 6.32 mmol) in DMF (12.64 mL) was added K₂CO₃ (2.84 g, 20.55 mmol), followed by ethyl iodide (0.76 mL, 9.48 mmol). The

reaction was stirred at room temperature for 18 h, then quenched with 16 mL H₂O and 79 mL ethyl acetate. The organic layer was extracted three times with 40 mL H₂O, then dried with sodium sulfate and concentrated. Column chromatography (gradient elution, 50 to 75% ethyl acetate/hexanes) afforded the compound as a yellow, crystalline solid (72% yield, 1.2 g). ¹H NMR (300 MHz, CDCl₃) δ 7.17 (t, *J* = 7.8 Hz, 1H), 6.76 (dd, *J* = 8.1, 5.7 Hz, 1H), 6.66 (m, 2H), 6.04 (d, *J* = 7.5 Hz, 1H), 4.85 (dt, *J* = 7.8, 5.7 Hz, 1H), 3.97 (q, *J* = 6.9 Hz, 2H), 3.71 (s, 3H), 3.07 (m, *J* = 8.1, 5.7 Hz, 2H), 1.97 (s, 3H), 1.38 (t, *J* = 6.9 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 172.2, 169.7, 159.2, 137.3, 129.6, 121.5, 115.6, 113.2, 63.4, 53.1, 52.4, 37.9, 23.3, 14.9.

4.3.2 Methyl 2-acetamido-3-(3-propoxymphenyl)propanoate (29)

Synthesized according to the general procedure with *m*-tyrosine (1.5 g, 6.32 mmol), DMF (12.64 mL), K₂CO₃ (2.84g, 20.55 mmol), and n-propyl bromide (0.86 mL, 9.48 mmol) to yield a yellow, crystalline solid, (68% yield, 1.2 g). ¹H NMR (300 MHz, CDCl₃) δ 7.18 (t, *J* = 8.1 Hz, 1H), 6.79 (m, *J* = 5.7 Hz, 1H), 6.66 (m, 2H), 5.88 (d, *J* = 7.8 Hz, 1H), 4.87 (dt, *J* = 7.8, 5.7 Hz, 1H), 3.88 (t, *J* = 6.6 Hz, 2H), 3.74 (s, 3H), 3.09 (m, *J* = 8.1, 5.7 Hz, 2H), 1.99 (s, 3H), 1.79 (sx, *J* = 7.2 Hz, 2H), 1.03 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 172.1, 169.7, 159.3, 137.3, 129.6, 121.4, 115.6, 113.2, 69.5, 53.1, 52.4, 37.9, 23.2, 22.6, 10.6.

4.3.3 Methyl 2-acetamido-3-(3-butoxyphenyl)propanoate (30)

Synthesized according to the general procedure with *m*-tyrosine (1.0 g, 4.21 mmol), DMF (8.42 mL), K₂CO₃ (1.89g, 13.69 mmol), and n-butyl iodide (0.72 mL, 6.32 mmol) to yield a yellow, crystalline solid (65% yield, 0.8 g). ¹H NMR (300 MHz, CDCl₃) δ 7.19 (t, *J* = 7.8 Hz, 1H), 6.77 (d, *J* = 7.2 Hz, 1H), 6.64 (m, 2H), 5.87 (d, *J* = 6.9 Hz, 1H), 4.87 (dt, *J* = 7.8, 5.4 Hz, 1H), 3.92 (t, *J* = 6.6 Hz, 2H), 3.74 (s, 3H), 3.09 (dd, *J* = 8.1, 5.7 Hz, 2H), 1.99 (s, 3H), 1.76 (p, *J* = 6.9, 6.6, 6.3 Hz, 2H), 1.46 (m, 2H), 0.97 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 172.2, 169.7, 159.4, 137.3, 129.6, 121.4, 115.6, 113.2, 67.7, 53.1, 52.4, 37.9, 31.4, 23.2, 19.3, 13.9.

