

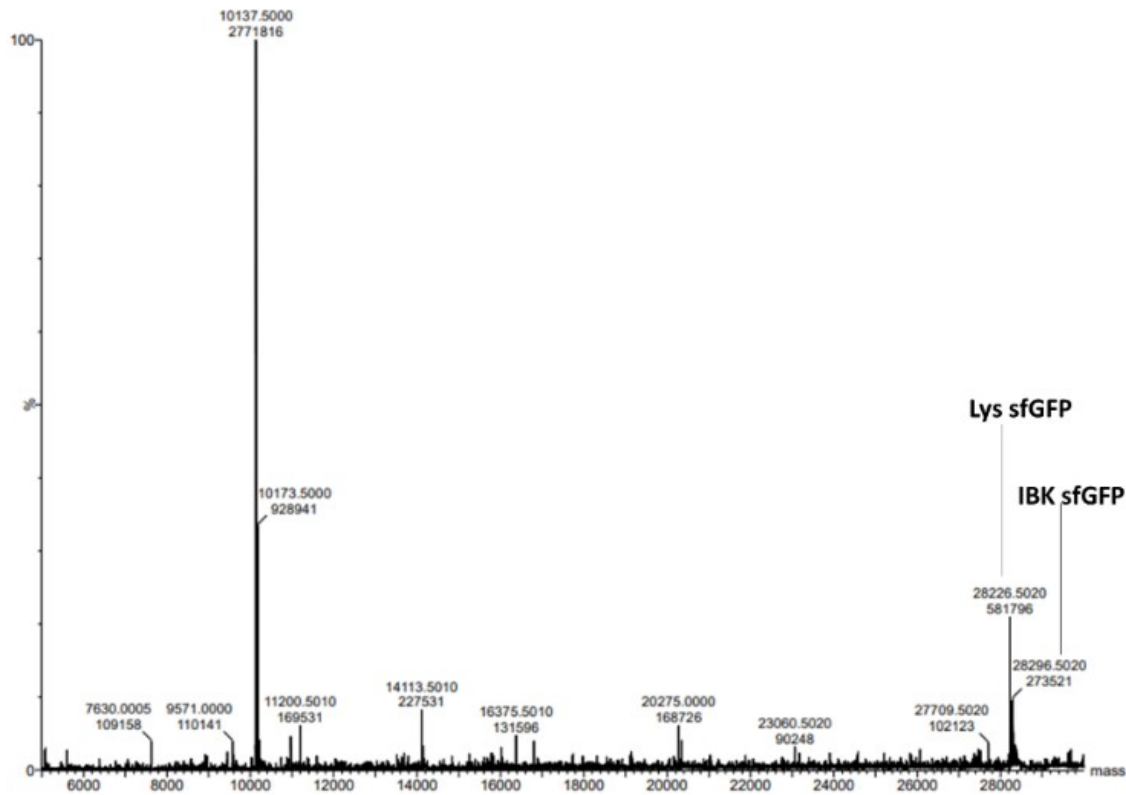
**Supplementary Information for**

**Genetic encoding of isobutyryl-, isovaleryl-, and  $\beta$ -hydroxybutryl-lysine in *E coli*.**

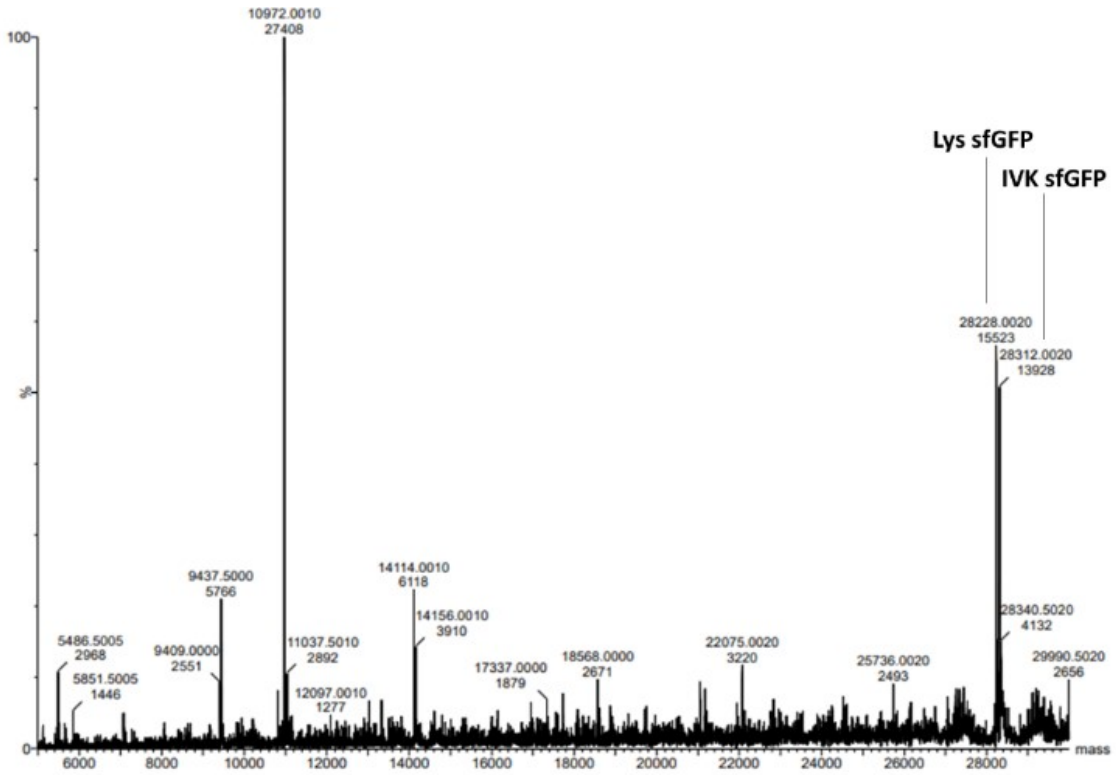
Jayani A. Christopher, Sahan A. Galbada Liyanage, Eve M. Nicholson, William D. Kinney, T. Ashton Cropp\*

**Figure S1. Mass spectrometry of isolated His6-sfGFP variants.** ESI- MS spectra of A) IBK B) IVK C) HBK His6-sfGFP. The masses correspond to the masses of His6-sfGFP with the acylated lysines and lysine with no acylation. The masses are calculated by the protein sequence minus the N-terminal methionine (-131) and GFP chromophore maturation (+20). Observed masses: (A) IBK, 28296 (calc 28299), (B) IVK, 28312 (calc 28313), (C) HBK, 28312 (calc 28314).

