SUPPORTING INFORMATION

FOR

SERENDIPITOUS LATE-STAGE MODIFICATION OF DIPEPTIDE BY USING AIBN AND THIOACETIC ACID

Sambasivarao Kotha,*[a] Naveen K. Gupta[a] and Saima Ansari[a]

[a] Department of Chemistry, Indian Institute of Technology Bombay, Mumbai, 400076, India.

E-mail: srk@chem.iitb.ac.in

Table of Contents

General information

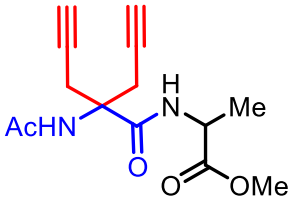
Without further purification, all commercially available reagents were used as received unless otherwise mentioned. Commercial sources such as Aldrich, Avra and Spectrochem were used for buying chemicals such as ethyl isocyanoacetate, DBU, HBTU, DIPEA, and propargyl bromide. Air or moisture sensitive reactions were carried out under inert atmosphere using a nitrogen balloon and in anhydrous and/or degassed solvents. The paraffin oil (purchased from Merck, paraffin liquid heavy) bath was utilized for all heating/reflux reactions, and the continuous water outflow was maintained by attaching the double-surface condenser to a water tap to assure cool water supply throughout the reaction. The drying and distillation of solvents such as dichloromethane (DCM) and acetonitrile were achieved using calcium hydride (CaH_2) and further stored over activated molecular sieves (3Å). The reaction progress was monitored by thin-layer chromatography (2.0 × 4.0 cm^2 alumina plates) using appropriate combination of ethyl acetate (EtOAc) and petroleum ether. By utilizing an appropriate mixture of EtOAc and petroleum ether, column chromatography was performed by using Acme's silica gel (100–200 mesh) for the purification and isolation of compounds from the reaction mixture. High-resolution mass spectrometry (HRMS) measurements of unknown compounds were done by using Bruker or Micromass Q-ToF spectrometers. NMR spectra of all newly synthesized compounds were obtained by using Bruker (AVANCE IIITM) 500 MHz and Bruker (AVANCE IIITM) 400 MHz spectrometers and solvent residual peaks as an internal standard (^{1}H NMR: 500 and 400 MHz, $CDCl_3$ at 7.26 ppm; ^{13}C NMR: 125 and 100 MHz, $CDCl_3$ at 77.2 ppm). ^{1}H NMR data expressed in chemical shift (δ ppm), multiplicity (s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet), and coupling constants (*J* in hertz). The melting point (mp's) of solid compounds was taken from a Buchi 560 melting point instrument and is uncorrected. Single-crystal X-Ray diffraction data was collected on a Bruker diffractometer equipped with graphite monochromated MoKα radiation.

Synthesis of compounds 7

To a solution of acid **6** (1 equiv., 1.04 mmol) in dry DCM (20 mL), HBTU (1.3 equiv., 1.35 mmol) and DIPEA (3 equiv., 3.12 mmol) was added and the reaction mixture was stirred for 10 min. then we added methyl ester hydrochloride of L-amino acid (1.3 equiv., 1.35 mmol). The reaction mixture was stirred at rt for 4 h. After completion of reaction (TLC monitoring), we added 2N HCl, 10% Na_2CO_3, water and brine solution. The aqueous layer was extracted with ethyl acetate

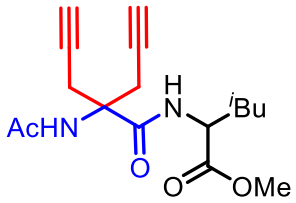
(3 × 10 mL) and dried over anhydrous Na_2SO_4. Evaporation of the solvent gave the crude product, which was purified by silica gel column chromatography (35-45% EtOAc/ petroleum ether) afford the dipeptide **7** (Reaction was carried out in 100 mg scale).

Methyl(2-acetamido-2-(prop-2-yn-1-yl)pent-4-ynoyl)-L-alaninate (7a)



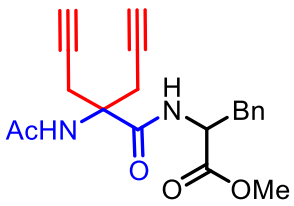
Yield 109 mg, 76%, **Appearance** white solid, **mp** 140-142 ºC, $\mathbf{R_f}$ = 0.1 (40 % EtOAc-petroleum ether), **^{1}H NMR (400 MHz, $CDCl_3$)**: δ 7.26 (d, *J* = 4.4 Hz, 1H), 6.58 (s, 1H), 4.59-4.52 (m, 1H), 3.73 (s, 3H), 3.15-2.96 (m, 4H), 2.13 (t, *J* = 2.4 Hz, 1H), 2.09 (t, *J* = 2.4 Hz, 1H), 2.05 (s, 3H), 1.41 (d, *J* = 7.2 Hz, 3H) ppm, **^{13}C NMR (100 MHz, $CDCl_3$)**: δ 173.2, 170.5, 169.8, 78.9, 78.9, 72.8, 72.5, 60.7, 52.7, 48.9, 25.2, 25.1, 24.0, 18.3 ppm, **HRMS (ESI,Q-ToF) *m/z***: calculated for $C_{14}H_{19}N_2O_4$ $[M+H]^+$ 279.1345, found: 279.1351.