4.3.4 Methyl 2-acetamido-3-(3-(pentyloxy)phenyl)propanoate (31)

Synthesized according to the general procedure with *m*-tyrosine (1.5 g, 6.32 mmol), DMF (12.64 mL), K₂CO₃ (2.84g, 20.55 mmol), and n-pentyl bromide (1.18 mL, 9.48 mmol) to yield a yellow, crystalline solid (72% yield, 1.4 g). ¹H NMR (300 MHz, CDCl₃) δ 7.19 (t, *J* = 8.1 Hz, 1H), 6.78 (m, 1H), 6.63 (m, 2H), 5.87 (d, *J* = 6.9 Hz, 1H), 4.82 (dt, *J* = 7.8, 5.7 Hz, 1H), 3.91 (t, *J* = 6.6 Hz, 2H), 3.74 (s, 3H), 3.10 (dq, *J* = 8.1, 6.0 Hz, 2H), 1.99 (s, 3H), 1.77 (p, *J* = 7.2, 6.9, 6.6 Hz, 2H), 1.39 (m, 4H), 0.93 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 172.1, 169.7, 159.2, 137.3, 129.5, 121.2, 115.5, 113.0, 67.8, 53.1, 52.3, 37.8, 28.9, 28.2, 23.0, 22.4, 14.0.

4.4.5 Methyl 2-acetamido-3-(3-(hexyloxy)phenyl)propanoate (32)

Synthesized according to the general procedure with *m*-tyrosine (1.0 g, 4.21 mmol), DMF (8.42 mL), K₂CO₃ (1.89g, 13.69 mmol), and 1-iodohexane (0.93 mL, 6.32 mmol) to yield a yellow, crystalline solid (44% yield, 0.6 g). ¹H NMR (300 MHz, CDCl₃) δ 7.18 (t, *J* = 8.1 Hz, 1H), 6.77 (dd, *J* = 8.1, 5.7 Hz, 1H), 6.63 (m, 2H), 5.88 (d, *J* = 7.2 Hz, 1H), 4.87 (dt, *J* = 7.8, 5.7 Hz, 1H), 3.91 (t, *J* = 6.6 Hz, 1H), 3.74 (s, 3H), 3.09 (m, *J*

= 8.1, 5.7 Hz, 2 H), 1.99 (s, 3H), 1.77 (p, J = 8.1, 6.6 Hz, 2 H), 1.44 (m, 4 H), 1.33 (m, 4H), 0.91 (t, J = 7.2 Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 172.2, 169.7, 159.4, 137.3, 129.6, 121.4, 115.6, 113.2, 68.0, 53.1, 52.4, 37.9, 31.7, 29.3, 25.8, 23.3, 22.7, 14.1.

4.5 Representative Procedure for *meta*-alkoxy Phenylalanine Deprotection:

4.5.1 2-amino-3-(3-ethoxyphenyl)propanoic acid hydrochloride (6)

A round bottomed flask charged with compound **28** (1.1 g, 4.15 mmol) was treated with 14.3 mL 2 M HCl, and the resulting suspension was heated to 100 °C, which was refluxed overnight until complete consumption of starting material as observed on TLC. The reaction mixture was then cooled to room temperature and concentrated *in vacuo*. If necessary, the compound was resubjected to the reaction conditions due to the persistence of methyl ester and/or N-acyl amide signals in ^1H and ^{13}C NMR. The title compound was obtained as a white solid in 45% yield (463 mg). ^1H NMR (300 MHz, D_2O) δ 7.35 (t, J = 7.8 Hz, 1 H), 6.93 (m, 3 H), 4.31 (t, J = 7.5 Hz 1 H), 4.11 (q, J = 6.9 Hz, 2 H), 3.24 (dq, J = 14.7, 5.7 Hz, 2 H), 1.37 (t, J = 6.9 Hz, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 171.4, 158.3, 135.6, 130.4, 122.1, 115.5, 114.1, 64.2, 54.0, 35.5, 13.8.

4.5.2 2-amino-3-(3-propoxypheNyl)propanoic acid hydrochloride (7)

Prepared according to the general procedure using compound **29** (1.0 g, 3.58 mmol) and 12.34 mL 2 M HCl. Obtained as a white solid in 43% yield (402 mg). ^1H NMR (300 MHz, D_2O) δ 7.34 (t, J = 7.5 Hz, 1 H), 6.90 (m, 3 H), 4.29 (t, J = 5.4 Hz, 1 H), 3.97 (t, J = 6.6 Hz), 3.21 (dd, J = 14.7, 5.7 Hz, 2 H), 1.73 (sx, J = 7.5, 7.2, 6.9, 6.6 Hz, 2 H), 0.96 (t, J = 7.5, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 171.3, 158.5, 135.5, 130.3, 122.0, 115.5, 114.1, 70.1, 53.9, 35.4, 21.7, 9.5.

