

Supporting Information for:

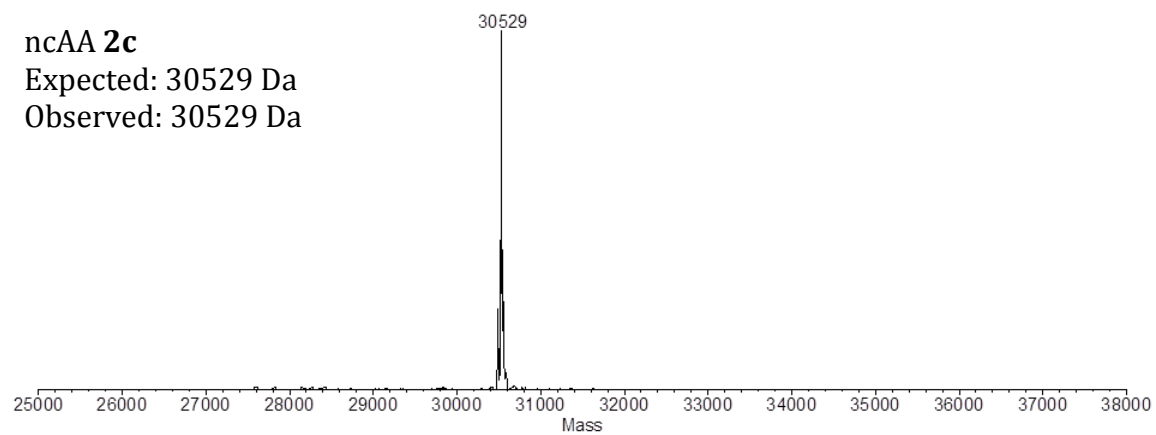
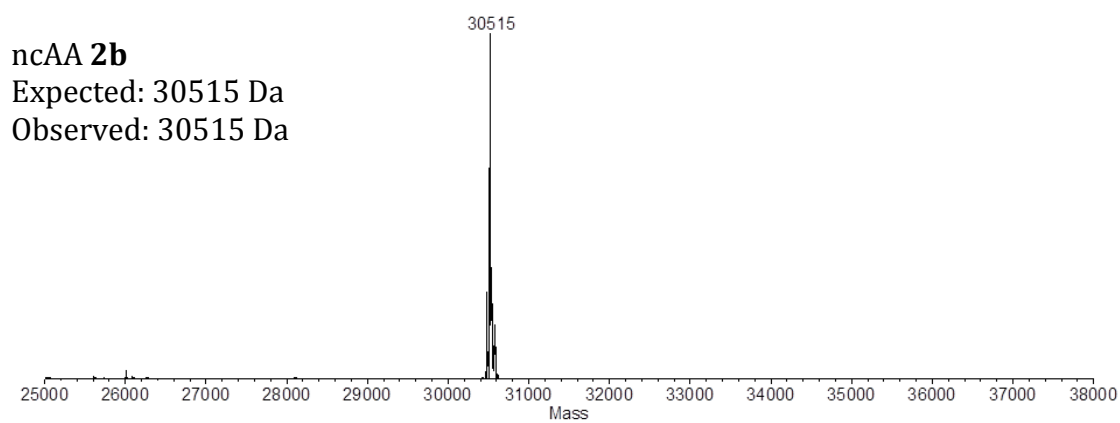
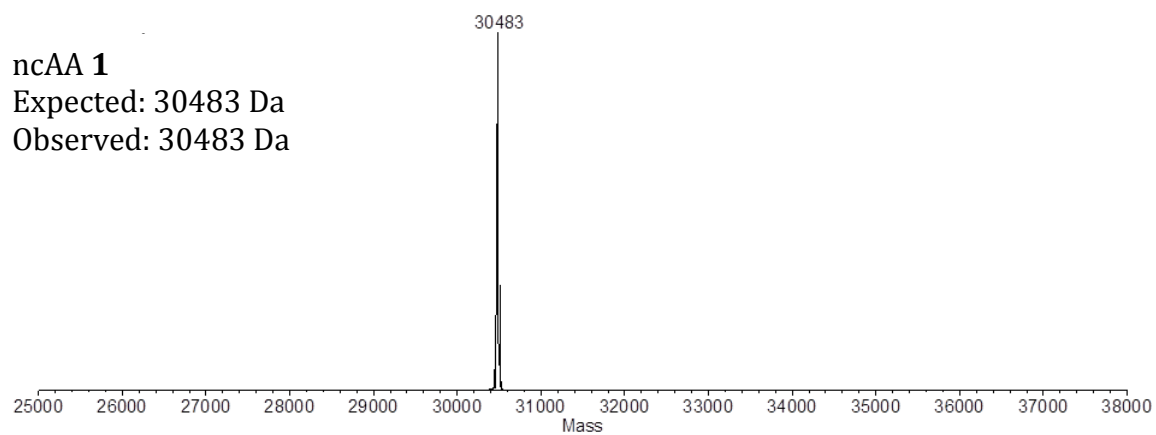
**Expanding the Scope of Single- and Double- Noncanonical Amino Acid Mutagenesis
in Mammalian Cells Using Orthogonal Polyspecific Leucyl-tRNA Synthetases**

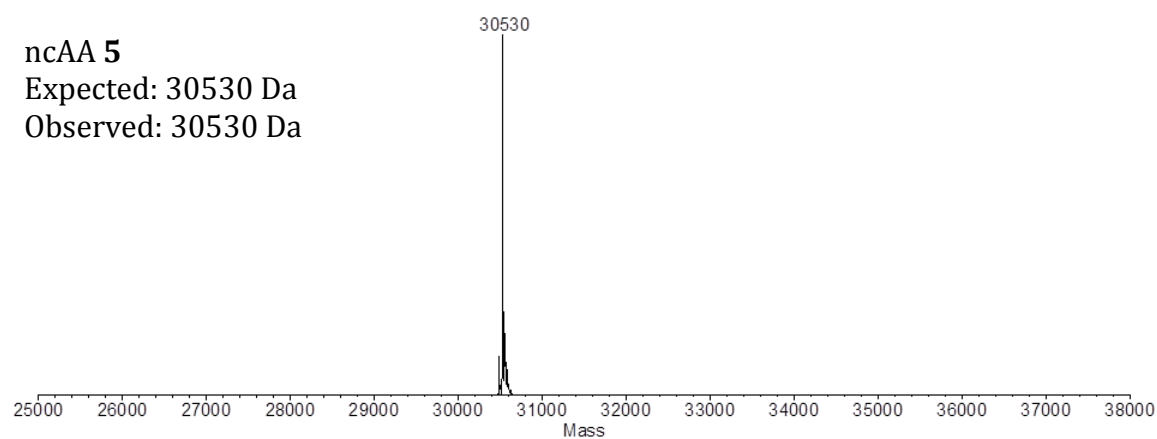
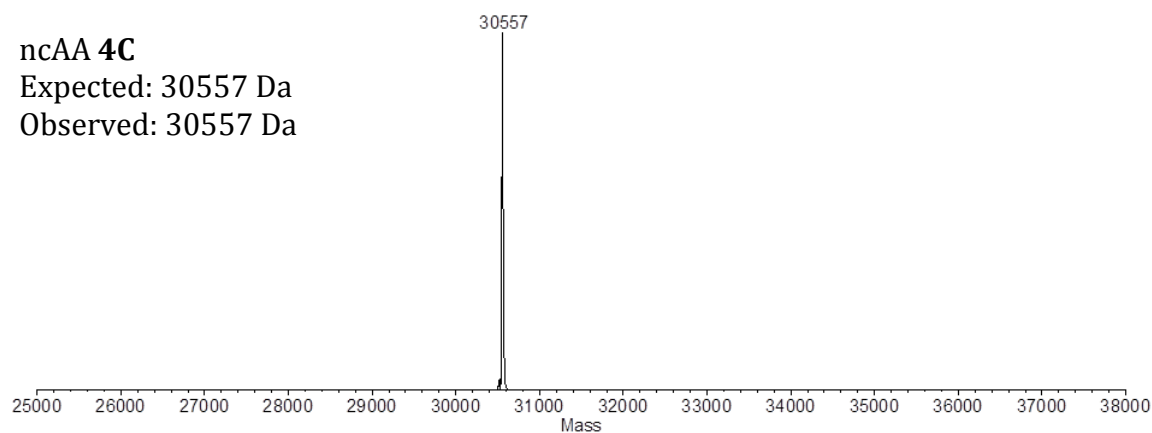
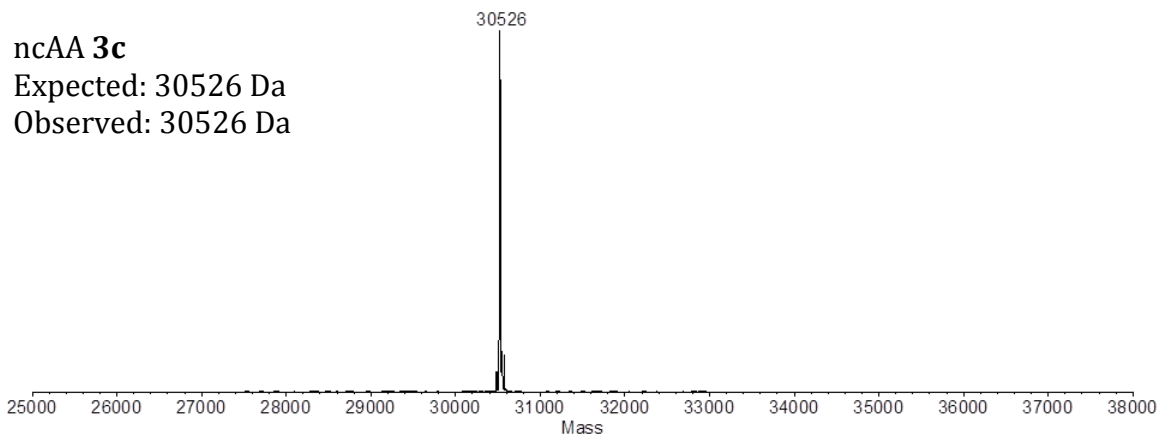
Yunan Zheng,[†] Raja Mukherjee,[†] Melissa A. Chin, Peter Igo, Martin J. Gilgenast,
Abhishek Chatterjee*

Department of Chemistry, Boston College, 2609 Beacon Street, Chestnut Hill, MA
02467, USA.