Methyl(2-acetamido-2-(prop-2-yn-1-yl)pent-4-ynoyl)-L-leucinate (7b)



Yield 134 mg, 81%, **Appearance** white solid, **mp** 136-138 ºC, $\mathbf{R_f}$ = 0.2 (40 % EtOAc-petroleum ether), **^{1}H NMR (500 MHz, $CDCl_3$)**: δ 7.14 (d, *J* = 7.5 Hz, 1H), 6.59 (s, 1H), 4.62-4.58 (m, 1H), 3.72 (s, 3H), 3.16-2.99 (m, 4H), 2.12 (t, *J* = 2.5 Hz, 1H), 2.09 (t, *J* = 2.5 Hz, 1H), 2.05 (s, 3H), 1.73-1.55 (m, 3H), 0.93-0.91 (m, 6H) ppm, **^{13}C NMR (125 MHz, $CDCl_3$)**: δ 173.1, 170.6, 170.3, 78.9, 78.9, 72.7, 72.5, 60.8, 52.5, 51.5, 41.7, 25.2, 25.1, 24.9, 24.0, 22.9, 22.1 ppm, **HRMS (ESI,Q-ToF) *m/z***: **HRMS (ESI,Q-ToF) *m/z***: calculated for $C_{17}H_{24}N_2O_4K$ $[M+K]^+$ 359.1373, found: 359.1377.

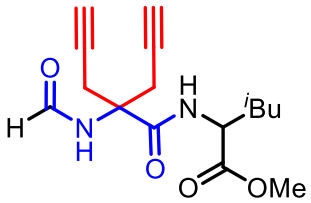
Methyl(2-acetamido-2-(prop-2-yn-1-yl)pent-4-ynoyl)-L-phenylalaninate (7c)



Yield 141 mg, 77%, **Appearance** white solid, **mp** 150-152 ºC, $\mathbf{R_f}$ = 0.2 (40 % EtOAc-petroleum ether), **^{1}H NMR (400 MHz, $CDCl_3$)**: δ 7.26-7.19 (m, 3H), 7.14-7.12 (m, 3H), 6.67 (s, 1H), 4.862\-4.81 (m, 1H), 3.67 (s, 3H), 3.11 (d, J = 6.0 Hz, 2H), 3.06-2.88 (m, 4H), 2.03 (t, *J* = 2.4 Hz, 1H), 2.01 (t, *J* = 2.4 Hz, 1H), 1.98 (s, 3H) ppm, **^{13}C NMR (100 MHz, $CDCl_3$)**: δ 171.5, 170.5, 170.1, 135.8, 129.4, 128.6, 127.1, 78.8, 78.7, 72.6, 72.6, 60.5, 53.8, 52.3,

37.8, 25.0, 24.9, 23.7 ppm, **HRMS (ESI,Q-ToF) *m/z***: calculated for $C_{20}H_{22}N_2O_4K$ $[M+K]^+$ 393.1217, found: 393.1220.

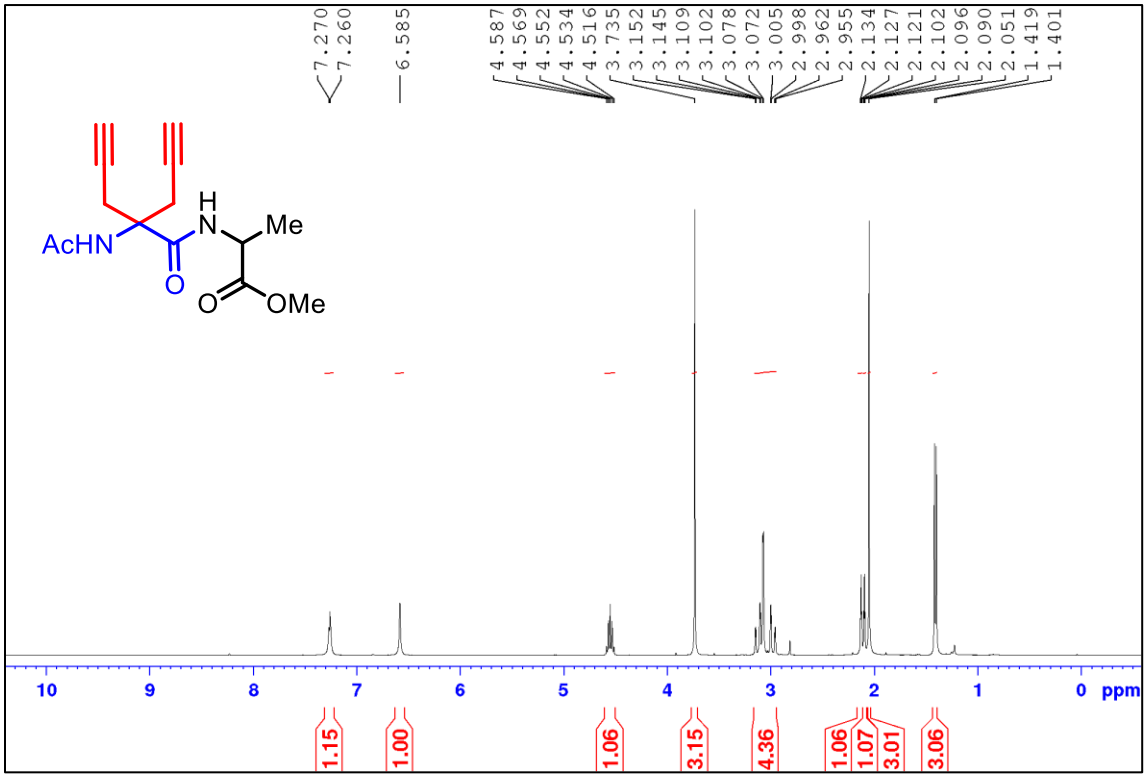
Methyl (2-formamido-2-(prop-2-yn-1-yl)pent-4-ynoyl)-L-leucinate (9)