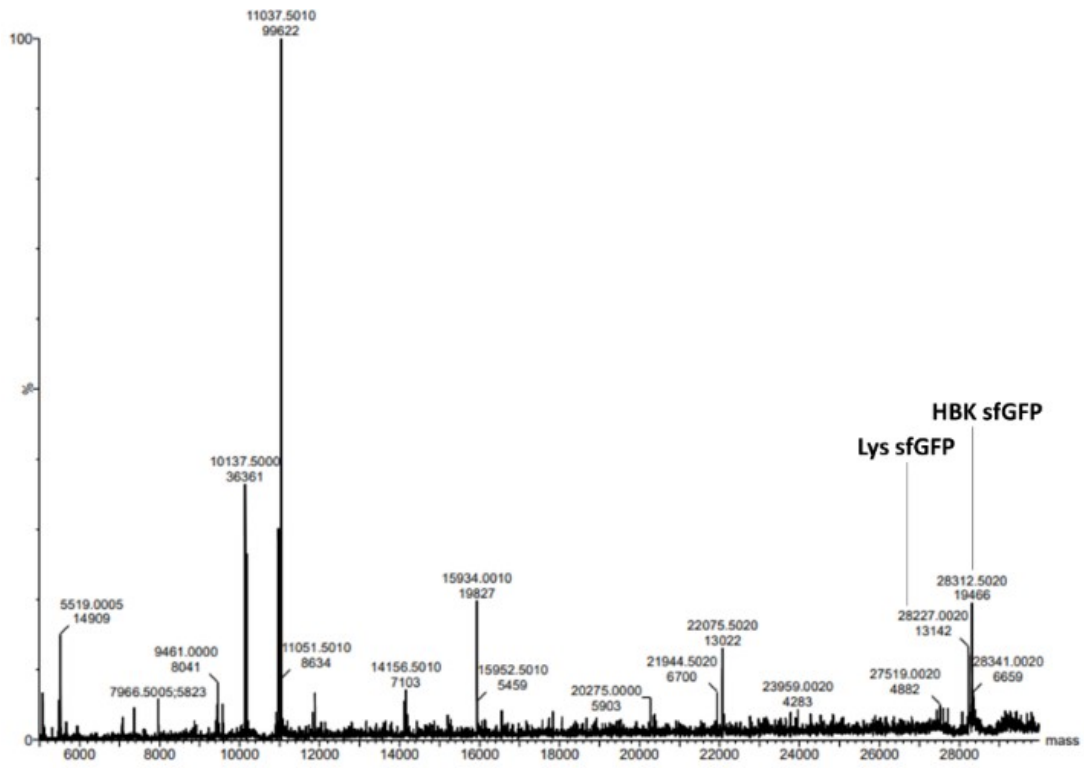
A.



B.

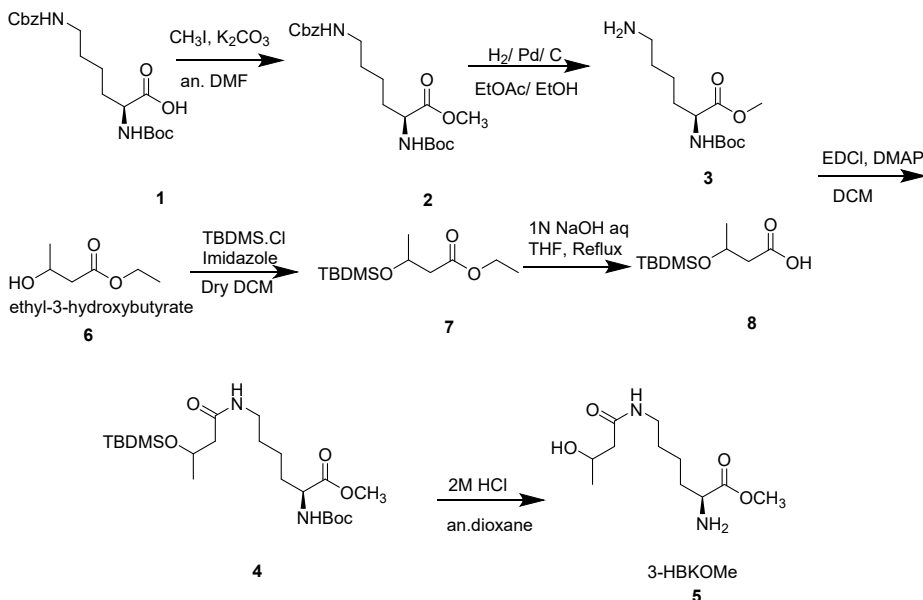


C.



## General synthetic methods of amino acids.

All starting materials were of ACS grade or better. TLC plates used were aluminum backed Silica XG TLC Plates w/ UV254 coating. Plates were visualized with UV shadowing or ninhydrin staining. Purity of all compounds was determined by TLC and NMR. While we did not observe any, when working with



**Figure S2.** Synthesis of 3-HBKOMe (5)

amines, the researcher is advised to be aware of the potential toxicity of nitrosamines.<sup>1</sup>

### Synthesis of methyl N6-((benzyloxy)carbonyl)-N2-(tert-butoxycarbonyl)-L-lysinate (2)

Boc-Lys(Z)-OH (1) (7.23 g, 19.0 mmol) was dissolved in 50 ml of dry DMF.  $\text{K}_2\text{CO}_3$  (5.25 g, 38.0 mmol) was added to the reaction flask, and it was loaded with a stir bar. The reaction flask was cooled to 0 °C in an ice bath. To the cooled reaction mixture, methyl iodide (1.05 ml, 24.7 mmol) was added and warmed to room temperature overnight. The reaction was monitored by TLC for the consumption of the starting material. The reaction mixture was diluted with EtOAc (120 ml) and washed with water (100 ml) and brine (2×100 ml). The organic layer was dried over  $\text{MgSO}_4$  and concentrated under reduced pressure to obtain the product methyl (2) as a viscous light-yellow oil (5.98 g, 80 %).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  1.30 - 1.38 (m, 2H), 1.43 (s, 9H), 1.45 - 1.58 (m, 2H), 1.60 - 1.70 (m, 2H), 3.15 - 3.23 (q,  $J = 6.6$  Hz, 2H), 3.72 (s, 3H), 4.21 - 4.35 (m, 1H), 5.09 (s, 2H), 7.27 - 7.37 (m, 5H).