[†]These authors contributed equally

*email: abhishek.chatterjee@bc.edu





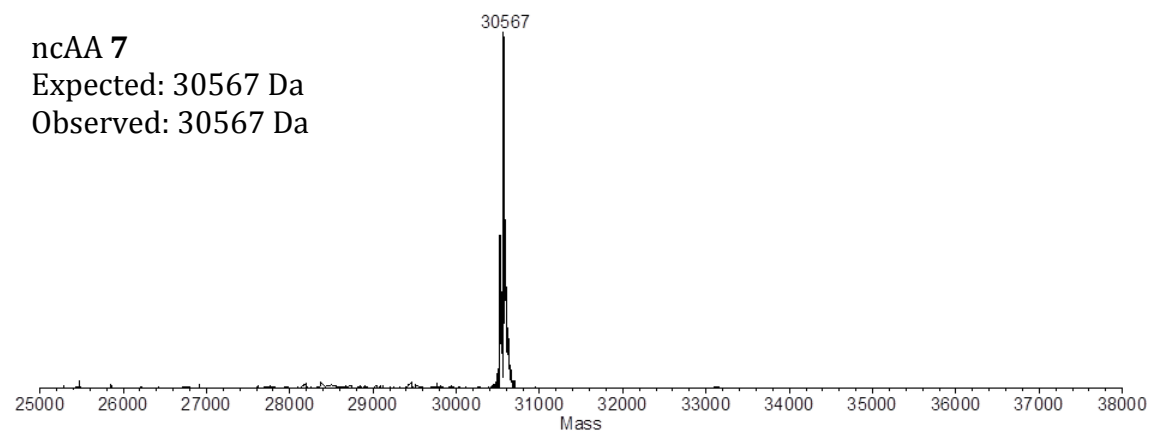
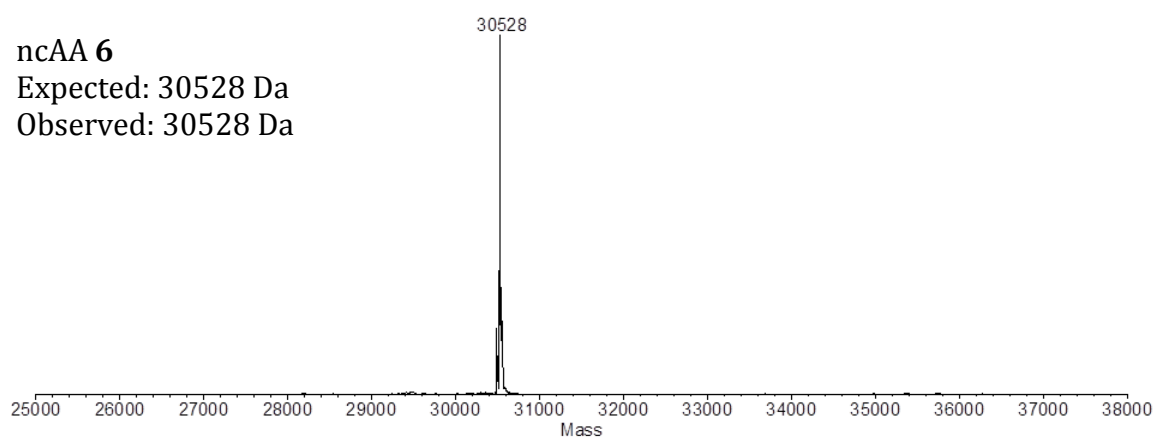


Figure S1: ESI-MS analysis of EGFP-39-TAG expressed in the presence of various ncAAs and the PLRS1/tRNA_{CUA}^{EcLeu} pair show successful ncAA incorporation.

Table S1: Representative yields for different EGFP reporters incorporating one or two different ncAAs using the PLRS1/tRNA_{CUA}^{EcLeu} pair

Reporter	ncAA	EGFP yield (µg/10 cm dish)
EGFP-39-TAG	1	48
EGFP-39-TAG	2b	40
EGFP-39-TAG	2c	40
EGFP-39-TAG	3c	55
EGFP-39-TAG	4c	63
EGFP-39-TAG	5	53
EGFP-39-TAG	6	45
EGFP-39-TAG	7	36
EGFP-39-TAG-151-TGA	4c + 9	2
EGFP-39-TAG-151-TGA	4c + 10	0.6
Wild-type EGFP	--	175

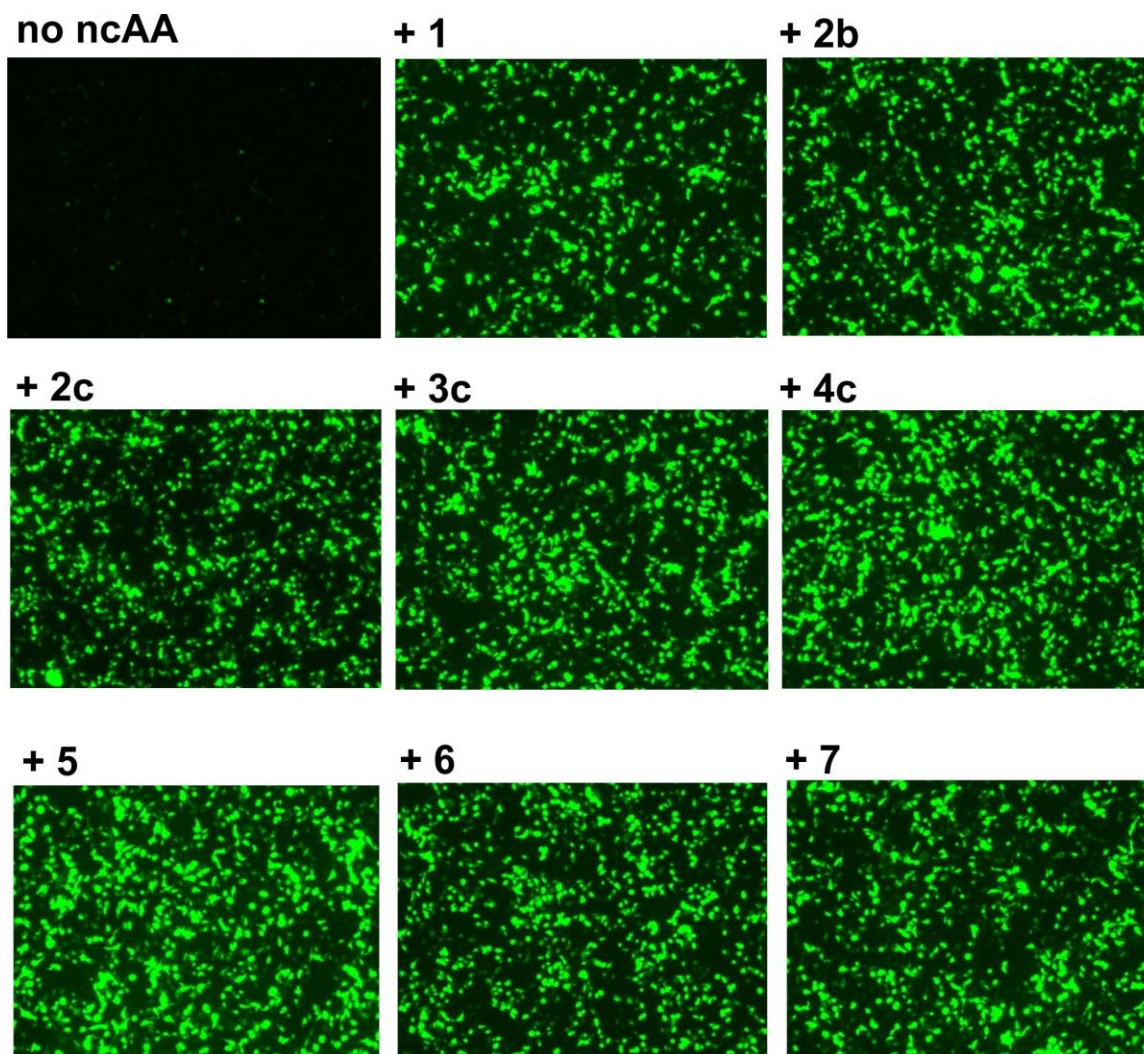


Figure S2: Fluorescence images of HEK293T cells (48 hr post-transfection) transfected with the pAcBac3-EcLeu_{TAG}-EGFP* plasmid in the absence or presence of indicated ncAAs (1 mM) in the growth medium.

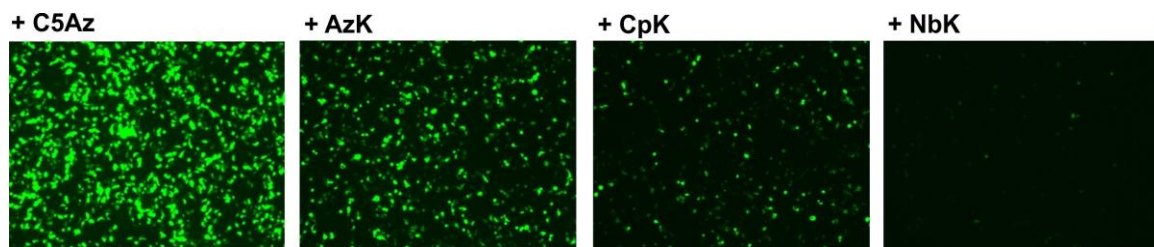


Figure S3: Evaluating if PLRS1 can charge PylRS substrates. Fluorescence images of HEK293T cells (48 hr post-transfection) transfected with the pAcBac3-EcLeu_{TAG}-EGFP* plasmid in the absence or presence of indicated ncAAs (1 mM) in the growth medium.