4.5.3 2-amino-3-(3-butoxyphenyl)propanoic acid hydrochloride (8)

Prepared according to the general procedure using compound **30** (0.8 g, 2.73 mmol) in 9.4 mL 2 M HCl to obtain the title compound in 30% yield (222 mg). ^1H NMR (300 MHz, DMSO) δ 8.53 (s, 2 H), 7.21 (t, J = 7.8 Hz, 1 H), 6.89 (s, 1 H), 6.82 (d, J = 7.8 Hz, 2 H), 4.14 (t, J = 5.4 Hz, 1 H), 3.94 (t, J = 6.6 Hz, 2 H), 3.12 (d, J = 6.0 Hz, 2 H), 1.69 (p, J = 7.5, 6.9, 6.6 Hz, 2 H), 1.45 (sx, J = 7.5, 7.2 Hz, 2 H), 0.93 (t, J = 7.2 Hz, 3 H). ^{13}C NMR (75 MHz, DMSO) δ 170.2, 158.7, 136.4, 129.5, 121.5, 115.6, 113.1, 66.9, 53.1, 35.5, 30.8, 18.8, 13.7.

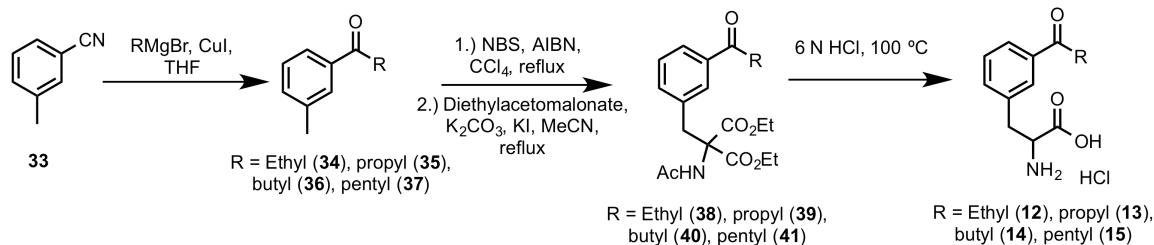
4.5.4 2-amino-3-(3-(pentyloxy)phenyl)propanoic acid hydrochloride (9)

Prepared according to the general procedure using compound **31** (1.0 g, 3.25 mmol) in 11.2 mL 2 M HCl to obtain the title compound in 61% yield (570 mg). ^1H NMR (300 MHz, DMSO) δ 8.49 (s, 2 H), 7.21 (t, J = 7.5 Hz, 1 H), 6.88 (s, 1 H), 6.82 (d, J = 8.1 Hz, 2 H), 4.14 (t, J = 6.0 Hz, 1 H), 3.94 (t, J = 6.6 Hz, 1 H), 3.11 (d, J = 6.0 Hz, 2 H), 1.71 (p, J = 6.9, 6.6, 6.3 Hz, 2 H), 1.36 (m, 4 H), 0.89 (t, J = 6.6 Hz, 3 H). ^{13}C NMR (75 MHz, DMSO) δ 170.2, 158.7, 136.4, 129.5, 121.6, 115.6, 113.1, 67.2, 53.0, 35.6, 28.4, 27.7, 21.9, 13.9.

4.5.5 2-amino-3-(3-(hexyloxy)phenyl)propanoic acid hydrochloride (10)

Prepared according to the general procedure using compound **32** (0.6 g, 1.87 mmol) in 6.44 mL 2 M HCl to yield the compound as a white solid in 65% yield (367

mg). ^1H NMR (300 MHz, DMSO) δ 8.48 (s, 3 H), 7.21 (t, J = 7.8 Hz, 1 H), 6.88 (s, 1 H), 6.82 (d, J = 7.8 Hz, 2 H), 4.13 (m, J = 5.1, 4.5 Hz, 1 H), 3.93 (t, J = 6.6 Hz, 2 H), 3.10 (d, J = 6.3 Hz, 2 H), 1.69 (p, J = 7.8, 6.6 Hz, 2 H), 1.41 (m, 2 H), 1.29 (m, 4 H), 0.88 (t, J = 6.6 Hz, 3 H). ^{13}C NMR (75 MHz, DMSO) δ 170.2, 158.7, 136.3, 129.5, 121.5, 115.6, 113.1, 67.2, 53.0, 35.6, 31.0, 28.7, 25.2, 22.1, 13.9.