### Synthesis of methyl (tert-butoxycarbonyl)-L-lysinate (3)

Compound **2** (5.70 g, 14.45 mmol) was dissolved in EtOAc (80 ml) / EtOH (20 ml). Palladium on carbon (5% Pd on carbon, 450 mg) was added to the reaction flask. The reaction flask was sealed and H<sub>2</sub> gas was bubbled through the reaction mixture while stirring the reaction 6 hours. TLC confirmed the total consumption (complete transformation) of the starting material to the product (Solvent EtOAc: Hexane 1:1). Palladium on carbon was removed by filtering the reaction mixture through a pad of Celite. The filtrate was concentrated under reduced pressure to obtain the product (3.75 g, 99.0 %) as a clear dense oil (**3**), which was used in the next step without further purification.

<sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>) δ 3.92 (q, J = 8.2 Hz, 1H), 3.62 (s, 3H), 2.53 – 2.48 (m, 2H), 1.67 – 1.50 (m, 2H), 1.37 (s, 9H), 1.35 – 1.26 (m, 4H).

#### Synthesis of ethyl 3-((tert-butyldimethylsilyl)oxy)butanoate (**7**)

Tert-butyldimethylsilyl chloride (7.8 g, 52 mmol, 1.30 equiv.) was added to a round bottom flask, which contained imidazole (6.8 g, 100 mmol). To the round bottom flask, dry CH<sub>2</sub>Cl<sub>2</sub> (25 ml) was added, and the mixture was stirred for 5 mins. Ethyl-3-hydroxybutyrate (**6**) (5.2 ml, 40 mmol.) was added to the reaction flask dropwise and the reaction was stirred for 2 hours. After the completion of the reaction, white precipitate was formed, and water (25 ml) was added and stirred for another 10 mins. Layers were separated in a separatory funnel and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3×20 ml). Combined organic layers were dried over MgSO<sub>4</sub> and concentrated in a rotary evaporator to obtain the final product as a clear oil (**7**) (9.85 g, 99.9%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 0.02 (s, 6H), 0.80 (s, 9H), 1.13 (d, J = 6.1 Hz, 3H), 1.19 (t, J = 7.2 Hz, 3H), 2.35 (m, 2H), 4.00 – 4.12 (m, 2H), 4.16 - 4.26 (m, 1H).

#### Synthesis of 3-((tert-butyldimethylsilyl)oxy)butanoic acid (**8**)

(**7**) (3.72 g, 15.099 mmol, 1.00 equiv.) was dissolved in MeOH (35 ml) and added to 1M NaOH (65 ml). The reaction stirred overnight and MeOH evaporated under reduced pressure and the remaining aqueous layer acidified to pH 4 with dil. HCl. The mixture was extracted with EtOAc (3×100 ml). Organic phases were combined and dried over MgSO<sub>4</sub>. It was concentrated under reduced pressure to obtain the product 3-((tert-butyldimethylsilyl)oxy)butanoic acid as a clear oil (**8**) (3.01 g, 91.5 %)

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 0.00 (s, 6H), 0.79 (s, 9H), 2.25 (m, 2H), 3.88 – 4.06 (m, 2H), 4.09 - 4.21 (m, 1H).

#### Synthesis of methyl N<sub>2</sub>-(tert-butoxycarbonyl)-N<sub>6</sub>(3((tertbutyldimethylsilyl)oxy)butanoyl)-L-lysinate (**4**)

(**3**) (3.75 g, 14.40 mmol, 1.00 equiv.) was dissolved in dry CH<sub>2</sub>Cl<sub>2</sub> (40 ml) and to the flask, **8** (6.29 g, 28.80 mmol, 2.00 equiv.) was added while stirring. In a separate flask, EDCI (3.588 g, 18.72 mmol, 1.30 equiv.) and DMAP (70.36 mg, 0.576 mmol, 0.04 equiv.) were mixed in dry CH<sub>2</sub>Cl<sub>2</sub> (10 ml). This was added to the initial reaction mixture dropwise and the reaction stirred while monitoring the total consumption of the starting material (methyl (tert-butoxycarbonyl)-L-lysinate) (**3**). The CH<sub>2</sub>Cl<sub>2</sub> layer was removed by reduced pressure and the crude was dissolved in EtOAc (75 ml). The organic layer was washed with water (2×60 ml) and brine (60 ml). The organic layer was dried over MgSO<sub>4</sub> and concentrated to obtain the crude product. Crude was further purified through silica column chromatography (starting with a gradient of EtOAc: Hexanes 1:3 and changing the gradient to 3:1). Combining the fractions and rotary evaporation resulted in the white crystalline product (3.10 g, 46.8 %).