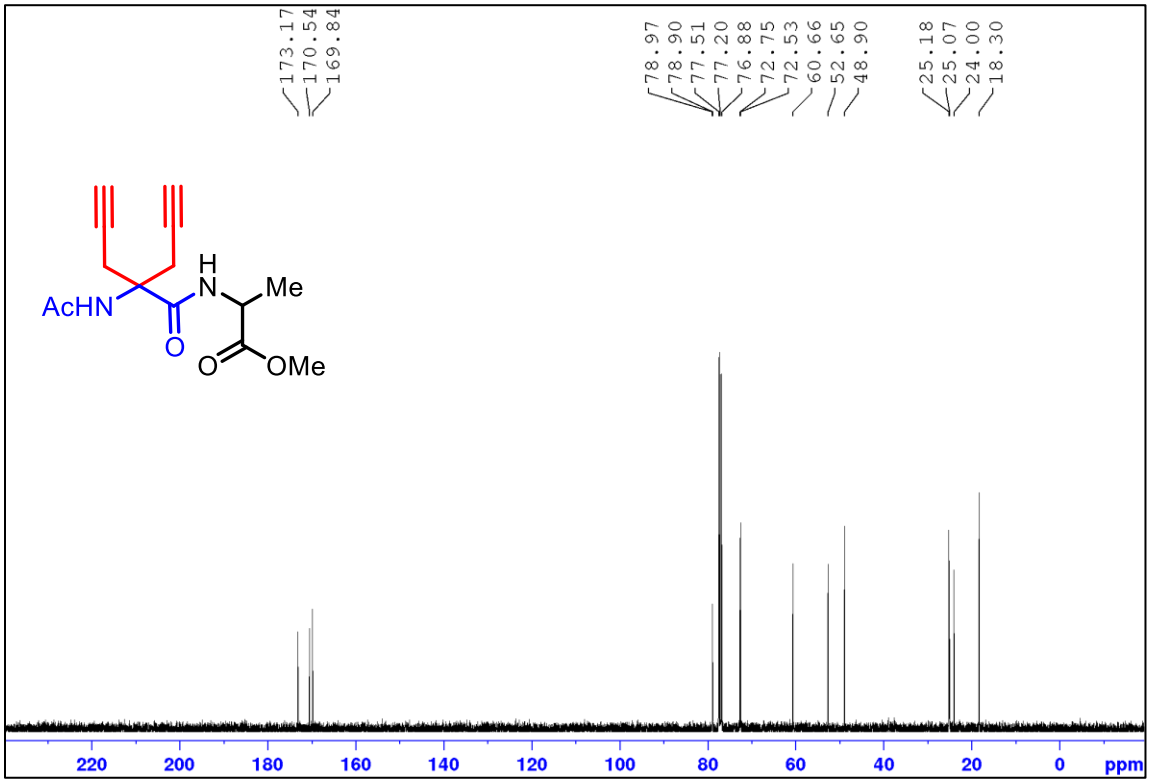


9

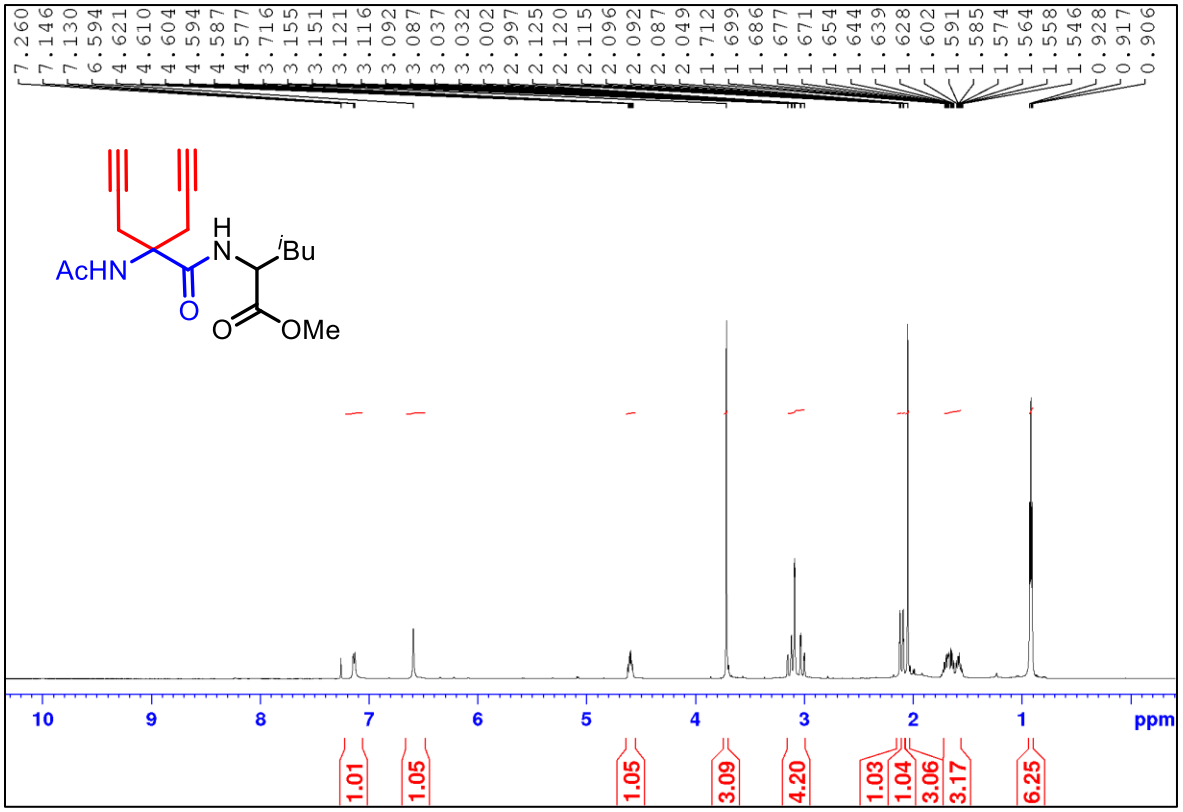
Yield 84 mg, 79 %, **Appearance** colorless solid, **MP** 144-146 ºC **R_f =** 0.6 (50 % EtOAc-petroleum ether), **^{1}H NMR (500 MHz, CDCl3)**: δ 8.25 (s, 1H), 7.08 (d, J = 7.5 Hz, 1H), 6.72 (s, 1H), 4.66-4.61 (m, 1H), 3.74 (s, 3H), 3.22-3.03 (m, 4H), 2.15 (t, J = 2.5 Hz, 1H), 2.12 (t, J = 2.5 Hz, 1H), 