4.6 Representative Procedure for Grignard Addition to *m*-Tolunitrile

4.6.1 1-(*m*-tolyl)propan-1-one (34)

Ethyl magnesium bromide (1 M/THF, 85.36 mL, 85.36 mmol) was added dropwise to a solution of *m*-tolunitrile (10.25 mL, 85.36 mmol) and copper (I) iodide (40.64 mg, 0.213 mmol) in anhydrous THF (170.72 mL) under argon. After stirring for 22 h, the reaction was quenched with approx. 5 mL 1 M HCl at 0 °C and stirred for 4 h, allowing the reaction to warm to room temperature. The resulting layers were separated and the organic layer was dried with MgSO_4 , filtered, and concentrated. The crude, yellow oil was purified via silica gel chromatography (gradient, 0 to 10% EtOAc/Hex) to afford the known compound⁷ 34 in 94% yield (9.4 g) as a yellow oil. ^1H NMR (300 MHz, CDCl_3) δ 7.76 (m, 1H), 7.38 (m, 3H), 2.99 (q, J = 7.2 Hz, 2H), 2.41 (s, 3H), 1.22 (t, J = 7.2 Hz, 3H).

4.6.2 1-(*m*-tolyl)butan-1-one (35)

Prepared according to the general procedure with Propylmagnesium bromide (2 M/THF, 42.68 mL, 85.36 mmol), *m*-tolunitrile (10.25 mL, 85.36 mmol), and CuI (40.64 mg, 0.213 mmol) to yield the known compound⁸ in 96 % yield (13.24 g). ^1H NMR (300 MHz, CDCl_3) δ 7.77 (s, 2H), 7.36 (s, 2H), 2.94 (t, J = 7.2 Hz, 2H), 2.41 (s 3H), 1.76 (sx, J = 7.2 Hz, 2H), 1.00 (t, J = 7.2 Hz, 3H).

4.6.3 1-(*m*-tolyl)pentan-1-one (36)

Prepared according to the general procedure with butylmagnesium chloride (2 M/THF, 42.68 mL, 85.36 mmol), *m*-tolunitrile (10.25 mL, 85.36 mmol), and CuI (40.64 mg, 0.213 mmol) to yield the known compound⁹ in 76% yield (11.43 g). ^1H NMR (300 MHz, CDCl_3) δ 7.75 (m, 2H), 7.34 (m, 2H), 2.95 (t, J = 7.2 Hz, 2H), 2.41 (s, 3H), 1.71 (p, J = 7.5 Hz, 2H), 1.41 (sx, J = 7.2 Hz, 2H), 0.95 (t, J = 7.5 Hz, 3H).

4.6.4 1-(*m*-tolyl)hexan-1-one (37)

Prepared according to the general procedure with Pentylmagnesium bromide (1 M/THF, 42.68 mL, 85.36 mmol), *m*-tolunitrile (10.25 mL, 85.36 mmol), and CuI (40.64 mg, 0.213 mmol) to yield the known compound¹⁰ in 94% yield (15.3 g). ^1H NMR (300 MHz, CDCl_3) δ 7.74 (m, 1H), 7.46 (s, 1H), 7.37 (m, 2H), 2.94 (t, J = 6.9 Hz, 2H), 2.41 (s, 3H), 1.73 (p, J = 7.2 Hz, 2H), 1.35 (m, 4H), 0.91 (t, J = 6.6 Hz, 3H).

4.7 Representative Procedure for the Synthesis of Protected Ketones

4.7.1 Diethyl 2-acetamido-2-(3-propionylbenzyl)malonate (38)

To a solution of *m*-Keto toluene **34** (8.0 g, 53.98 mmol) in CCl_4 (134.95 mL) was added NBS (10.57 g, 59.38 mmol) and AIBN (2.67 g, 16.19 mmol), and the resulting suspension was refluxed overnight. Upon completion, the reaction was filtered and concentrated, and the resulting crude material was subjected to the next step without further purification.