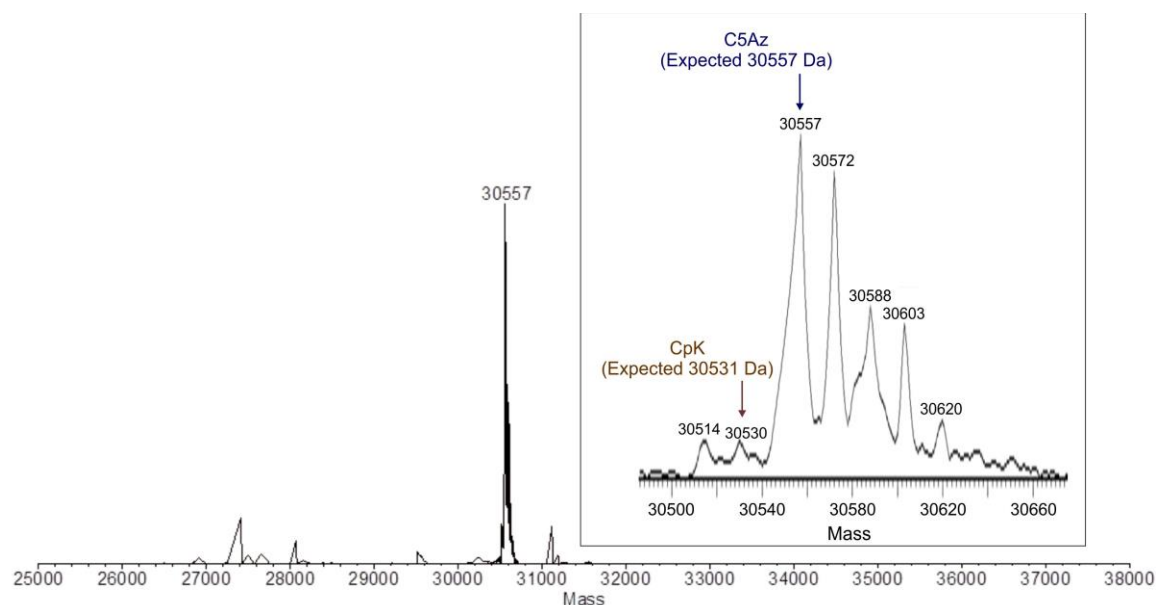


Figure S4: ESI-MS analysis of purified EGFP reporter from HEK293T cells transfected with pAcBac3-EcLeu(PLRS1)_{TAG}-EGFP* plasmid in the presence of 1 mM each of C5Az and CpK. The observed mass (zoomed-in spectra shown) corresponds to near-exclusive incorporation of C5Az (at least >90%). The additional major peaks (30572, 30588, 30603, 30620) are oxidation products of EGFP-39-C5Az, arising likely due to the fact that reducing agents were avoided during protein purification (to prevent unwanted reduction of the azide in C5Az).

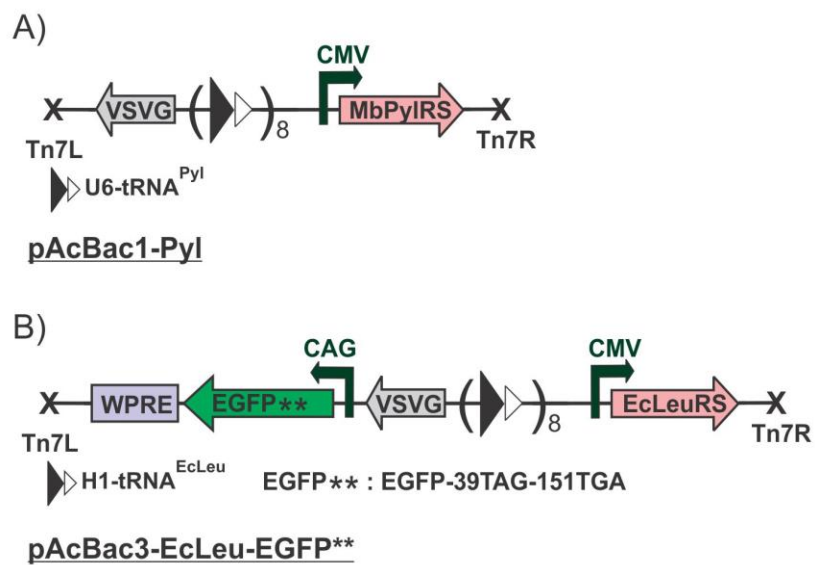


Figure S5: Maps of pAcBac1-Pyl_{TGA} and pAcBac3-EcLeu_{TAG}-EGFP** plasmids

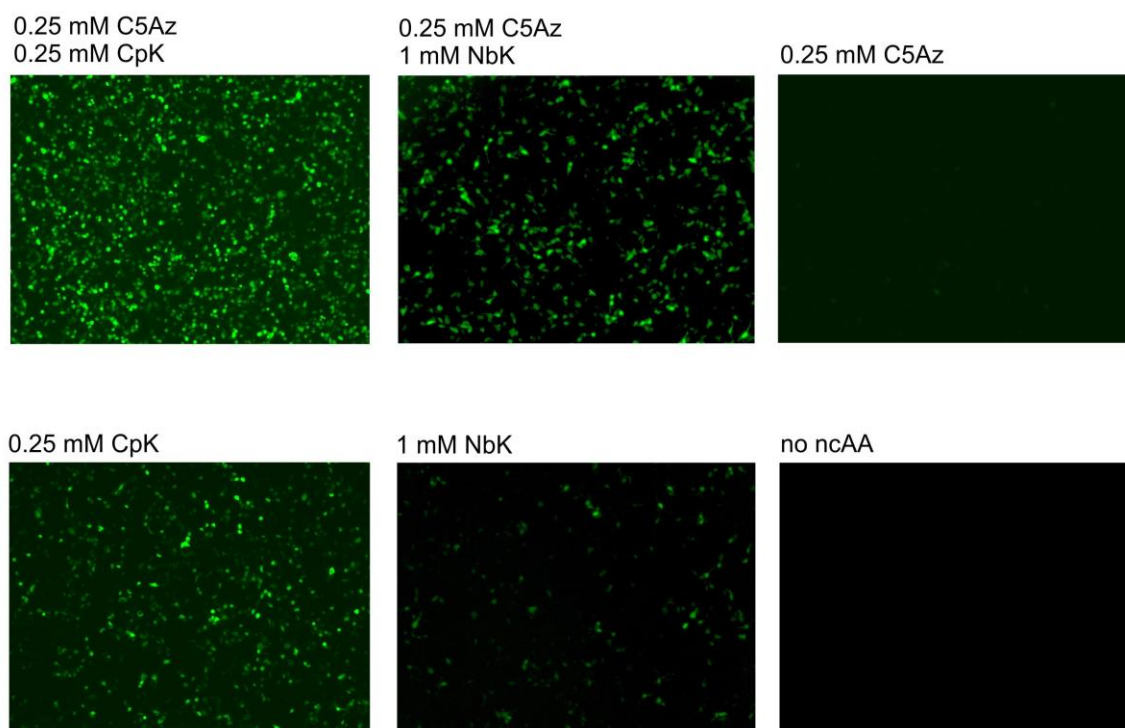


Figure S6: Fluorescence images of cells associated with Figure 4B. 48 hr post-transfection fluorescence images of HEK293T cells co-transfected with pAcBac3-EcLeu(PLRS1)_{TAG}-EGFP** and pAcBac1-Pyl_{TGA} in the absence or presence of indicated ncAAs in the growth medium.

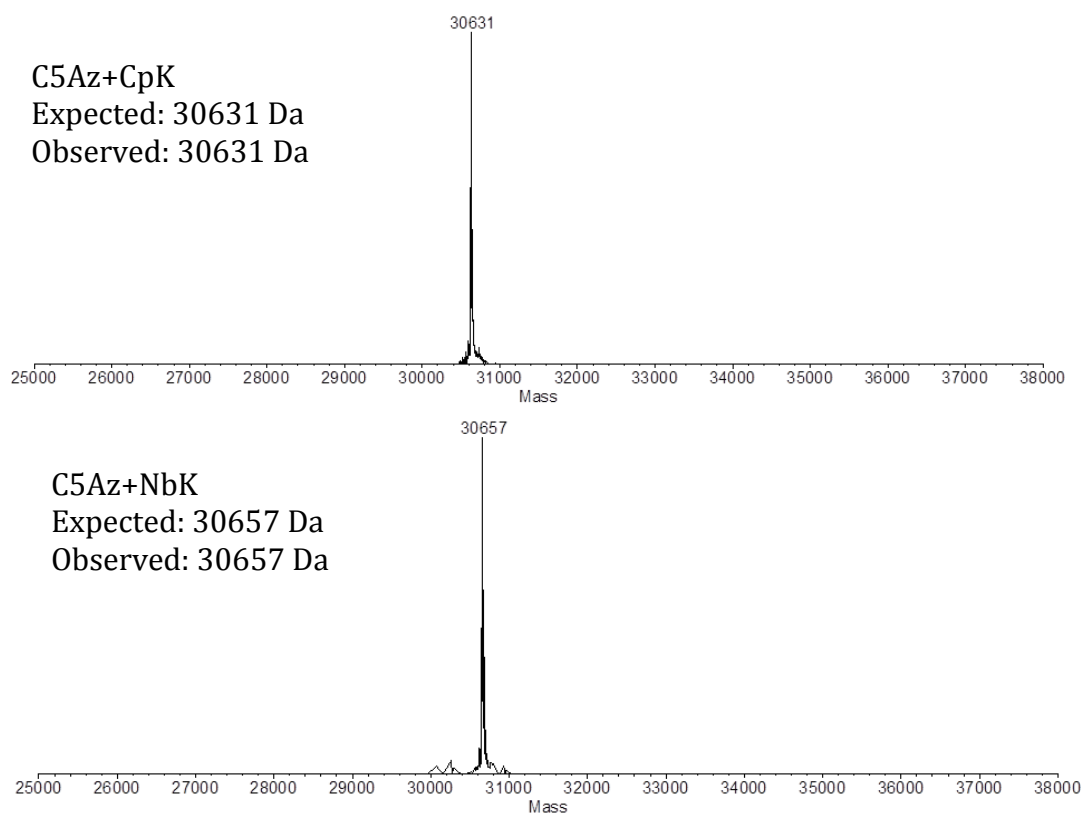


Figure S7: ESI-MS analyses of EGFP-39-TAG-151-TGA expressed in the presence of indicated ncAAs using co-transfection of pAcBac3-EcLeu(PLRS1)_{TAG}-EGFP** and pAcBac1-Pyl_{TGA} reveal mass consistent with successful dual ncAA incorporation.