1.74-1.58 (m, 3H), 0.95-0.92 (m, 6H) ppm, **^{13}C NMR (125 MHz, CDCl3)**: δ 172.8, 169.6, 160.9, 78.5, 78.4, 72.8, 72.6, 60.6, 52.4, 51.5, 41.6, 25.2, 25.2, 24.8, 22.7, 21.9 ppm, **HRMS (ESI,Q-ToF) *m/z***: calculated for $C_{16}H_{22}N_2NaO_4$ $[M+Na]^+$ 329.1472, found: 329.1471.

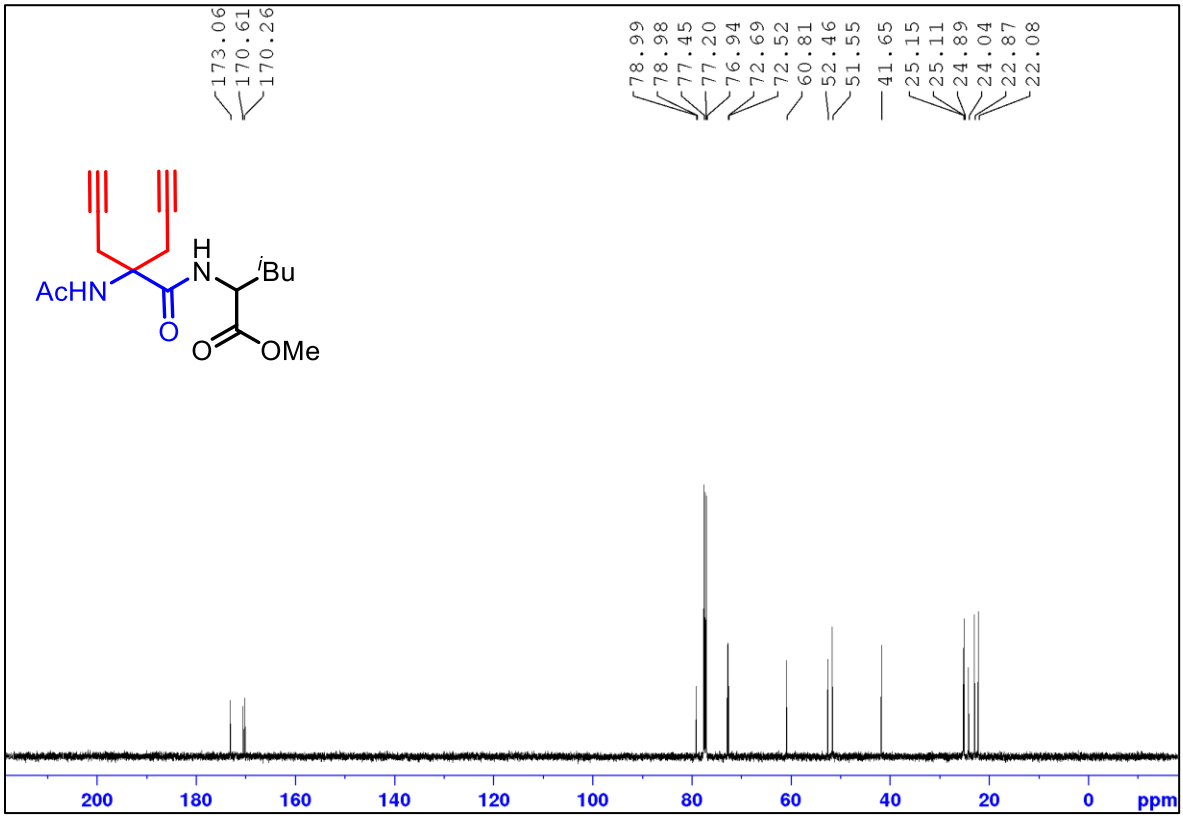
^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) of compound 7a