A round-bottom flask was charged with brominated **34** (4.24 g, 19.09 mmol), diethyl 2-acetamidomalonate (3.73 g, 17.18 mmol), K_2CO_3 (5.28 g, 38.18 mmol), and KI (3.17 g, 19.09 mmol), followed by 119.3 mL of MeCN, and the resulting suspension was heated to reflux and stirred overnight. Upon completion, the reaction was cooled to room temperature, filtered with celite, and concentrated. Purification via silica gel chromatography (gradient elution, 0 to 30% EtOAc/Hex) afforded **38** as a yellow solid in 58% yield (3.6 g). ^1H NMR (300 MHz, CDCl_3) δ 7.83 (d, $J = 7.8$ Hz, 1 H), 7.64 (s, 1 H), 7.36 (t, $J = 7.8, 7.5$ Hz, 1 H), 7.20 (d, $J = 8.4$ Hz, 1 H), 6.52 (s, 1 H), 4.28 (q, $J = 7.2, 6.9$ Hz, 4 H), 2.96 (q, $J = 7.2$ Hz, 2 H), 2.05 (s, 3 H), 1.31 (t, $J = 7.2$ Hz, 6 H), 1.21 (t, $J = 7.2$ Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 200.5, 169.3, 167.4, 137.0, 136.0, 134.5, 129.3, 128.6, 127.1, 67.2, 63.0, 37.7, 31.9, 23.1, 14.1, 8.3.

4.7.2 Diethyl 2-acetamido-2-(3-butyrylbenzyl)malonate (39)

Synthesized according to the general procedure with ketone **35** (8.0 g, 49.31 mmol), NBS (9.65 g, 54.24 mmol), AIBN (2.43 g, 14.79 mmol), and 123.28 mL CCl_4 . The crude product was subjected to the next step without further purification.

Malonate synthesis was performed according to the general procedure using brominated **35** (5.73 g, 23.75 mmol), diethyl 2-acetamidomalonate (4.64 g, 21.37 mmol), K_2CO_3 (6.56 g, 47.5 mmol), KI (3.94 g, 23.75 mmol), and 148 mL MeCN. 37% yield, 2.9 g. ^1H NMR (300 MHz, CDCl_3) δ 7.82 (d, $J = 7.8$ Hz, 1 H), 7.63 (s, 1 H), 7.34 (t, $J = 7.5$ Hz, 1 H), 7.20 (d, $J = 7.2$ Hz, 1 H), 6.52 (s, 1 H), 4.28 (q, $J = 7.2, 6.9$ Hz, 4 H), 2.89 (t, $J = 7.2$ Hz, 2 H), 2.05 (s, 3 H), 1.75 (sx, $J = 7.5, 7.2$ Hz, 2 H), 1.31 (t, $J = 7.2$ Hz, 6 H), 0.99 (t, $J = 7.5$ Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 200.1, 169.3, 167.4, 137.2, 136.0, 134.5, 129.4, 128.6, 127.1, 67.2, 63.0, 40.6, 37.7, 23.1, 17.8, 14.1, 14.0.

4.7.3 Diethyl 2-acetamido-2-(3-pentanoylbenzyl)malonate (40)

Synthesized according to the general procedure with ketone **36** (9.5 g, 53.9 mmol), NBS (10.55 g, 59.29 mmol), AIBN (2.66 g, 16.17 mmol), and 134.75 mL CCl_4 . The crude product was subjected to the next step without further purification.

Malonate synthesis was performed according to the general procedure using brominated **36** (6.26 g, 24.53 mmol), diethyl 2-acetamidomalonate (4.79 g, 22.08 mmol), K_2CO_3 (6.78 g, 49.07 mmol), KI (4.07 g, 24.53 mmol), and 153.3 mL MeCN. 51% yield as a yellow solid (4.4 g). ^1H NMR (300 MHz, CDCl_3) δ 7.82 (d, $J = 8.1$ Hz, 1 H), 7.62 (s, 1 H), 7.36 (t, $J = 7.8$ Hz, 1 H), 6.52 (s, 1 H), 4.28 (q, $J = 7.2$ Hz, 4 H), 3.7 (s, 2 H), 2.92

(t, $J = 7.5$ Hz, 2 H), 2.05 (s, 3 H), 1.70 (p, $J = 7.8, 7.5, 7.2$ Hz, 2 H), 1.38 (p, $J = 7.8, 7.5, 7.2$ Hz, 2 H), 1.31 (t, $J = 6.9, 7.2$ Hz, 6 H), 0.95 (t, $J = 7.2$ Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 200.2, 169.3, 167.4, 137.2, 135.9, 134.5, 129.3, 128.6, 127.1, 67.2, 62.9, 38.4, 37.7, 26.4, 23.1, 22.5, 14.0.