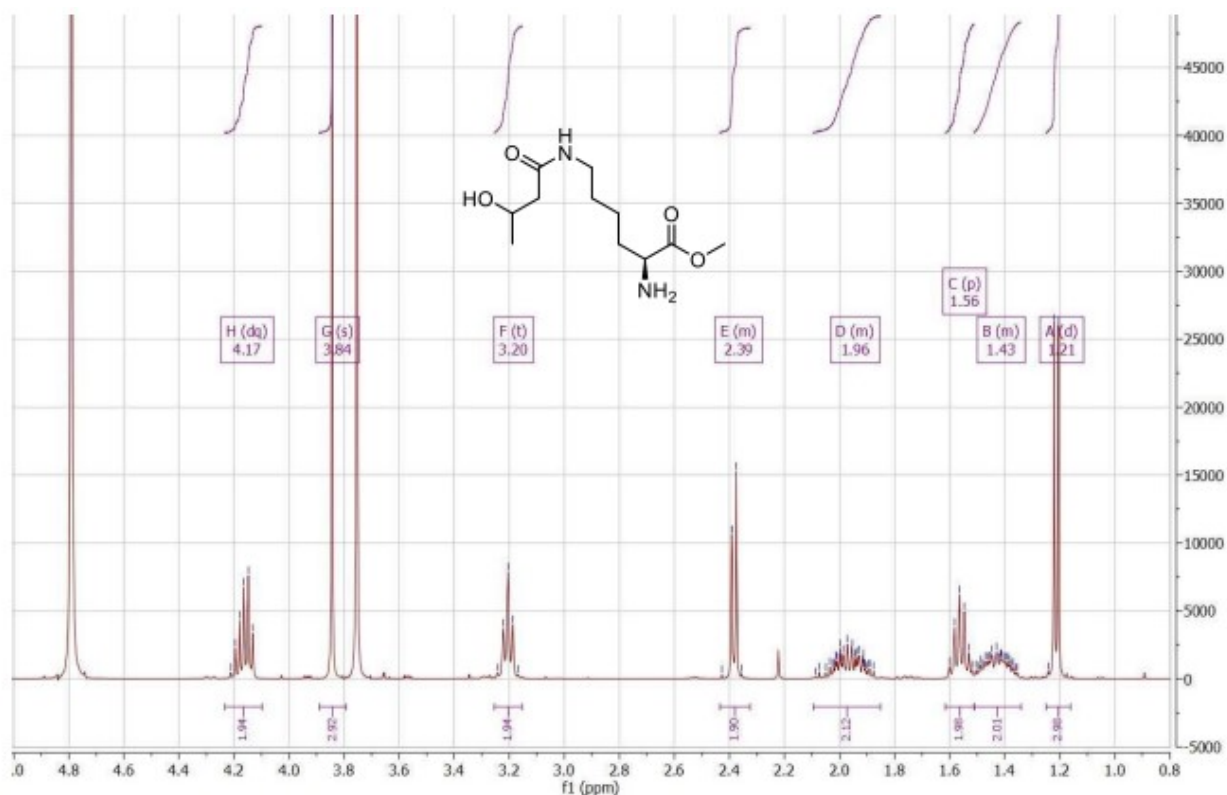
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.32 – 4.24 (m, 1H), 4.21 (td,  $J = 6.2, 4.3$  Hz, 1H), 3.72 (s, 3H), 3.22 (ddq,  $J = 67.3, 12.7, 6.4$  Hz, 2H), 2.43 – 2.24 (m, 2H), 1.85 – 1.61 (m, 2H), 1.57 – 1.47 (m, 2H), 1.43 (s, 9H), 1.40 – 1.31 (m, 2H), 0.87 (s, 9H), 0.07 (d,  $J = 7.6$  Hz, 6H).

#### Synthesis of methyl N6-(3-hydroxybutanoyl)-L-lysinate (3-HBKOMe) (**5**)

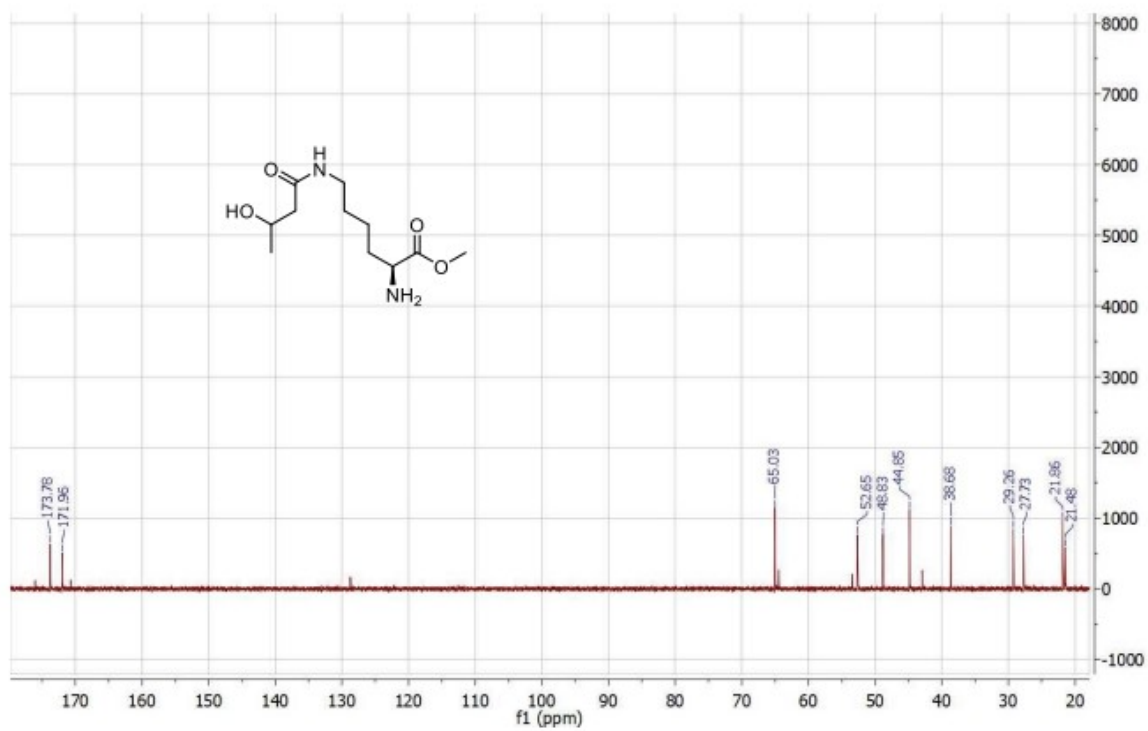
Compound **4** (775 mg, 1.682 mmol, 1.00 equiv.) was dissolved in 25 ml of 4M HCl in dioxane. The reaction was stirred for 4 hours while monitoring by TLC for the conversion of the starting material. The reaction was stopped by adding diethyl ether to precipitate the product. The organic layer was decanted, and the above step was repeated several times remove dioxane. The remaining white precipitate was dried more under reduced pressure (350 mg, 71.6 %).

$^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.18 – 4.00 (m, 2H), 3.77 (s, 2H), 3.21 – 3.09 (m, 2H), 2.36 – 2.28 (m, 2H), 2.00 – 1.79 (m, 2H), 1.49 (p,  $J = 6.5$  Hz, 2H), 1.37 (m, 2H), 1.14 (d,  $J = 6.3$  Hz, 3H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{D}_2\text{O}$ )  $\delta$  173.78, 171.96, 65.03, 52.65, 48.83, 44.85, 38.68, 29.26, 27.73, 21.86, 21.48.



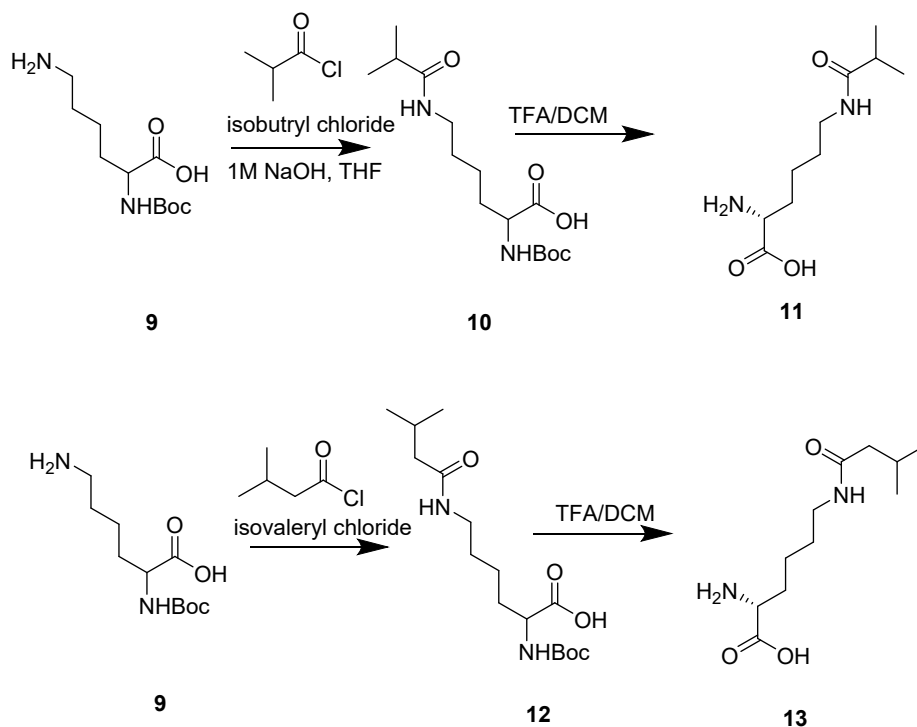
$^1\text{H}$  NMR of 3-HBKOMe (**5**)



<sup>13</sup>C NMR of 3-HBKOMe (5)



## Synthesis of IBK



**Figure S3.** Synthesis of IBK (11) and IVK (13)

### N2-(tert-butoxycarbonyl)-N6-isobutyryl-L-lysine (10)

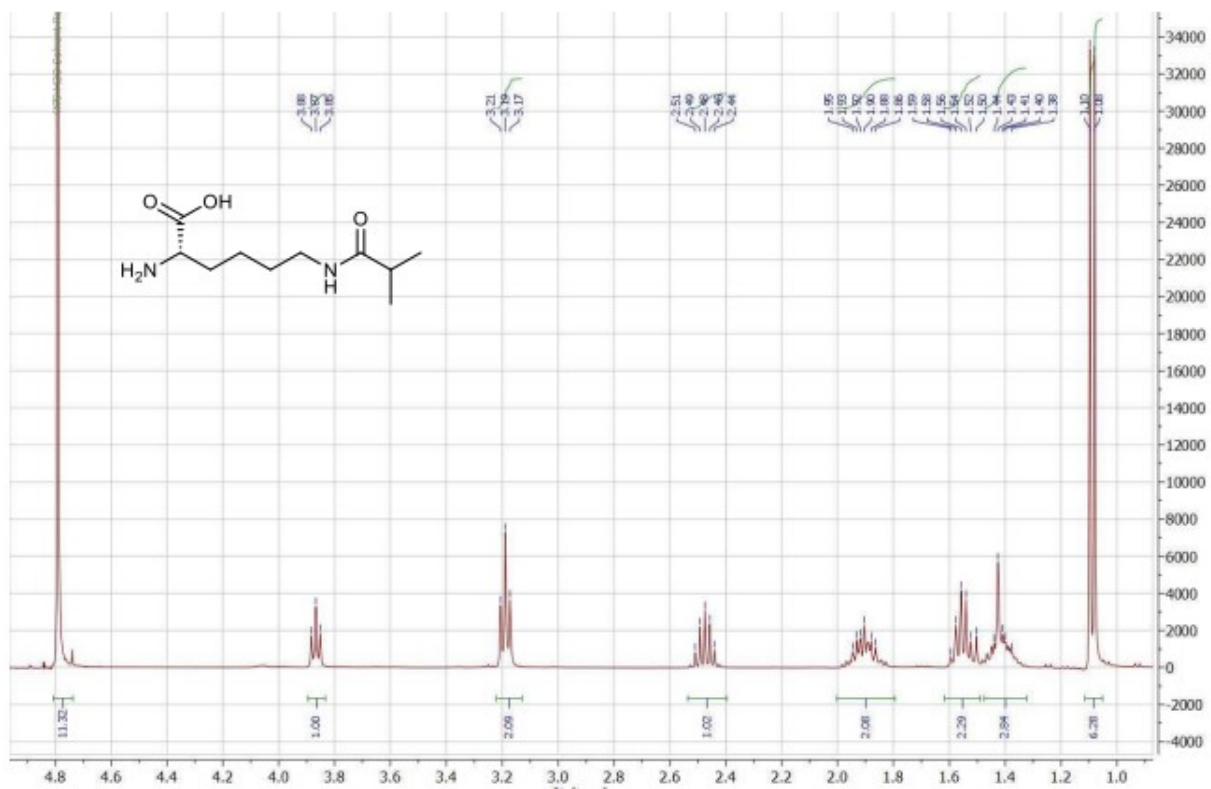
Boc-lysine (9) (5 g, 20.30 mmol, 1.0 Equiv.) was dissolved in a 1M NaOH (50ml)/THF (50 ml) solvent mixture and was cooled to 0 °C. The reaction flask was kept under a nitrogen atmosphere and isobutyryl chloride (2.50 ml, 26.39 mmol, 1.30 equiv.) was added to the mixture dropwise while stirring. The reaction was stirred overnight at room temperature. It was washed with ice-cold diethyl ether (100 ml), and the aqueous phase acidified to pH 4 by adding 1M HCl. The acidified aqueous layer was extracted with ice-cold EtOAc (3×200 ml). The organic phases were combined and dried over MgSO<sub>4</sub>, and the solvent evaporated under reduced pressure to obtain the crude product (10) as a viscous light-yellow oil (5.8 g) which was directly used in the next step without any further purification.