Materials and Methods

Materials:

The TOP10 *E. coli* strain was used for cloning and plasmid propagation. All oligonucleotides are synthesized by Integrated DNA Technologies (Coralville, IA). *E. coli* cells were grown in LB liquid medium and on LB-agar plates. Antibiotics were used as described: 10 µg/ml gentamycin, 50 µg/ml ampicillin. Restriction enzymes (NEB, Beverly, MA), Phusion Hot Start II DNA polymerase (Fisher Scientific, MA) and T4 DNA ligase (Enzymatics, Beverly, MA) were used for plasmid construction by using standard methods suggested by the suppliers. For maintaining mammalian cell culture, Dulbecco's Modified Eagle's medium (DMEM), fetal bovine serum (FBS) and penicillin/streptomycin solution were obtained from Fisher Scientific. For ncAAs: 2-aminocaproic acid was obtained from TCI America (Portland, Oregon), CpK and AzK was obtained from SicheM (Bremen, Germany). All other ncAAs used were synthesized in lab (see later). For fluorescent probes: DBCO-TAMRA was purchased from Click Chemistry Tools (Scottsdale, Arizona). Tetrazine-fluorescein was synthesized as previously described.¹

Construction of plasmids:

The previously reported pAcBac3-EcLeu_{TAG}-EGFP** (EGFP** = EGFP-39TAG-151TGA) plasmid was used to construct additional plasmids.¹ Site-directed mutagenesis was performed to create polyLeuRS with primers listed below. PLRS1 and PLRS2 were cloned into pAcBac3-EcLeu_{TAG}-EGFP** using NheI/EcoRI restriction sites to create pAcBac3-EcLeu(PLRS1/PLRS2)_{TAG}-EGFP**. For incorporation of a single ncAA, we replaced the EGFP** in the aforementioned plasmids with EGFP* (EGFP*=EGFP-39TAG) using the SfiI restriction sites to generate pAcBac3-EcLeu_{TAG}-EGFP*.

Mammalian cell culture and transfection:

HEK293T cells were maintained in 37 °C humidified incubator supplemented with 5 % CO₂. Transfection was performed following procedures described previously.¹ Cells were seeded as 700,000 cell per well for a 12-well plate 24 hr before transfection. For transfecting cells in one well of a 12-well plate, 1 µg plasmid was incubated with 4 µl PEI (1 mg/mL; Polysciences, Warrington, PA) + 18 µl DMEM for 10 min at room temperature before the mixture was added to cells. For transfecting cells in a 10 cm culture dish, 12 µg plasmid was incubated with 50 µl PEI + 200 µl DMEM. ncAA was added at the same time to a final concentration of 1 mM or 0.25 mM (for dual suppression with CpK and C5Az). 2 mM sodium butyrate was added to enhance protein expression for dual ncAA incorporation.

EGFP fluorescence analysis and purification:

For EGFP fluorescence analysis, cells from a 12-well plate were harvested 48 hr after transfection. 50 µl CellLytic M buffer (Sigma, St. Louis, MO) was used to lyse cells from one well of a 12-well plate. Lysate was centrifuged at 16,000 xg for 5 min and the supernatant was transferred to a clear-bottom 96-well plate for fluorescence measurement following previously described protocol. For EGFP protein purification, 600 µl CellLytic M buffer was used to lyse cells from one 10 cm culture dish. The clarified cell-free

extract was subjected to Ni-NTA affinity chromatography using HisPur resin (Fisher Scientific) following manufacturer's protocol.

Protein labeling:

For dual-labeling EGFP** proteins incorporating either C5Az+CpK or C5Az+NbK, 50 μ M DBCO-TAMRA was added to 5 μ M protein and allowed to incubate for 30 min at RT. Next, 200 μ M tetrazine-fluorescein was added to the mixture and allowed to incubate for an additional 30 min. The mixture was then dialyzed and subjected to SDS-PAGE followed by fluorescence imaging (ChemiDoc MP, Bio-Rad) and ESI-MS (Agilent TOF HPLC-MS) analysis.

List of Primers:

LeuRS-40I-41L-F:

GAGAAAGTATTACTGCCTGTCTATCCTCCCCTATCCTTCTGGTCGACTACAC

LeuRS-40I-41L-R:

GTGTAGTCGACCAGAAGGATAGGGGAGGATAGACAGGCAGTAATACTTCTC

LeuRS-252A-F:

CCGTTTACACTACCCGCCCGGACGCGTTTATGGGTTGTACCTACCTGGC

LeuRS-252A-R:

GCCAGGTAGGTACAACCCATAAACGCGTCCGGGCGGGTAGTGTAACCG

LeuRS-499I-F:

CACCTTTATGGAGTCCTCCTGGATCTATGCGCGCTACACTTGCCC

LeuRS-499I-R:

GGGCAAGTGTAGCGCGCATAGATCCAGGAGGACTCCATAAAGGTG

LeuRS-527A-537G-F:

GCCGGTGGATATCGCTATTGGTGGTATTGAACACGCCATTATGGGTCTGCTCTACTTCCG

LeuRS-527A-537G-R:

CGGAAGTAGAGCAGACCCATAATGGCGTGTTCAATACCACCAATAGCGATATCCACCGGC

LeuRS-NheI-F:

aataatGGCTAGCGTTTAACTTAAGCTTGCCGCCACCATGGAAGAGCAATACC

LeuRS-EcoRI-R:

attattaGAATTCTTAAACGGGCCCCGCCAACGACCAGATTGAG

Sequence of PLRS1:

atggaagagcaataccgcccgaagagatagaatccaaagtacagcttcattgggatgagaagcgcacatttgaagtaaccgaagacgagagcaaaagag
aagtattactgctgtctAtcCTCccctatcctctgctgactacacatgggccacgtacgtaactacaccatcggtgacgtgatgcccgtaccagcgt
atgctgggcaaaacgctcgcagccgatcggtgggacgcgttggctgcctcggaaggcgcggcggtgaaaaacaacaccgctccggcaccgtg
gacgtacgacaacatcgctatatgaaaaaccagctcaaaatgctgggcttgggtatgactggagccgcgagctggcaacctgtacccggaatactacc
gttgggaacagaaattctcaccgagctgtataaaaaagcctggtatataagaagacttctcgggtcaactggtgtccgaacgaccagaccgtactggcga
acgaacaagtatcgacggctgctgctggcgtgcgataccaaagtgaacgtaaagagatcccgagtggttatcaaatcactgcttacgtgacgagc
tgctcaacgatctggataaactggatcactggccagacaccgttaaaaccatgcagcgttaactggatcggtcgttccgaaggcgtggagatcacctcaac
gttaacgactatgacaacacgctgaccgtttacactacccgcccggacGCGtttatgggtgtacctacctggcggtagctgcccgtcatccgctggcgc
agaaagcggcgaaaataatcctgaactggcggcctttattgacgaatgccgtaacaccaaagtgccgaagctgaaatggcgacgatggagaaaaaag
gcgtcgatactggctttaaagcgggttaccattaacggcggaagaaattcccgttgggcagcaaaacttcgtattgatggagtaggcacggcgcgagtta
tggcggtaccgggacgaccagcgcgactacgatttgcctctaataacggcctgaacatcaaaccgggtatcctggcagctgacggctctgagccaga
tcttctcagcaagccctgactgaaaaaggcgtgctgttcaactctggcgagttcaacggcttggacctgaagcggccttcaacgccatcgccgataaactg
actgcgatggcggttggcgagcgtaaagtgaactaccgctgcgcgactggggtgtttcccgtagcggttactggggcgccgattccgatggtgactct
agaagacggtagcgtaatccgacccggacgaccagctccgggtgactcgtccggaagatgtgtaattggacggcattaccagcccattaaagcaga
tccggagtggcgcaaaactaccgttaacggatgccagcactgcgtgaaccgacactttcgacaccttatggagtcctctggATCtatgcgcgtac
acttccccgagtaacaagaaggtatgctggattccgaagcggctaactactgctgcccgttgatcGCTattggtgtattgaacacgccattatgG
Gtctgcttacttccgttctccaaaaactgatgcgtgatgcaggcatggtgaactctgacgaaccagcgaacagttgctgtgcagggtatggtgctggc
agatgccttactatgttggcgaacggcgaacgttaactgggtttccccggttgatgctatcgttgaaactgacgagaaaggccgtatcgtgaaagcgaa
agatgcggcaggccatgaactggtttataccggcatgagcaaatgtccaagtgaagaacaacggatcgaaccgcaggtgatggttgaaactgacggc

gcggacaccgttcgtctgtttatgatgtttcctccggctgatatgactctcgaatggcaggaatccgggtgggaaggggctaaccgttctctgaaacgtgt
 ctggaaactggtttacgagcacacagcaaaaggatgttgcggcactgaacgttgatgcgctgactgaaatcagaaagcgtgcgtcgcgatgtgcata
 aaacgatcgttaagtaccgatgatatcggcgtcgtcagacctcaacaccgcaattgcggcgattatggagctgatgaacaaactggcgaaagcacc
 aaccgatggcgagcaggatcgcgtctgatgcaggaagcactgctggcgttgctccgtatgcttaaccgttcacccgcacatctgcttcacgtgtggca
 ggactgaaaggcgaaaggcgatcgcacacgcgccgtggccggttgctgacgaaaaagcgatggagggaagactccacgtgtgctggtgcaggtta
 acggtaaagtccgtgccaaaatcaccgttcgggtggacgcaacgggaagacaggttcgcgaacgtgctggccaggaacatctgtagcaaaatatctga
 tggcgttactgtacgtaaaagtatttaccgtaccaggtaaactcctcaatctggtcgttggcgggcccgtttaa