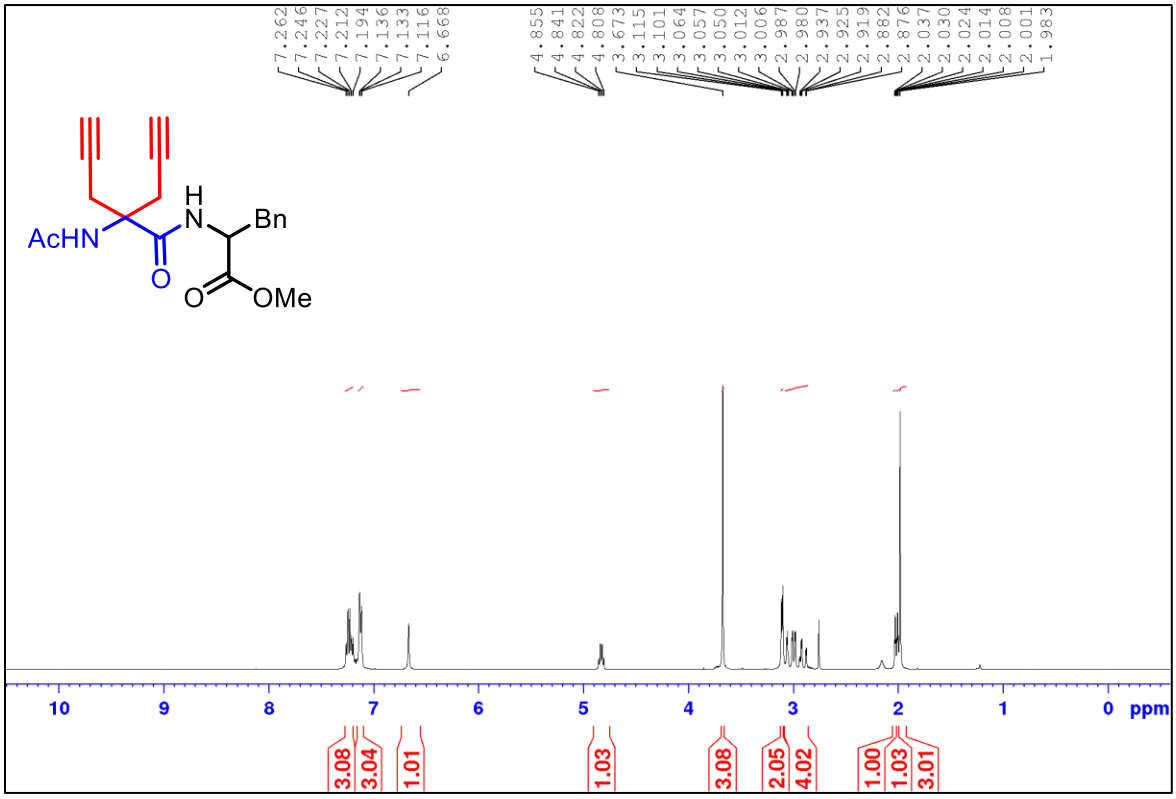


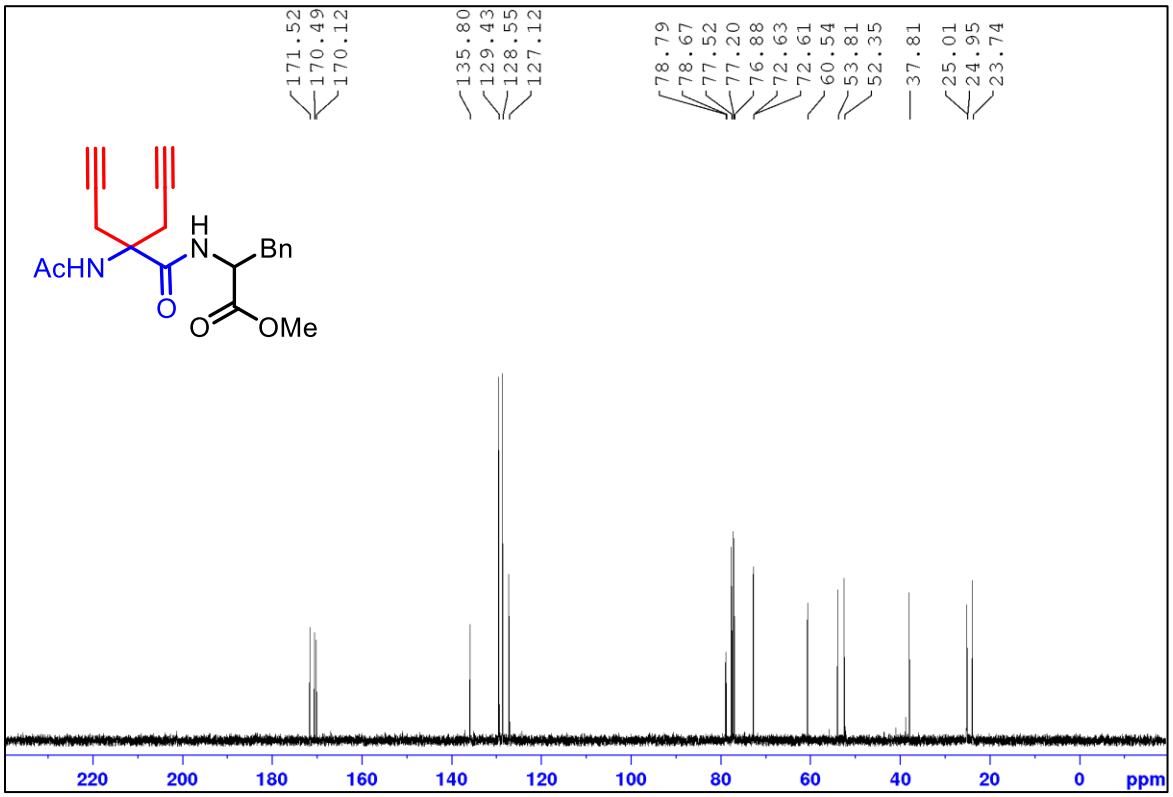
^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) of compound 7b



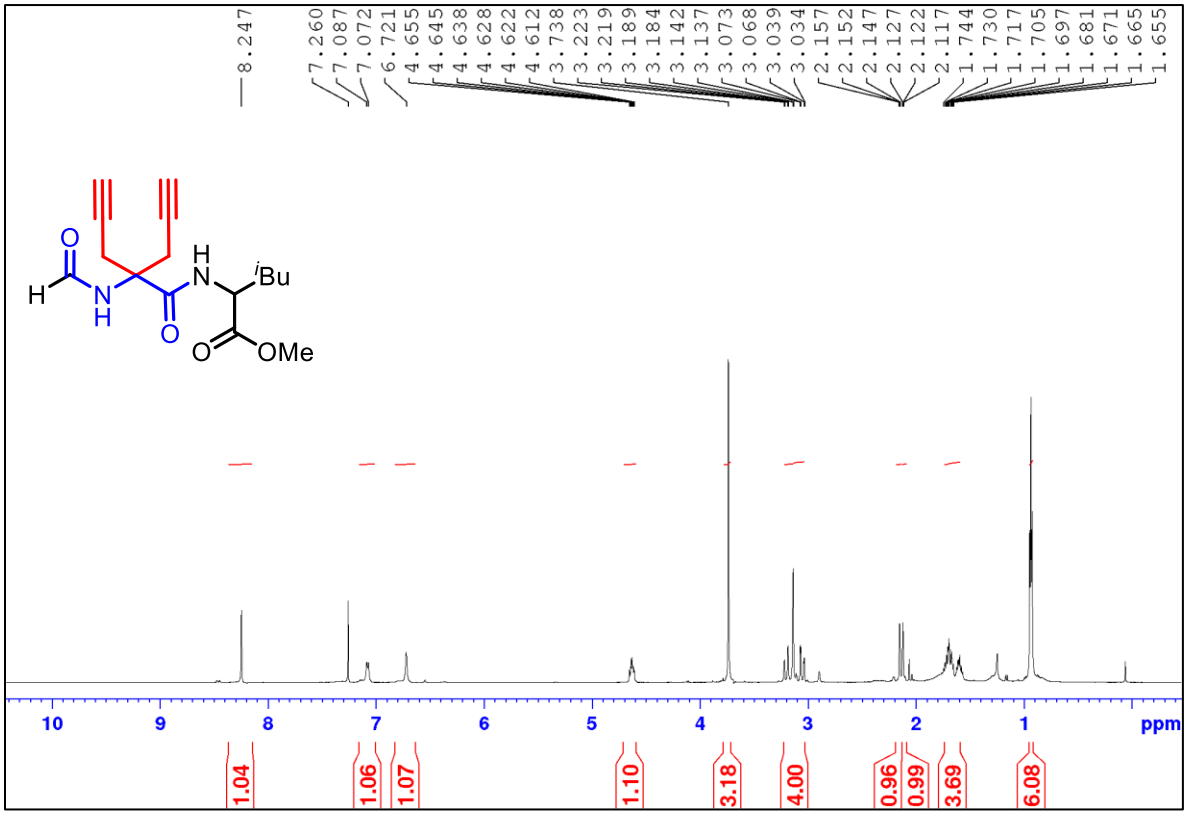


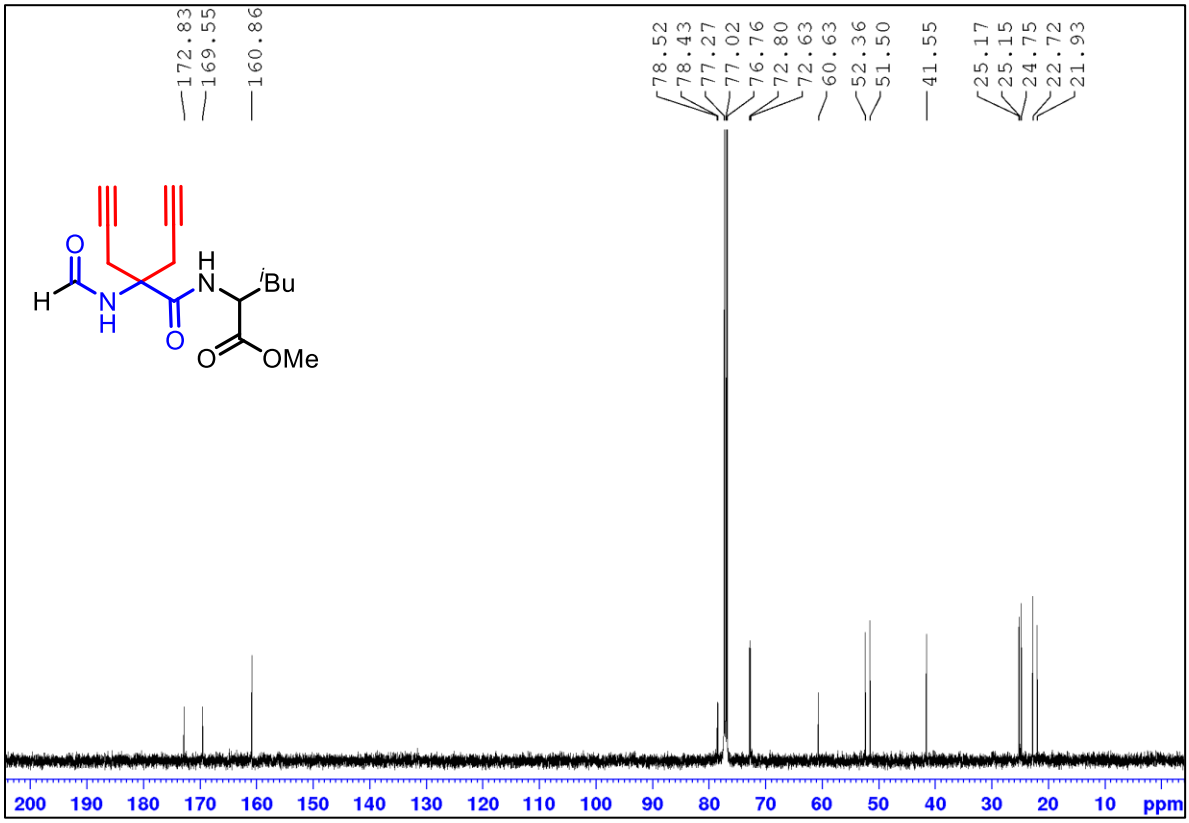
^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) of compound 7c