4.7.4 Diethyl 2-acetamido-2-(3-hexanoylbenzyl)malonate (41)

Synthesized according to the general procedure with ketone **37** (10.0 g, 52.55 mmol), NBS (10.3 g, 57.81 mmol), AIBN (2.59 g, 15.77 mmol), and 131.38 mL CCl_4 . The crude product was subjected to the next step without further purification.

Malonate synthesis was performed according to the general procedure using brominated **37** (5.73 g, 23.75 mmol), diethyl 2-acetamidomalonate (4.52 g, 20.79 mmol), K_2CO_3 (6.39 g, 46.21 mmol), KI (3.84 g, 23.11 mmol), and 144.44 mL MeCN. 37% yield as a yellow solid (3.1 g). ^1H NMR (300 MHz, CDCl_3) δ 7.82 (d, $J = 6.6$ Hz, 1 H), 7.62 (s, 1 H), 7.36 (t, $J = 7.5$ Hz, 1 H), 7.20 (d, $J = 7.5$ Hz, 1 H), 6.52 (s, 1 H), 4.28 (q, $J = 7.2$ Hz, 4 H), 3.71 (s, 2 H), 2.91 (t, $J = 7.5$ Hz, 2 H), 2.05 (s, 3 H), 1.72 (m, 2 H), 1.34 (m, 4 H), 1.31 (m, 6 H), 0.91 (t, $J = 6.6$ Hz, 3 H). ^{13}C NMR (75 MHz, CDCl_3) δ 200.2, 169.4, 167.4, 137.2, 135.9, 134.5, 129.3, 128.6, 127.1, 67.2, 62.9, 38.7, 37.7, 31.6, 24.0, 23.1, 22.6, 14.1.

4.8 Representative Procedure for Malonate Deprotection

4.8.1 2-amino-3-(3-propionylphenyl)propanoic acid hydrochloride (12)

A suspension of **38** (1.0 g, 2.75 mmol) in 6 M HCl was refluxed overnight, until disappearance of protecting groups was verified via ^1H NMR. The resulting solution was concentrated *in vacuo* to yield **12** as a yellow solid (36% yield, 256 mg). ^1H NMR (300 MHz, D_2O) δ 7.93 (d, $J = 6.9$ Hz, 1 H), 7.87 (s, 1 H), 7.55 (m, 2 H), 4.16 (t, $J = 7.2$ Hz, 1 H), 3.30 (dq, $J = 14.1, 5.7$ Hz, 2 H), 3.09 (q, $J = 7.2$ Hz, 2 H), 1.14 (t, $J = 6.9$ Hz, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 206.4, 172.4, 136.8, 135.1, 134.4, 129.3, 128.7, 127.5, 54.8, 35.7, 31.9, 7.5.

4.8.2 2-amino-3-(3-butyrylphenyl)propanoic acid hydrochloride (13)

Synthesized according to the representative procedure using **39** (1.0 g, 2.65 mmol) in 9.14 mL 6 M HCl. 80% yield, 576 mg. ^1H NMR (300 MHz, D_2O) δ 7.95 (d, $J = 7.2$ Hz, 1 H), 7.85 (s, 1 H), 7.53 (m, 2 H), 4.28 (t, $J = 7.2$ Hz, 1 H), 3.31 (dq, $J = 14.7, 5.7$ Hz, 2 H), 3.02 (t, $J = 7.2$ Hz, 2 H), 1.66 (q, $J = 7.2$ Hz, 2 H), 0.92 (t, $J = 7.2$ Hz, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 206.1, 171.5, 136.9, 134.8, 134.5, 129.4, 128.8, 127.8, 54.2, 40.4, 35.4, 17.6, 12.8.