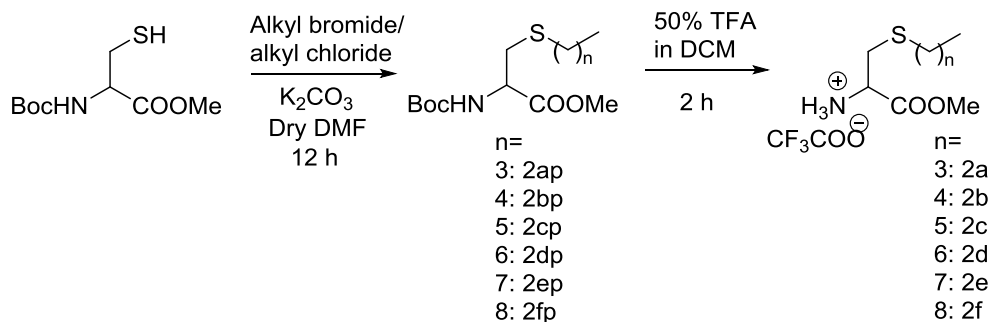
References:

1. Y. Zheng, P. S. Addy, R. Mukherjee, and A. Chatterjee. Chem. Sci. (2017) DOI: 10.1039/c7sc02560b

Synthesis of various ncAAs

General: All chemicals were obtained from commercial sources (Sigma-Aldrich or Fisher Scientific) and used without further purification. ^1H - and ^{13}C -NMR spectra were obtained either on a Varian INOVA-600 MHz, 500 MHz or a 400 MHz spectrophotometer. The nmr spectra were recorded in chloroform-*d* unless otherwise mentioned. Mass spectra were acquired at the Mass Spectrometry Center (Merkert Chemistry Center, Boston College). Alkaline KMnO_4 and ethanolic ninhydrine solution were used as TLC stains.

Synthesis of ncAAs 2a-f:



General experimental procedure:

S-alkylation:

In an oven dried round bottom flask methyl (tert-butoxycarbonyl) cysteinate (500 mg, 2.1mmol) was dissolved in 10 mL dry DMF and potassium carbonate (879 mg, 6.3 mmol) was added under argon. After 10 min, the respective alkyl bromide/chloride (1.1 molar equivalents) was added, followed by the addition of a catalytic amount of KI and the reaction mixture was allowed to stir at room temperature for 12 h. Next, 20 mL of EtOAc and 30 mL of water were added to the reaction mixture and it was transferred to a separating funnel. The aqueous layer was extracted twice with 20 mL of EtOAc. The combined organic layer was washed first with 50 mL of water followed by brine, and dried over anhydrous sodium sulfate. The solvent was evaporated under reduced pressure and the crude material was purified over silica gel column chromatography using hexane-ethyl acetate as eluent.

2ap: Yield: 369 mg, 92%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.34 (d, J = 8.0 Hz, 1H), 4.55 – 4.45 (m, 1H), 3.74 (s, 3H), 2.93 (dd, J = 5.1, 2.9 Hz, 1H), 2.53 – 2.45 (m, 1H), 1.58 – 1.47 (m, 1H), 1.43 (s, 5H), 1.40 – 1.32 (m, 1H), 0.88 (t, J = 7.3 Hz, 3H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.60, 155.09, 53.25, 52.42, 34.43, 32.38, 31.55, 28.26, 21.81, 13.57.

2bp: Yield: 383 mg, 88%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.34 (d, J = 8.1 Hz, 1H), 4.52 (d, J = 7.0 Hz, 1H), 3.75 (s, 3H), 2.95 (dd, J = 5.4, 2.3 Hz, 2H), 2.55 – 2.43 (m, 2H), 1.60 – 1.49 (m, 2H), 1.44 (s, 9H), 1.38 – 1.25 (m, 4H), 0.88 (td, J = 6.8, 0.9 Hz, 3H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.62, 154.44, 53.25, 52.44, 34.47, 32.72, 30.88, 29.19, 28.27, 22.21, 13.91.

2cp: Yield: 400 mg, 87%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.35 (d, J = 8.0 Hz, 1H), 4.52 (t, J = 6.4 Hz, 1H), 3.78 (s, 3H), 3.04 – 2.87 (m, 2H), 2.57 – 2.42 (m, 2H), 1.60 – 1.51 (m, 2H), 1.48 – 1.40 (m, 9H), 1.40 – 1.19 (m, 6H), 0.87 (ddd, J = 7.0, 4.5, 2.1 Hz, 3H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.61, 155.11, 53.25, 52.44, 34.47, 32.76, 31.34, 29.47, 28.40, 28.27, 22.49, 13.97.

2dp: Yield: 464 mg, 95%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.35 (d, J = 7.8 Hz, 1H), 4.51 (d, J = 7.0 Hz, 1H), 3.75 (s, 3H), 2.94 (s, 2H), 2.54 – 2.43 (m, 2H), 1.59 – 1.49 (m, 2H), 1.44 (s, 9H), 1.39 – 1.19 (m, 8H), 0.91 – 0.82 (m, 3H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.60, 155.10, 53.24, 52.42, 34.44, 32.72, 31.66, 29.50, 28.80, 28.67, 28.26, 22.54, 14.01.

2ep: Yield: 482 mg, 93%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.35 (s, 1H), 4.49 (d, J = 7.7 Hz, 1H), 3.73 (s, 3H), 3.00 – 2.82 (m, 2H), 2.48 (td, J = 7.5, 1.6 Hz, 2H), 1.57 – 1.48 (m, 2H), 1.42 (s, 9H), 1.37 – 1.14 (m, 10H), 0.89 – 0.80 (m, 3H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.60, 155.10, 53.24, 52.41, 34.42, 32.71, 31.73, 29.48, 29.11, 29.09, 28.71, 28.25, 22.57, 14.02.

2fp: Yield: 482 mg, 88%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.35 (d, J = 8.1 Hz, 1H), 4.52 (q, J = 5.7 Hz, 1H), 3.75 (s, 3H), 3.02 – 2.86 (m, 2H), 2.50 (td, J = 7.3, 1.0 Hz, 2H), 1.60 – 1.49 (m, 2H), 1.47 – 1.41 (m, 9H), 1.41 – 1.32 (m, 1H), 1.32 – 1.21 (m, 12H), 0.91 – 0.84 (m, 3H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.62, 155.11, 53.25, 52.44, 34.46, 32.75, 31.82, 29.51, 29.43, 29.20, 29.16, 28.73, 28.27, 22.62, 14.05.

Boc-deprotection:

The compound (2ap-2fp, 300 mg) was dissolved in 5 mL dry DCM and treated with 50% TFA in DCM (5 mL). After 2 hours, all the starting material disappeared as revealed by TLC analysis. The solvent was then removed using a rotovap. The crude product was washed thrice with hexane (3 X 30 mL) and diethylether (3 X 30 mL) separately, and the remaining TFA was co-evaporated after dissolving in methanol. Lastly, the compound was dissolved in 3 mL water and cryodesiccated overnight to get pure ncAAs as their TFA salt in near-quantitative yield.

2a: Yield, 295 mg, 94%. ^1H NMR (500 MHz, Deuterium Oxide) δ 4.20 (dd, $J = 7.2, 4.6$ Hz, 1H), 3.69 (s, 3H), 3.03 (dd, $J = 15.0, 4.6$ Hz, 1H), 2.93 (dd, $J = 15.0, 7.3$ Hz, 1H), 2.45 – 2.39 (m, 2H), 1.43 – 1.35 (m, 2H), 1.25 – 1.14 (m, 2H), 0.71 (t, $J = 7.4$ Hz, 3H). ^{13}C NMR (125 MHz, D_2O) δ 169.23, 53.64, 52.20, 31.25, 30.90, 30.55, 21.03, 12.70. HRMS (DART) calcd. for $\text{C}_8\text{H}_{18}\text{NO}_2\text{S}$ ($\text{M}^+ + 1$) 192.1058, found 192.1061.

2b: Yield, 301 mg. ^1H NMR (500 MHz, Deuterium Oxide) δ 4.23 (dd, $J = 7.2, 4.6$ Hz, 1H), 3.72 (s, 3H), 3.06 (dd, $J = 15.0, 4.6$ Hz, 1H), 2.96 (dd, $J = 15.0, 7.2$ Hz, 1H), 2.48 – 2.40 (m, 2H), 1.50 – 1.39 (m, 2H), 1.24 – 1.11 (m, 4H), 0.72 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (125 MHz, D_2O) δ 169.27, 53.67, 52.23, 31.56, 30.92, 30.00, 28.12, 21.45, 13.12. HRMS (DART) calcd. for $\text{C}_9\text{H}_{20}\text{NO}_2\text{S}$ ($\text{M}^+ + 1$) 206.1215, found 206.1211.