^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) of compound 9





High Resolution Mass Spectrometry of compound 9

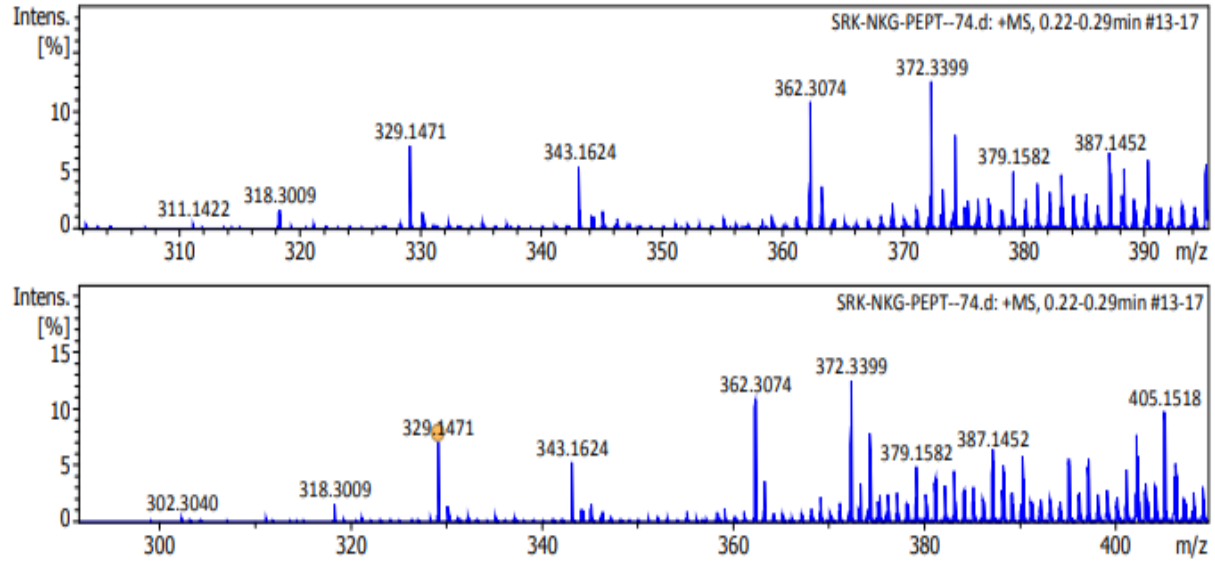
DEPARTMENT OF CHEMISTRY, I.I.T.(B)

Analysis Info

Analysis Name	D:\Data\FEB-21\SRK-NKG-PEPT--74.d	Acquisition Date	2/17/2021 9:23:52 PM
Method	DEFAULT.m	Operator	PPIOUT
Sample Name	SRK-NKG-PEPT--74	Instrument	maXis impact 282001.00081
Comment	C16H22N2O4		

Acquisition Parameter

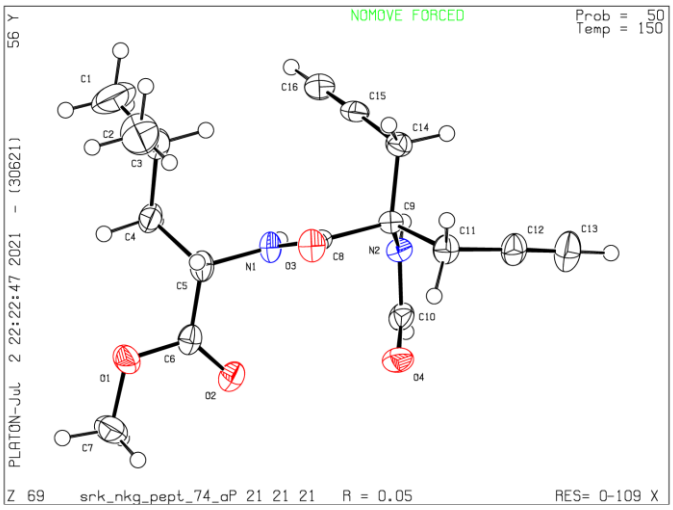
Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.3 Bar
Focus	Not active	Set Capillary	3700 V	Set Dry Heater	180 °C
Scan Begin	50 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Charging Voltage	2000 V	Set Divert Valve	Source
		Set Corona	0 nA	Set APCI Heater	0 °C



Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e⁻ Conf	N-Rule
329.1471	1	C16H22N2NaO4	329.1472	0.2	9.3	1	100.00	7.0	even	ok

Methyl (2-formamido-2-(prop-2-yn-1-yl)pent-4-ynoyl)-L-leucinate (9)

(CCDC Number = 2094117)



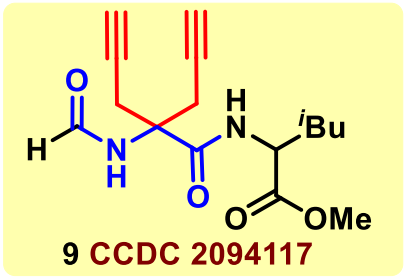


Table S1. X-ray crystallographic data and refinement parameters for **9 (CCDC 2094117)**

Identification code	SRK_NKG_PEPT_74_autored
Empirical formula	$C_{16}H_{22}N_2O_4$
Formula weight	306.35
Temperature	150 K
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Unit cell dimensions	a = 7.5095(5) Å α = 90°
	b = 14.9922(11) Å β = 90°
	c = 15.1453(13) Å γ = 90°
Volume	1705.1(2) $Å^3$
Z	4
Density (calculated)	1.193 g/cm^3
Absorption coefficient (μ)	0.086 mm^{-1}
Absorption correction	Multi-scan
Max. and Min. transmission	1.000-0.281
F (000)	656.0
Crystal size	0.369 x 0.233 x 0.119 mm^3

Index ranges	$-8 \leq h \leq 8, -17 \leq k \leq 17, -15 \leq l \leq 17$
Two-theta range for data collection	4.648 to 56.672°
Reflections collected	15237
Diffraction radiation wavelength	0.71073
Independent reflections	2940 [$R_{(int)} = 0.0724$]
Completeness to $\theta = 24.999°$	99%
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	2940/0/202
Goodness-of-fit on F^2	1.051
Final *R* indices [I>=2σ (I)]	*R*1 = 0.0462, *wR*2 = 0.0975
R indices (all data)	*R*1 = 0.0611, *wR*2 = 0.1051
Largest diff. peak and hole	0.19/-0.18 e $Å^{-3}$