4.8.3 2-amino-3-(3-pentanoylphenyl)propanoic acid hydrochloride (14)

Synthesized according to the representative procedure using **40** (1.08 g, 2.75 mmol) in 9.48 mL 6 M HCl. 29% yield, 200 mg. ^1H NMR 7.96 (dt, $J = 7.2$ Hz, 1 H), 7.87 (s, 1 H), 7.58 (m, 2 H), 4.31 (t, $J = 7.2$ Hz, 1 H), 3.35 (dq, $J = 14.7, 5.7$ Hz, 2 H), 3.07 (t, $J = 7.2$ Hz, 2 H), 1.65 (p, $J = 7.2$ Hz, 2 H), 1.34 (sx, $J = 7.5, 7.2$ Hz, 2 H), 0.91 (t, $J = 7.2$ Hz, 3 H). ^{13}C NMR (75 MHz, D_2O) δ 206.2, 171.5, 136.9, 134.7, 134.5, 129.4, 128.8, 127.8, 54.2, 38.3, 35.4, 26.2, 21.6, 13.0.

4.8.4 2-amino-3-(3-hexanoylphenyl)propanoic acid hydrochloride (15)

Synthesized according to the representative procedure using **41** (1.03 g, 2.55 mmol) in 8.79 mL 6 M HCl. 34% yield, 229 mg. ¹H NMR (300 MHz, D₂O) δ 7.95 (d, *J* = 7.2 Hz, 1 H), 7.87 (s, 1 H), 7.57 (m, 2 H), 4.22 (t, *J* = 7.2 Hz, 1 H), 3.33 (dq, *J* = 14.7, 7.8 Hz, 2 H), 3.08 (t, *J* = 7.2 Hz, 2 H), 1.68 (m, 2 H), 1.33 (m, 4 H), 0.86 (m, 3 H). ¹³C NMR (75 MHz, D₂O) δ 205.2, 181.6, 139.1 136.3, 134.6, 128.8, 128.7, 126.3, 57.3, 40.9, 30.7, 23.8, 21.8, 13.3.

5. sfGFPs2TAG' Protein Sequence

MAXKGEELFTGVVPILVELGDVNGHKFSVRGEGERGDATNGKLTLKFICTTGKL
PVPWPTLVTTLYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYK
TRAEVKFEGLDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIK
ANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRD
HMVLLEFVTAAGITHGMDELYKGSHHHHH

X denotes an amber stop codon for NAA incorporation in this study.

Compounds **5-15** were expressed using the sequence provided above. sfGFP expression of **1-4** used a sequence described previously.^{1,2}

References

1. Wang, Y.-S.; Fang, X.; Wallace, A. L.; Wu, B.; Liu, W. R. *J. Am. Chem. Soc.* **2012**, *134*, 2950-2953.
2. Wang, Y.-S.; Fang, X.; Chen, H.-Y.; Wu, B.; Wang, Z. U.; Hilty, C.; Liu, W. R. *ACS Chem. Biol.* **2013**, *8*, 405-415.
3. Wu, B.; Wang, Z.; Huang, Y.; Liu, W. R. *ChemBiochem.* **2012**, *13*, 1405-1408.
4. a.) Cabiddu, M. G.; Cadoni, E.; Montis, S. D.; Fattouni, C.; Melis, S.; Usai, M. *Tetrahedron* **2003**, *59*, 4383-4387. b.) Archer, A. W.; Claret, P.A.; Hayman, D. F. *J. Chem. Soc. (B)* **1971**, 1231-1240.
5. Drew, S. L.; Lawrence, A. L.; Sherburn, M. S. *Ang. Chem. Int Ed.* **2013**, *52*, 4221-4224.
6. Humphrey, C. E.; Furegati, M.; Laumen, K.; Vecchia, L. L.; Leutert, T.; Muller-Hartwig, C. D.; Vogtle, M. *Org. Process Res. Dev.* **2007**, *11*, 1069-1075.
7. Carroll, F. I.; Blough, B. E.; Abraham, P.; Mills, A. C.; Holleman, J. A.; Wolkenhauer, S. A.; Decker, A. M.; Landavazo, A.; McElroy, K. T.; Navarro, H. A.; Gatch, M. B.; Forster, M. J. *J. Med. Chem.* **2009**, *52*, 6768-6781.
8. Liu, Y.; Yao, B.; Deng, C.-L.; Tang, R.-Y.; Zhang, X.-G.; Li, J.-H. *Org. Lett.* **2011**, *13*, 2184-2187.
9. Meltzer, P. Z.; Butler, D.; Deschamps, J. R.; Madras, B. K. *J. Med. Chem.* **2006**, *49*, 1420-1432.
10. Ruan, J.; Saidi, O.; Iggo, J. A.; Xiao, J. *J. Am. Chem. Soc.* **2008**, *130*, 10510-10511.

6. NMR Data