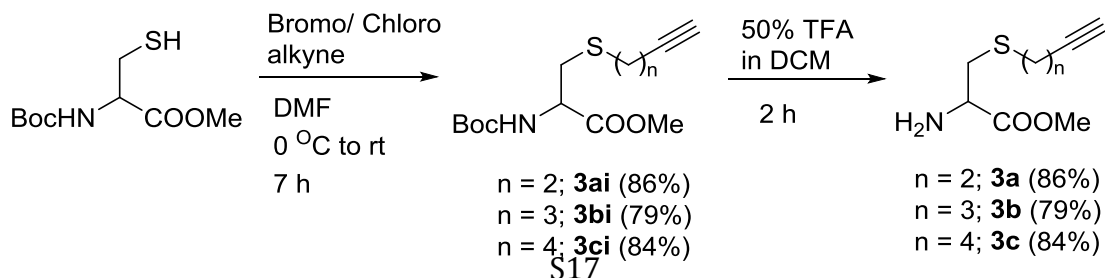
2c: Yield, 310 mg, 99%. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 8.61 (s, 3H), 4.28 (td, $J = 5.9, 1.7$ Hz, 1H), 3.73 (s, 3H), 2.96 (t, $J = 5.2$ Hz, 2H), 2.60 – 2.44 (m, 2H), 1.55 – 1.42 (m, 3H), 1.36 – 1.20 (m, 8H), 0.84 (t, $J = 6.8$ Hz, 3H). ^{13}C NMR (125 MHz, $\text{DMSO}-d_6$) δ 169.16, 53.32, 52.35, 32.02, 31.73, 31.21, 29.22, 28.21, 22.42, 14.29. HRMS (DART) calcd. for $\text{C}_{10}\text{H}_{22}\text{NO}_2\text{S}$ ($\text{M}^+ + 1$) 220.1371, found 220.1264.

2d: Yield, 297 mg, 95%. ^1H NMR (600 MHz, $\text{DMSO}-d_6$) δ 8.58 (bs, 3H), 4.29 (t, $J = 5.8$ Hz, 1H), 3.74 (s, 3H), 2.96 (d, $J = 5.8$ Hz, 2H), 2.57 – 2.50 (m, 2H), 1.54 – 1.43 (m, 2H), 1.33 – 1.18 (m, 8H), 0.88 – 0.81 (m, 3H). ^{13}C NMR (150 MHz, $\text{DMSO}-d_6$) δ 169.17, 53.35, 52.34, 32.02, 31.74, 31.60, 29.21, 28.66, 28.47, 22.47, 14.35. HRMS (DART) calcd. for $\text{C}_{11}\text{H}_{24}\text{NO}_2\text{S}$ ($\text{M}^+ + 1$) 234.1528, found 234.1533.

2e: Yield, 306 mg, 98%. ^1H NMR (600 MHz, $\text{DMSO}-d_6$) δ 8.58 (bs, 3H), 4.29 (t, $J = 5.8$ Hz, 1H), 3.74 (s, 3H), 2.96 (d, $J = 5.8$ Hz, 2H), 2.56 – 2.47 (m, 2H), 1.49 (td, $J = 8.4, 7.5, 6.1$ Hz, 2H), 1.30 – 1.22 (m, 10H), 0.86 – 0.83 (m, 3H). ^{13}C NMR (150 MHz, $\text{DMSO}-d_6$) δ 169.18, 53.37, 52.33, 32.02, 31.75, 31.66, 29.21, 29.03, 28.96, 28.52, 22.51, 14.38. HRMS (DART) calcd. for $\text{C}_{12}\text{H}_{26}\text{NO}_2\text{S}$ ($\text{M}^+ + 1$) 248.1684, found 248.1680.

2f: Yield, 297 mg, 95%. ^1H NMR (600 MHz, $\text{DMSO}-d_6$) δ 8.55 (bs, 3H), 4.29 (t, $J = 5.8$ Hz, 1H), 3.74 (s, 3H), 2.95 (dd, $J = 6.5, 5.4$ Hz, 2H), 2.53 (dtd, $J = 14.7, 7.3, 4.2$ Hz, 2H), 1.54 – 1.41 (m, 2H), 1.34 – 1.19 (m, 12H), 0.88 – 0.81 (m, 3H). ^{13}C NMR (150 MHz, $\text{DMSO}-d_6$) δ 169.18, 53.38, 52.33, 32.02, 31.76, 31.70, 29.34, 29.21, 29.08, 29.01, 28.52, 22.53, 14.38. HRMS (DART) calcd. for $\text{C}_{13}\text{H}_{28}\text{NO}_2\text{S}$ ($\text{M}^+ + 1$) 262.1841, found 262.1837.

Synthesis of ncAAs **3a-c**:



Synthesis: To an ice cold solution of methyl (tert-butoxycarbonyl) cysteinate (590 mg, 2.5 mmol) in 10 mL dry DMF, potassium carbonate (690 mg, 5 mmol) was added under argon, followed by the dropwise addition of the respective chloro or bromo alkyne (1.1 equivalents, 3 mmol) and the reaction mixture was allowed to stir at room temperature for 7 h. After which, 20 mL of EtOAc and 30 mL of water was added to the reaction mixture. The aqueous layer was extracted twice with 20 mL of EtOAc. The combined organic layer was washed first with 50 mL of water followed by 30 mL of brine and dried over anhydrous sodium sulfate. The solvent was evaporated under reduced pressure and the crude material was purified over silica gel column chromatography using EtOAc-petroleum ether as eluent.

3ai: Yield 634 mg, 86%. ¹H NMR (500 MHz, Chloroform-*d*) δ 5.36 (d, *J* = 8.1 Hz, 1H), 4.53 (d, *J* = 6.2 Hz, 1H), 3.76 (s, 3H), 3.08 – 2.96 (m, 2H), 2.70 (t, *J* = 7.1 Hz, 2H), 2.46 (td, *J* = 7.3, 2.6 Hz, 2H), 2.02 (t, *J* = 2.6 Hz, 1H), 1.44 (s, 9H). ¹³C NMR (125 MHz, CDCl₃) δ 171.37, 155.07, 82.12, 80.17, 69.68, 53.32, 52.56, 34.52, 31.42, 28.27, 19.81.

3bi: Yield 569 mg, 79%. ¹H NMR (500 MHz, Chloroform-*d*) δ 5.34 (d, *J* = 8.0 Hz, 1H), 4.53 (d, *J* = 7.2 Hz, 1H), 3.76 (s, 3H), 2.96 (t, *J* = 5.8 Hz, 2H), 2.64 (t, *J* = 7.2 Hz, 2H), 2.30 (td, *J* = 6.9, 2.6 Hz, 2H), 1.96 (t, *J* = 2.7 Hz, 1H), 1.77 (p, *J* = 7.1 Hz, 2H), 1.44 (s, 9H). ¹³C NMR (125 MHz, CDCl₃) δ 171.48, 155.07, 83.16, 80.12, 69.08, 53.29, 52.51, 34.47, 31.45, 28.28, 28.11, 17.32.

3ci: Yield 634 mg, 84%. ¹H NMR (500 MHz, Chloroform-*d*) δ 5.33 (d, *J* = 8.1 Hz, 1H), 4.50 (q, *J* = 5.9 Hz, 1H), 3.73 (s, 3H), 3.01 – 2.83 (m, 2H), 2.52 (t, *J* = 7.1 Hz, 2H), 2.18 (td, *J* = 6.9, 2.7 Hz, 2H), 1.96 – 1.89 (m, 1H), 1.72 – 1.63 (m, 2H), 1.63 – 1.55 (m, 2H), 1.42 (s, 9H). ¹³C NMR (125 MHz, CDCl₃) δ 171.54, 155.07, 83.81, 80.05, 68.68, 53.22, 52.46, 34.41, 32.08, 28.31, 28.26, 27.24, 17.93.

Boc deprotection:

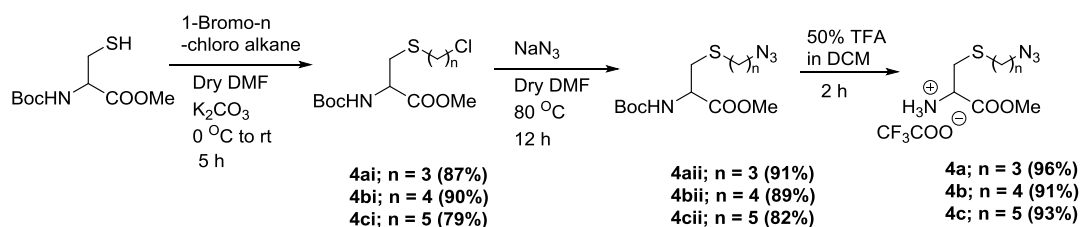
The compound (**3ai-3ci**, 550 mg) was dissolved in 5 mL dry DCM and treated with 50% TFA in DCM (5 mL). After 2 hours, all the starting material disappeared as followed by TLC. The solvent was then removed using a rotovap. The crude product was washed thrice with hexane (3 X 30 mL) and diethyl ether (3 X 30 mL) separately and the remaining TFA was co evaporated by dissolving the crude product in 1:1 mixture of methanol-diethyl ether, and removing the solvent under reduced pressure. Lastly, the compound was dissolved in 3 mL water and the solvent was removed using a lyophilizer overnight to get pure respective ncAAs as their TFA salts.