### Deprotection step to obtain N6-isobutyryl-L-lysine (IBK)

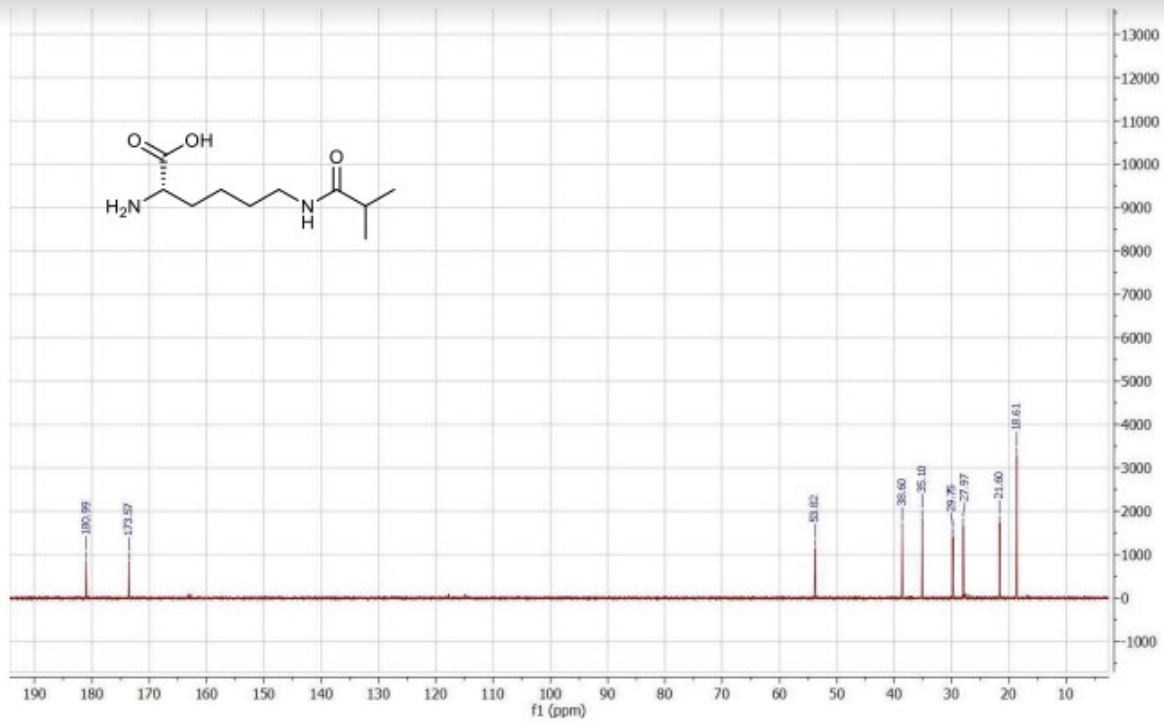
Trifluoroacetic acid (10 ml) was added to a solution of 10 (5.8 g) in DCM (50 ml) while stirring. The reaction mixture was stirred for 4 hours until the starting material disappeared by TLC. The was removed under reduced pressure and the oily crude was dissolved in a minimum amount of DCM. The product was precipitated by adding diethyl ether (50 ml) to the flask and decanting the solvents. The product was further kept under a high vacuum to remove residual solvent and reactants. IBK was obtained as a white precipitate (11) (4.10 g, 93.4%)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 1.09 (d, J = 6.9 Hz, 5H), 1.28 – 1.48 (m, 2H), 1.49 – 1.64 (m, 2H), 1.77 – 1.97 (m, 2H), 2.48 (p, J = 6.9 Hz, 1H), 3.19 (t, J = 6.9 Hz, 2H), 3.69 – 3.76 (m, 1H).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) δ 180.99, 173.57, 53.82, 38.60, 35.10, 29.75, 27.97, 21.60, 18.61



$^1\text{H}$  NMR of IBK (**11**)