3a: Yield 559 mg, 97%. ¹H NMR (500 MHz, Deuterium Oxide) δ 4.28 (dd, *J* = 7.5, 4.5 Hz, 1H), 3.73 (s, 3H), 3.17 (dd, *J* = 15.1, 4.5 Hz, 1H), 3.03 (dd, *J* = 15.0, 7.5 Hz, 1H), 2.64 (dq, *J* = 14.9, 6.6 Hz, 2H), 2.42 (td, *J* = 7.0, 2.7 Hz, 2H), 2.31 (td, *J* = 2.6, 0.5 Hz, 1H). ¹³C NMR (125 MHz, D₂O) δ 169.12, 83.31, 70.65, 53.76, 52.33, 31.08, 30.36, 18.73. HRMS (DART) calcd. for C₈H₁₄NO₂S (M⁺+1) 188.0745, found 188.0749

3b: Yield 541 mg, 88%. ^1H NMR (500 MHz, Deuterium Oxide) δ 4.26 (dd, $J = 7.3, 4.5$ Hz, 1H), 3.74 (s, 3H), 3.10 (dd, $J = 15.0, 4.6$ Hz, 1H), 2.98 (dd, $J = 15.0, 7.3$ Hz, 1H), 2.60 – 2.54 (m, 2H), 2.25 (t, $J = 2.7$ Hz, 1H), 2.20 (td, $J = 6.9, 2.7$ Hz, 2H), 1.72 – 1.63 (m, 2H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.78, 87.22, 72.42, 56.30, 54.81, 33.49, 32.90, 29.68, 19.02. HRMS (DART) calcd. for $\text{C}_9\text{H}_{16}\text{NO}_2\text{S}$ ($\text{M}^+ + 1$) 202.0902, found 202.0899

3c: Yield 563 mg, 98%. ^1H NMR (500 MHz, Deuterium Oxide) δ 4.24 (ddd, $J = 6.8, 4.5, 2.2$ Hz, 1H), 3.72 (d, $J = 2.4$ Hz, 3H), 3.07 (ddd, $J = 14.9, 4.3, 2.4$ Hz, 1H), 3.01 – 2.91 (m, 1H), 2.47 (t, $J = 7.3$ Hz, 2H), 2.22 (q, $J = 2.4$ Hz, 1H), 2.10 (td, $J = 7.0, 2.9$ Hz, 2H), 1.57 (dq, $J = 14.9, 7.4, 6.5$ Hz, 2H), 1.46 (p, $J = 7.2$ Hz, 2H). ^{13}C NMR (125 MHz, D_2O) δ 169.25, 85.72, 69.37, 53.72, 52.25, 30.96, 30.90, 27.50, 26.47, 17.05. HRMS (DART) calcd. for $\text{C}_{10}\text{H}_{18}\text{NO}_2\text{S}$ ($\text{M}^+ + 1$) 216.1058, found 216.1068.

Synthesis of ncAAs 4a-c:



Chloroalkylation:

In an oven dried round bottom flask methyl (tert-butyloxycarbonyl) cysteinate (470 mg, 2 mmol) was dissolved in 10 mL dry DMF and potassium carbonate (552 mg, 4 mmol) was added under argon. The flask was cooled to 0°C and after 10 min the respective 1-bromo-n-chloro alkane ($n = 3, 4, 5$; 1.1 equivalents) was added and the reaction mixture was allowed to stir at room temperature for 5 h. Next, the reaction mixture was diluted with 20 mL EtOAc and transferred to a separating funnel. 30 mL of water was added and the aqueous layer was extracted twice (2 X 30 mL) with ethyl acetate. The combined organic layer was washed with 50 mL each of water and brine, and dried over anhydrous sodium sulfate. The solvent was removed using rotovap and the crude product was isolated using column chromatography (silica-gel) with hexane-ethyl acetate as eluent.

4ai: Yield 541 mg, 87 %. ^1H NMR (500 MHz, Chloroform- d) δ 5.34 (d, $J = 8.1$ Hz, 1H), 4.53 (d, $J = 6.0$ Hz, 1H), 3.76 (s, 3H), 3.62 (t, $J = 6.3$ Hz, 2H), 2.97 (qd, $J = 13.9, 5.2$ Hz, 2H), 2.68 (t, $J = 7.0$ Hz, 2H), 2.07 – 1.95 (m, 2H), 1.44 (s, 9H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.40, 155.05, 53.31, 52.56, 43.16, 34.56, 32.03, 29.67, 28.27, 28.10.

4bi: Yield 585 mg, 90%. ^1H NMR (500 MHz, Chloroform- d) δ 5.33 (d, $J = 8.0$ Hz, 1H), 4.50 (dd, $J = 8.9, 4.4$ Hz, 1H), 3.73 (s, 3H), 3.51 (t, $J = 6.5$ Hz, 2H), 2.93 (qd, $J = 13.8, 5.2$ Hz, 2H), 2.53 (t, $J = 7.1$ Hz, 2H), 1.83 (dq, $J = 8.6, 6.2$ Hz, 2H), 1.75 – 1.64 (m, 2H), 1.42 (s, 9H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.49, 155.07, 53.23, 52.49, 44.34, 34.42, 31.84, 31.27, 28.26, 26.53

4ci: Yield 535 mg, 79%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.34 (d, J = 8.0 Hz, 1H), 4.46 (q, J = 6.1 Hz, 1H), 3.70 (s, 3H), 3.47 (t, J = 6.6 Hz, 2H), 2.96 – 2.82 (m, 2H), 2.48 (t, J = 7.2 Hz, 2H), 1.76 – 1.67 (m, 2H), 1.59 – 1.51 (m, 2H), 1.51 – 1.41 (m, 2H), 1.41 – 1.39 (m, 9H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.49, 155.05, 76.84, 53.23, 52.43, 44.67, 34.41, 32.39, 32.03, 28.68, 28.24, 25.89.

Azide formation:

To the chloroalkyl cysteine methyl ester (**4ai-4ci**, 500 mg) in 10 mL dry DMF, sodium azide (2 equivalents) was added and the reaction mixture was stirred for 12 h at 80 °C under argon atmosphere. The reaction mixture was then cooled at room temperature, diluted with ethyl acetate (25 mL) and 30 mL water was added before taking in a separating funnel. The aqueous layer was extracted with EtOAc (2 X 25 mL). The combined organic layer was dried over anhydrous sodium sulfate before purifying via silica gel chromatography using EtOAc-hexane as eluent.

4aii: Yield 465 mg, 91%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.35 (d, J = 8.0 Hz, 1H), 4.53 (d, J = 7.3 Hz, 1H), 3.76 (s, 3H), 3.39 (t, J = 6.6 Hz, 2H), 2.95 (qd, J = 13.9, 5.2 Hz, 2H), 2.61 (t, J = 7.1 Hz, 2H), 1.87 – 1.78 (m, 2H), 1.44 (s, 9H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.40, 155.06, 53.29, 52.54, 49.85, 34.57, 29.62, 28.62, 28.26.

4bii: Yield 454 mg, 89%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.33 (d, J = 8.0 Hz, 1H), 4.50 (d, J = 7.2 Hz, 1H), 3.74 (s, 3H), 3.25 (t, J = 6.8 Hz, 2H), 2.92 (dt, J = 13.6, 5.8 Hz, 2H), 2.51 (t, J = 7.3 Hz, 2H), 1.63 – 1.52 (m, 4H), 1.43 (s, 9H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.50, 155.08, 53.25, 52.51, 50.93, 34.49, 32.13, 28.27, 27.79, 26.52.

4cii: Yield 418 mg, 82%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.33 (d, J = 8.0 Hz, 1H), 4.50 (d, J = 7.2 Hz, 1H), 3.74 (s, 3H), 3.25 (t, J = 6.8 Hz, 2H), 2.92 (dt, J = 13.6, 5.8 Hz, 2H), 2.51 (t, J = 7.3 Hz, 2H), 1.63 – 1.52 (m, 4H), 1.43 (m, 10H). ^{13}C NMR (125 MHz, cdcl_3) δ 171.53, 155.08, 53.25, 52.47, 51.22, 34.50, 32.45, 28.96, 28.26, 25.76.

Boc deprotection:

The compound (**4aii-4cii**, 405 mg) was dissolved in 5 mL dry DCM and treated with 50% TFA in DCM (5 mL). After 2 hours the solvent was removed using a rotatory evaporator. The crude product was washed thrice with hexane (3 X 30 mL) and diethyl ether (3 X 30 mL) separately, and the remaining TFA was co evaporated by dissolving the crude product in 1:1 mixture of methanol-diethyl ether. Lastly, the compound was dissolved in 3 mL water and the solvent was removed in a lyophilizer overnight to get pure respective ncAAs as their TFA salt in excellent yields.

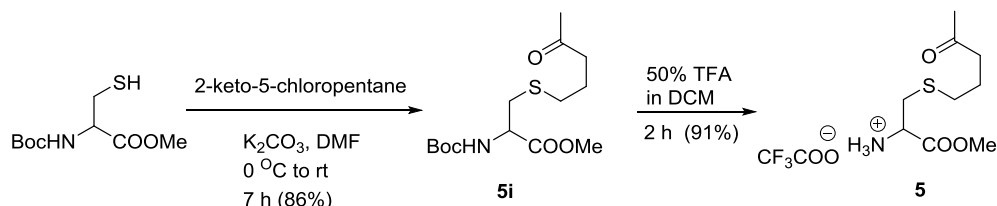
4a: Yield 406 mg, 96%. ^1H NMR (500 MHz, Deuterium Oxide) δ 4.39 – 4.19 (m, 1H), 3.69 (m, 3H), 3.41 – 3.28 (m, 2H), 3.19 – 3.09 (m, 1H), 3.09 – 2.99 (m, 1H), 2.62 – 2.54 (m, 2H), 1.83 – 1.74 (m, 2H). ^{13}C NMR (150 MHz, D_2O) δ 169.16, 53.65, 52.21, 50.57,

30.85, 26.91, 25.68. HRMS (DART) calcd. for C₇H₁₅N₄O₂S (M⁺+1) 219.0916, found 219.0920.