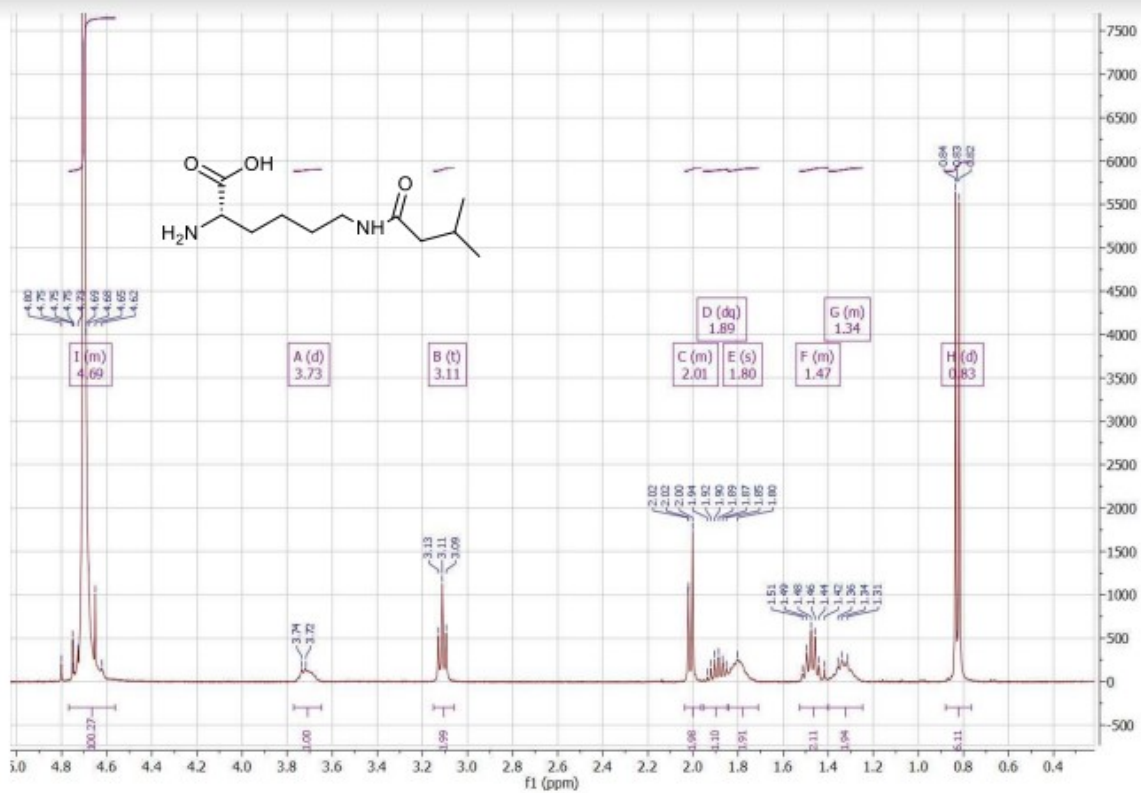
<sup>13</sup>C NMR of IBK (11)

*Synthesis of IVK*

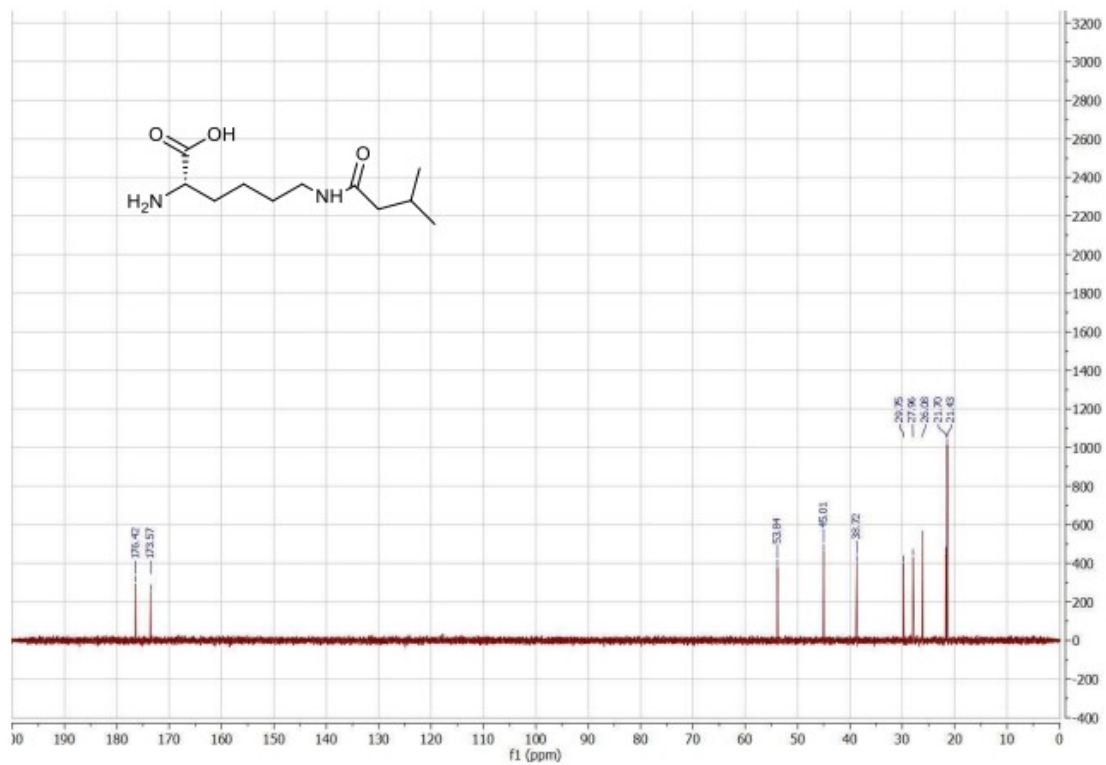
IVK (**13**) was synthesized by following the identical route as the above synthesis of synthesis of IBK. Instead of isobutyryl chloride, isovaleryl chloride was used in the first step.

$^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.83 – 4.59 (m, 3H), 3.73 (d,  $J = 7.5$  Hz, 1H), 3.11 (t,  $J = 6.9$  Hz, 2H), 2.05 – 1.98 (m, 2H), 1.89 (dq,  $J = 13.9, 6.6$  Hz, 1H), 1.80 (s, 2H), 1.54 – 1.39 (m, 2H), 1.35 (d,  $J = 7.4$  Hz, 1H), 1.31 (s, 1H), 0.83 (d,  $J = 6.6$  Hz, 5H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{D}_2\text{O}$ )  $\delta$  176.42, 173.57, 53.84, 45.01, 38.72, 29.75, 27.96, 26.08, 21.70, 21.43



<sup>1</sup>H NMR of IVK (13)



<sup>13</sup>C NMR of IVK (13)

### *Plate assay to confirm the Mb PylRS wildtype synthetase*

DH10 $\beta$  electrocompetent cells were transformed with pRepPylT (Plasmid with T7 RNA polymerase with a stop codon and GFP UV) and pBKMbWt (Plasmid with PylRS) and plated on LB agar with 15 $\mu$ g/mL tetracycline (Tet) and 50  $\mu$ g/mL kanamycin (Kan). A single colony was picked, and 5 mL culture was inoculated and grown at 37 °C. Using 5  $\mu$ L of the culture plates containing 4 mM acylated lysines Bock, IBK, IVK, 3-HBKOME, no amino acid and 0.02 % arabinose were streaked and incubated at 37 °C overnight. The plates were observed using Spectroline UV Transilluminator.

### *His6-sfGFP expression, purification and mass spectrometry*

DH10 $\beta$  electrocompetent cells were transformed with pPylTsfGFPY151TAG (Plasmid with sfGFP with a stop codon at 151 position) and pBKMbWt and plated on LB agar with 15 $\mu$ g/mL Tet and 50  $\mu$ g/mL Kan. A single colony was picked, and 20 mL culture was inoculated and grown at 37 °C. Using the overnight starter culture 200 mL cultures were inoculated by doing a 50-fold dilution and grown until the OD 600 was 0.6. The cultures were then treated with 2 mM acylated lysines and 0.002% arabinose and grown overnight. The cultures were harvested at 6000 rpm for 15 minutes at 4 °C. The cell pellets were washed with 30 mL Ni-NTA binding buffer (50 mM Tris base, 300 mM NaCl, 20 mM Imidazole, pH =8). The cell pellets were then lysed with cell lysis buffer (0.1 mg/ mL lysozyme in Ni-NTA binding buffer) and incubated in ice for 30 minutes. The cells were further lysed by sonicating. The sonication of cells was done in five rounds of 30-second bursts with 1-minute cooling intervals on the ice at 40% amplitude. The lysed cells were centrifuged at 10000 rpm for 15 minutes at 4 °C. Ni-NTA resin purification was conducted in batch mode. For each cell lysate, 300  $\mu$ L of Ni-Resin slurry (50 % resin in the slurry) was used. The resin was washed with 3 $\times$ 10 mL of Ni-NTA binding buffer. The washed resin was loaded with the cell lysate, and it was incubated at 4 °C in a rotator for 1 hour. The flow through was decanted to a vial. The resin was washed with a 5 $\times$ 30 mL binding buffer with. The resin was eluted with the elution buffer (50 mM Tris base, 300 mM NaCl and 500 mM imidazole, pH 8.00) in 5 $\times$ 200  $\mu$ L fractions. The collected fractions were run on an SDS-PAGE gel to detect the correct size of the intact proteins. The elutions of sfGFP protein showing the most protein by SDS-PAGE were combined, and desalted by dialysis to 0.1% acetic acid in water. The resulting sample was then directly injected onto a Synapt ESI -LC/MS.

*Sequence of His6-sfGFP protein. Site of mutation (Y151) shown as X:*

MGGSHHHHHHGMASMMMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFIC  
TTGKLPVPWPTLVTTLTYGVCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEV  
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNESHNVXITADKQKNGIKANFKIRHNVEDGSVQ  
LADHYQQNTPIGDGPVLL PDNHYLSTQSVLSKDPNEKRDHMLLEFVTAAGITHGMDELYK

The masses were calculated by deducting the mass of N-terminal methionine (-131) and the oxidized chromophore (-20).

### *Histone H3 expression and purification*

DH10 $\beta$  electrocompetent cells were transformed with pSAG\_h3K9TAGPylT (Plasmid with histone with a stop codon at 9<sup>th</sup> position) and pBKMbWt and plated on LB agar with 15 $\mu$ g/mL Tet and 50  $\mu$ g/mL Kan. A single colony was picked, and 20 mL culture was inoculated and grown at 37 °C. Using the overnight starter culture 50 mL cultures were inoculated by doing a 50-fold dilution and grown until the OD 600 was 0.6. The cultures were then treated with 4 mM acylated lysines and 0.2% arabinose and grown overnight. The cultures were harvested at 6000 rpm for 15 minutes at 4 °C. The cell pellets were resuspended in 5 mL Ni-

NTA binding buffer (50 mM Tris base, 300 mM NaCl, 20 mM imidazole, pH =8). The sonication of cells was done in five rounds of 30-second bursts with 1-minute cooling intervals on the ice at 40% amplitude. The lysed cells were centrifuged at 10000 rpm for 15 minutes at 4 °C. The insoluble and soluble portions were separated, and the insoluble cell pellet was washed with 5 mL H<sub>2</sub>O. The cell pellet was resuspended in 5 mL Ni-NTA denaturing binding buffer (50 mM Tris base, 300 mM NaCl, 20 mM imidazole, 8 M urea pH =8) and incubated at room temperature for 10 minutes. The lysates were centrifuged at 10000 rpm for 15 minutes at RT. Ni-NTA resin purification was conducted in batch mode. For each cell lysate, 200µL of Ni-Resin slurry (50 % resin in the slurry) was used. The resin was washed with 3×10 mL of Ni-NTA denaturing binding buffer. The washed resin was loaded with the cell lysate, and it was incubated at 4 °C in a rotator for 1 hour. The flow-through was decanted to a tube. The resin was washed with a 3×10 mL the same denaturing binding buffer for 10 minutes each at room temperature. The resin was eluted with the elution buffer (50 mM Tris base, 300 mM NaCl and 500 mM imidazole, 8M Urea, pH 8.00) in 4×50µL fractions. After each elution the resin was rotated for 10 minutes at RT. The collected fractions were run on an SDS-PAGE gel to detect the correct size of the intact proteins.

*Sequence of histone protein. Site of mutation (K9) shown as X:*

MGHHHHHLENLYFQGARTKQTARXSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTV  
ALREIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAI  
HAKRVTIMPKDIQLARRIGERA

## References

1. Beard, J.C, and Swager, T.M. An Organic Chemist's Guide to N-Nitrosamines: Their Structure, Reactivity, and Role as Contaminants. *J. Org. Chem.* 2021, 86, 3, 2037–2057