4b: Yield 384 mg, 91%. ¹H NMR (600 MHz, Deuterium Oxide) δ 4.18 (dd, *J* = 7.2, 4.5 Hz, 1H), 3.66 (s, 3H), 3.21 – 3.11 (m, 2H), 3.01 (dd, *J* = 15.0, 4.6 Hz, 1H), 2.91 (dd, *J* = 15.0, 7.2 Hz, 1H), 2.47 – 2.41 (m, 2H), 2.03 – 1.95 (m, 1H), 1.48 (q, *J* = 4.9, 3.5 Hz, 3H). ¹³C NMR (150 MHz, D₂O) δ 169.16, 53.67, 52.19, 50.56, 31.01, 30.85, 26.91, 25.68. HRMS (DART) calcd. for C₈H₁₇N₄O₂S (M⁺+1) 233.1072, found 233.1079.

4c: Yield 392 mg, 93%. ¹H NMR (600 MHz, Deuterium Oxide) δ 4.22 (dd, *J* = 7.2, 4.6 Hz, 1H), 3.70 (s, 3H), 3.17 (t, *J* = 6.8 Hz, 2H), 3.05 (dd, *J* = 15.0, 4.6 Hz, 1H), 2.95 (dd, *J* = 15.0, 7.2 Hz, 1H), 2.45 (t, *J* = 7.3 Hz, 2H), 1.51 – 1.41 (m, 4H), 1.34 – 1.26 (m, 2H). ¹³C NMR (150 MHz, D₂O) δ 169.24, 53.69, 52.25, 50.95, 31.37, 30.92, 27.93, 27.44, 24.94. HRMS (DART) calcd. for C₉H₁₉N₄O₂S (M⁺+1) 247.1229, found 247.123.

Synthesis of ncAA **5**:



Synthesis: To an ice cold solution of methyl (tert-butoxycarbonyl)cysteinate (500 mg, 2.13 mmol) in 5 mL dry DMF, potassium carbonate (882 mg, 6.39 mmol) and 2-keto-n-chloro-alkane was added successively and stirred under argon atmosphere until TLC showed the disappearance of the starting material. Then, 10 mL ethyl acetate and 15 mL water was added and the organic layer was extracted thrice with (3 X 15 mL) ethyl acetate. The organic layers were combined and washed with 15 mL each of water and brine, before drying over anhydrous sodium sulfate. The solvent was removed under reduced pressure and the crude product was purified by column chromatography (hexane-ethyl acetate).

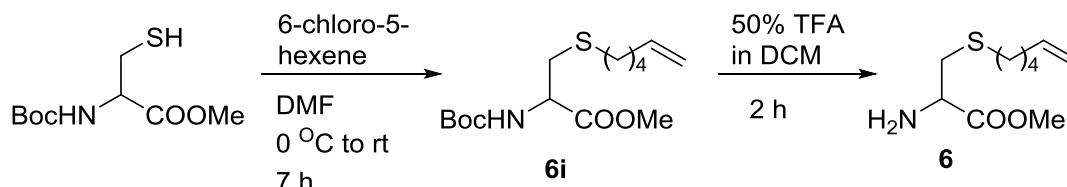
The boc-deprotection of purified **5i** was carried out as described earlier.

5i: Yield 584 mg, 86%. ¹H NMR (500 MHz, CDCl₃) δ 5.33 (bs, 1H), 4.49 (m, 1H), 3.73 (s, 3H), 2.96– 2.86 (m, 2H), 2.53–2.49 (m, 4H), 2.12 (s, 3H), 1.84– 1.78 (m, 2H), 1.42 (s, 9H). ¹³C NMR (125 MHz, CDCl₃) δ 171.49, 155.08, 53.18, 52.48, 41.82, 34.26, 31.94, 29.98, 28.25, 23.00.

5: Yield 522 mg. ¹H NMR (500 MHz, Deuterium Oxide) δ 4.23 (dd, *J* = 7.2, 4.6 Hz, 1H), 3.70 (d, *J* = 0.6 Hz, 3H), 3.05 (ddd, *J* = 15.0, 4.6, 0.6 Hz, 1H), 2.94 (ddd, *J* = 15.0, 7.3, 0.6 Hz, 1H), 2.52 (t, *J* = 7.1 Hz, 2H), 2.42 (t, *J* = 7.3 Hz, 2H), 2.05 (s, 3H), 1.73 – 1.62 (m, 2H). ¹³C NMR (125 MHz, D₂O) δ 169.27, 53.67, 52.23, 31.56, 30.92, 30.00, 28.12,

21.45, 13.12. HRMS (DART) calcd. for C₉H₁₈NO₃S (M⁺+1) 220.1007, found 220.1009.

Synthesis of ncAA 6:



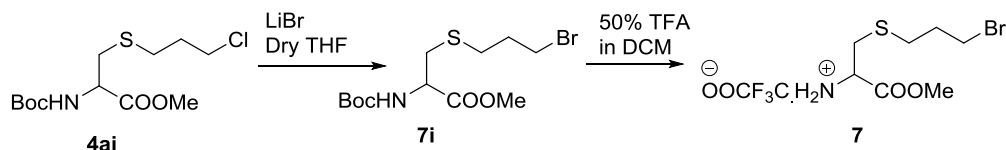
To an ice cold solution of methyl (tert-butoxycarbonyl) cysteinate (1g, 4.25 mmol) in 15 mL dry DMF, potassium carbonate (1.76 g, 12.75 mmol) was added under argon atmosphere. After 10 min the 6-chloro-5-hexene (0.67 mL, 5.1 mmol) was added in a dropwise manner and the reaction mixture was allowed to stir at room temperature for 7 h. After diluting the reaction mixture with 20 mL of EtOAc, it was partitioned with 30 mL of water in a separating funnel. The aqueous layer was extracted twice with 20 mL of additional EtOAc. The combined organic layer was washed first with 50 mL of water followed by 30 mL of brine and dried over anhydrous sodium sulfate. The solvent was evaporated under reduced pressure and the crude material was purified over silica gel column chromatography using EtOAc-petroleum ether as eluent.

Next, boc deprotection of **6i** was carried out as described earlier to get the alkene containing ncAA **6**.

6i: Yield 1.12 g, 83%. ¹H NMR (500 MHz, Chloroform-*d*) δ 5.76 (ddt, *J* = 17.0, 10.2, 6.7 Hz, 1H), 5.34 (d, *J* = 8.0 Hz, 1H), 5.02 – 4.89 (m, 2H), 4.50 (q, *J* = 6.1 Hz, 1H), 3.73 (s, 3H), 2.93 (t, *J* = 5.0 Hz, 2H), 2.50 (dd, *J* = 7.7, 6.9 Hz, 2H), 2.07 – 1.99 (m, 2H), 1.61 – 1.50 (m, 2H), 1.42 (s, 11H). ¹³C NMR (125 MHz, CDCl₃) δ 171.57, 155.09, 138.30, 114.71, 53.24, 52.44, 34.44, 33.19, 32.52, 28.87, 28.26, 27.85, -0.05.

6: Yield 950 mg, 91%. ¹H NMR (500 MHz, Deuterium Oxide) δ 5.74 (ddt, *J* = 17.1, 10.3, 6.7 Hz, 1H), 4.97 – 4.81 (m, 2H), 4.22 (dd, *J* = 7.2, 4.6 Hz, 1H), 3.71 (s, 2H), 3.05 (dd, *J* = 15.0, 4.6 Hz, 1H), 2.95 (dd, *J* = 15.0, 7.2 Hz, 1H), 2.45 (t, *J* = 7.3 Hz, 2H), 1.97 – 1.84 (m, 2H), 1.52 – 1.39 (m, 2H), 1.32 (p, *J* = 7.4 Hz, 2H). ¹³C NMR (125 MHz, D₂O) δ 169.23, 139.44, 114.40, 53.67, 52.22, 32.41, 31.35, 30.90, 27.87, 26.94. HRMS (DART) calcd. For C₁₀H₂₀NO₂S (M⁺+1) 218.1215, found 218.1217.

Synthesis of ncAA 7:



Synthesis: In an oven dried round-bottomed flask compound **4ai** (1 g, 3.21 mmol) was dissolved in 15 mL dry THF and LiBr (439 mg, 5.1 mmol) was added to it, and the solution was stirred overnight. The THF was then removed under reduced pressure, and the crude was loaded directly on silica-gel column for purification.

7i. Yield 1.03 g, 90%. ^1H NMR (500 MHz, Chloroform-*d*) δ 5.34 (d, J = 8.0 Hz, 1H), 4.54 (d, J = 7.4 Hz, 1H), 3.77 (s, 3H), 3.62 (t, J = 6.3 Hz, 2H), 2.97 (qd, J = 13.9, 5.1 Hz, 2H), 2.69 (t, J = 7.0 Hz, 2H), 2.01 (p, J = 6.6 Hz, 2H), 1.45 (s, 9H). ^{13}C NMR (125 MHz, CDCl_3) δ 171.40, 162.97, 53.32, 52.56, 43.16, 34.57, 32.04, 29.68, 28.27.

Boc-deprotection of **7i** (450 mg, 1.31 mmol) was performed as described earlier.

7. Yield 417mg, 89%. ^1H NMR (500 MHz, Deuterium Oxide) δ 4.27 (dd, J = 7.3, 4.5 Hz, 1H), 3.74 (d, J = 1.3 Hz, 3H), 3.58 (t, J = 6.3 Hz, 2H), 3.15 – 3.05 (m, 1H), 3.02 – 2.95 (m, 1H), 2.62 (t, J = 7.1 Hz, 2H), 1.97 – 1.89 (m, 2H). ^{13}C NMR (125 MHz, D_2O) δ 169.18, 53.74, 52.21, 43.53, 31.07, 30.95, 28.58.

NMR spectra of various compounds:

