Supplementary Materials: Incorporation of Amino Acids with Long-Chain Terminal Olefins into Proteins


Supplemental Materials and Methods

Plasmid Sequences

Suppression Plasmid

The pJZ plasmid was recently used for efficient incorporation of noncanonical amino acids into proteins [1] using a PylRS variant. The same setup was used for suppression of target genes with SacRS. Relevant sequence part (pylT-pylS with promoter/terminator):

CGCACCCGGTGcatactcccactccccctgtgacacttatctcaatagactatatctcagggagagagaGATACCGGGATCC
GGAAAACCTGATCACTGTAGATCGAATGGACTCTAAATCCGTTCAGCCGGTTAGATTCCGG
GGTTTCCGCCACTAGGTATTAAAAACTAAGGAGGTTAATTAA
ATGAGCGCTTGGAGCCACCCGC
AGTTCGAAAAACTCGAG
ATGGATAAAAAACCACTAAACACTCTGATATCTGCAACCGGG
CTCTGGATGTCCAGGAACACTCTGAGAAATGAGATTTCCTGAATTCCGGCAAGCCTTTCAGGGAGCTTGAGTCCGAATTGCTCTCTCGCAGA
AAAAGACCTGCAGCAGATCTACGGGAAGAAAGGGAGAATTATCTGGGGAAACTCGAGCGTGAAATTACCGGTTCTTTGTGGACAGGGGTTTTCTGGAAATAAAATCCCCGATCC
TGATCCCTCTTGAGTATATCGAAAGGATGGGCATTGATAATGATACCGAACTTTCAAACGATCTTCAGGGTTGACAAGAACTTCTGCCTGAGACCCATGCTTGCTCCAAACCTTTACAACTACCTGCGCAAGCTTGACAGGGCCCTGCCTGATCCAATAAAAATTTTTGAAATAGGCCCATGCTACAGAAAAGAGTCCGACGCAAAGAACCTCGAAGAGTTTACCATGCTGAACTTCTGCCAGATGGGATCGGGATGCACACGGGAAAATCTTGAAAGCATAATTACG
GACTTCCTGAACCACCTGGGAATTGATTTCAAGATCGTAGGCGATTCCTGCATGGTCTa

TGGGGATACCCTTGATGTAATGCACGGAGACCTGGAATCTTCTCTGCAGTAGTCGGA
CCCATAACCGGCTTGACCAGGAATGCGGTTAATGATGAAACCCCTGGATAGGGGCAGGTTTCGGGCTCGAACGCCTTCTAAAGGTTAAACACGACTTTAAAAATATCAAGAGAGCTGCAAGGTCCGAGTCTTACTATAACGGGATTTCTACCAACCTGTAA

Target Gene Plasmid

The model target gene was cloned on a commercially available high-copy vector pQE80L (Qiagen, Hilden, Germany) with an additional plasmid-based lac repressor for inducible target protein production in any host strain.

Relevant sequence part (EGFP with promoter/terminator):

CTCGAGaatactataaattaattttatatgtgacgacccgataaaactgactttaatttttttagctgttatttatttatttattattttttatttattttatttattttttatttttttatttttatttattattattttattttattttattttttattttattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Strain Genotypes

Table S1. Bacterial strains used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21(DE3)</td>
<td>fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS</td>
<td>NEB, Ipswich, USA</td>
</tr>
</tbody>
</table>

Primer Sequences

Table S2. Primers used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>MmPyIRS_Y306A_for</td>
<td>GCTGCTCCAACCTTGCCAACGCAACTCTGCAGCAAG</td>
<td>mutagenesis Y306A of MmPyIRS(Y384F)</td>
</tr>
<tr>
<td>MmPyIRS_Y306A_rev</td>
<td>CTTGCCAGCTTGCAAGGTTGGGCAACGCAAGCAAGC</td>
<td>mutagenesis Y306A of MmPyIRS(Y384F)</td>
</tr>
<tr>
<td>PyIRS_L309A_for</td>
<td>CAAACCTTGCAACTACGCGCAAGCTTGACAGG</td>
<td>mutagenesis L309A of MmPyIRS(Y384F)</td>
</tr>
<tr>
<td>PyIRS_L309A_rev</td>
<td>CCGCTCAAGCTTGCGCCGCTAGTTGGCAAGCTTTG</td>
<td>mutagenesis L309A of MmPyIRS(Y384F)</td>
</tr>
<tr>
<td>MmPyIRS_L309AvY306A_for</td>
<td>CAAACCTTGCAACTACGCGCAAGCTTGACAGG</td>
<td>mutagenesis L309A of MmPyIRS(Y306AY384F)</td>
</tr>
<tr>
<td>MmPyIRS_L309AvY306A_rev</td>
<td>CCGCTCAAGCTTGCGCCGCTAGTTGGCAAGCTTTG</td>
<td>mutagenesis L309A of MmPyIRS(Y306AY384F)</td>
</tr>
<tr>
<td>PyIRS_C348A_for</td>
<td>CATGCTGAACCTTGCCAAGATGGGATCGG</td>
<td>mutagenesis C348A of MmPyIRS(Y384F) or MmPyIRS(L309A/Y384F)</td>
</tr>
<tr>
<td>PyIRS_C348A_rev</td>
<td>CCGATCCCACATCGGCGCCGATGGGATCGG</td>
<td>mutagenesis C348A of MmPyIRS(Y384F) or MmPyIRS(L309A/Y384F)</td>
</tr>
</tbody>
</table>

Detailed Syntheses

All reagents are commercial grade (purchased from Sigma-Aldrich) and were used as received. All reactions were performed under inert atmosphere using anhydrous solvents which were dried and distilled before use. Thin-layer chromatography (TLC) and flash chromatography separations were respectively performed on precoated silica gel 60 F 254 plates (Merck, 0.25 mm) and on Merck silica gel 60 (230–400 mesh). 1H-NMR spectra were recorded at 400 MHz; shifts are relative to internal TMS. 13C-NMR spectra were recorded at 100 MHz.

Synthesis of Hek (1)

2,5-Dioxopyrrolidin-1-yl hept-6-enoate (5)

A mixture of 6-heptenoic acid 3 (0.53 g, 4.13 mmol, 1 eq), N-hydroxysuccinimide (0.47 g, 4.13 mmol, 1 eq) and DCC (0.85 g, 4.13 mmol, 1 eq) was stirred at room temperature for 48 h. The resulting white precipitate was filtered and washed with DMF. The filtrate was diluted with AcOEt and washed with water. The aqueous phase was extracted with AcOEt. The organic layers were combined, dried over Na2SO4, filtered and concentrated under reduced pressure. The residue was purified on silica gel flash chromatography (Heptane/AcOEt, 70:30) to give 2,5-dioxopyrrolidin-1-yl hept-6-enoate 5 as a white oil (0.707 g, 76%).

1H-NMR (400 MHz, CDCl3, δ in ppm) δ: 5.83–5.74 (m, 1H, CH=CH2), 5.05–4.96 (m, 2H, CH=CH2), 2.83 (s, 4H, NCOCH3), 2.60 (t, J = 8.0 Hz, 2H, COCH2CH2), 2.13–2.07 (m, 2H, CH2CH2CH2), 1.76 (q, J = 7.6 Hz, 2H, CH2CH2CH2), 1.51 (q, J = 7.6 Hz, 2H, CH2CH2CH2). 13C-NMR (100 MHz, CDCl3,
δ in ppm \(\delta\): 170.5 (2C, NCO), 169.3 (OCO), 138.6 (CH=CH₂), 115.2 (CH=CH₂), 36.0 (CH₃CH=CH₂), 30.5 (OCOCH₂), 27.6 (CH₂CH₂CH₂CH₂), 25.8 (2C, COCH₂CH₂CO), 24.2 (CH₂CH₃CH₂CH₂).

\(\text{(S)-2-Boc-amino-6-(hept-6-enamido)hexanoic acid (7)}\)

A mixture of 2,5-dioxopyrrolidin-1-yl hept-6-enoate 5 (0.718 g, 3.18 mmol, 1 eq), Boc-Lysine 6 (0.778 g, 3.18 mmol, 1 eq) in aqueous NaHCO₃ (1 N, 20 mL) was stirred at room temperature for 20 h and then acidified to pH 2 by adding 1 N HCl. The reaction mixture was then extracted with AcOEt. The combined organic layers were washed with HCl 1 N, H₂O, dried over Na₂SO₄, filtered and concentrated under reduced pressure to give \(\text{(S)-2-Boc-amino-6-(hept-6-enamido)hexanoic acid (7)}\) as a colorless oil (1.08 g, 95%).

\(\text{1H-NMR (400 MHz, CDCl₃, } \delta \text{ in ppm)}: 5.86–5.76 (m, 1H, CH=CH₂), 5.79 (s broad, 1H, NH amide), 5.30 (s broad, 1H, NH Boc), 5.04–4.96 (m, 2H, CH=C₃H₂), 4.29 (s, 1H, BocNCH), 3.30–3.26 (m, 2H, CH₂C₃H₂NH), 2.20 (t, \(J = 7.6 \text{ Hz, 2H, COCH₂CH₂)}), 2.09–2.07 (m, 2H, CH₃CH=CH₂), 1.77–1.53 (m, 10H, BocNHCH₂CH₂CH₂CH₂ + COCH₂CH₂CH₂CH₂), 1.47 (s, 9H, Boc).

\(\text{13C-NMR (100 MHz, CDCl₃, } \delta \text{ in ppm)}: 175.3 (NHCO), 174.0 (COOH), 157.9 (CO Boc), 138.3 (CH=CH₂), 114.6 (CH=CH₂), 79.8 (Cq, Boc), 49.3 (BocNHCH₂), 39.2 (CH₂NH), 36.4 (COCH₂), 33.3 (CH₂CH₂CH₂), 30.8 (BocNHCH₂CH₂), 28.4 (CH₂CH₂NH), 28.3 (CH₂CH₂CH₂CH₂), 28.3 (3 × CH₃ Boc), 24.8 (COCH₂CH₂), 22.4 (BocNHCH₂CH₂CH₂).\) HRMS (ESI+): calculated for C₁₈H₃₅N₂O₅ [M + H]⁺ 357.4705, found 357.4711.

\(\text{(S)-2-amino-6-(hept-6-enamido)hexanoic acid, TFA salt (1)}\)

To a solution of \(\text{(S)-2-Boc-amino-6-(hept-6-enamido)hexanoic acid (7)}\) (1.08 g, 3.03 mmol, 1 eq) in DCM, was added TFA (2 mL). The solution was stirred 4 h at room temperature and concentrated. The residue was dissolved in water, washed with AcOEt and concentrated to give \(\text{(S)-2-amino-6-(hept-6-enamido)hexanoic acid, TFA salt (1)}\) as a white solid (0.707 g, 64%).

\(\text{1H-NMR (400 MHz, D₂O, } \delta \text{ in ppm)}: 5.92–5.82 (m, 1H, CH=CH₂), 5.06–4.96 (m, 2H, CH=CH₂), 3.97 (t, \(J = 6.4 \text{ Hz, 1H, NH₂CHCOOH)}), 3.18 (t, \(J = 6.8 \text{ Hz, 2H, CH₂CH₂NH)}), 2.22 (t, \(J = 7.6 \text{ Hz, 2H, COCH₂CH₂)}), 2.07–1.83 (m, 4H, CH₂CH₂CH₂ + NH₂CHCH₂CH₂CH₂CH₂), 1.61–1.33 (m, 8H, NH₂CHCH₂CH₂CH₂CH₂ + COCH₂CH₂CH₂CH₂), 1.51–1.01 (m, 4H, COCH₂CH₂), 1.38–1.28 (m, 3 × CH₃ Boc), 0.88 (3 × CH₃ CO).\) HRMS (ESI+): calculated for C₁₃H₂₅N₂O₃ [M + H]⁺ 257.1860, found 257.1855.
Synthesis of Pek (2)

Commercially available N-α-Boc-L-lysine 6 (2.5 g, 10 mmol, 1 eq) was dissolved in 1 N NaOH (50 mL) solution in dioxane (1:1) and cooled in an ice bath. To this pre-cooled solution was added commercially available pentenoyl chloride 8 (2.3 mL, 21.0 mmol, 2.1 eq) dropwise and the solution left to stir overnight (0 °C to r.t.). After the completion of the reaction as indicated by a TLC analysis, the solution was cooled again and acidified with 1 N aq. HCl and extracted with dichloromethane. The organic extracts were combined, dried over Na2SO4, evaporated and chromatographed to obtain the Nε-pentenoyl product 9 in quantitative yield. This product was again dissolved in dry dioxane (10 mL) and cooled to 0 °C. 4 N HCl in dioxane (10 mL) was then added to this solution drop-wise and the reaction left to stir at r.t. overnight. The solvent was then evaporated and the yellowish residue dissolved in water and washed with diethyl ether. The aqueous layer was then lyophilized and purified on reverse phase HPLC (C-18; H2O:CH3CN) to obtain the product Nε-pentenoyllysine hydrochloride salt 2 as white powder.

1H-NMR (400 MHz, D2O, δ in ppm): δ 5.88-5.79 (m, 1H, CH=CH2), 5.10–5.01 (m, 2H, CH=C), 4.03 (t, J = 6 Hz, 1H, NH2CHOH), 3.18 (t, J = 7 Hz, 2H, CH2NH), 2.32 (m, 4H, COCH2CH2 + CH2CH2CH), 1.99–1.89 (m, 2H, NH2CHCH2CH2CH3), 1.57–1.49 (m, 2H, NH2CHCH2CH2CH3), 1.47–1.27 (m, 2H, NH2CHCH2CH2CH3); 13C-NMR (75 MHz, D2O): δ 175.7 (NHCO), 172.0 (COOH), 136.7 (CH=CH2), 115.3 (CH=CH2), 52.6 (NH2CHCOOH), 38.3 (CH2NH), 34.7 (COCH3), 29.1 (CH2CH2CHCH2), 27.5 (NH2CHCH2CH2), 24.8 (CH2CH2NH), 21.2 (NH2CHCH2CH2); HRMS (ESI+): calculated for C11H21N2O3 [M + H]+ 229.1546, found 229.1549.

Synthesis of HA-4-SH

A detailed synthesis of HA-4-SH will be available elsewhere.

Aqueous Metathesis with Protected 1 and 2

Cross Metathesis of NBoc-methylester pentenoyllysine (n = 1) as well as NBoc-methylester heptenoyllysine (n = 3) has been studied with various olefin partners in the presence of Hoveyda Grubbs II in aqueous media or organic conditions. Our full study (various experimental conditions and CM partners) will be reported elsewhere.

Scheme S1. Aqueous metathesis with protected 1 and 2 with allyl alcohol as coupling partner.
The metathesis results with allyl alcohol as coupling partner are summarized in Table S3.

Table S3. Results of cross metathesis in aqueous and organic solvent systems.

<table>
<thead>
<tr>
<th>Entry</th>
<th>n</th>
<th>Solvent</th>
<th>Additive MgCl₂</th>
<th>Yield of CM Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n = 1</td>
<td>CH₂Cl₂</td>
<td>no</td>
<td>&lt;10%</td>
</tr>
<tr>
<td>2</td>
<td>n = 1</td>
<td>H₂O/tBuOH 1/1</td>
<td>1.5 eq</td>
<td>55%</td>
</tr>
<tr>
<td>3</td>
<td>n = 3</td>
<td>CH₂Cl₂</td>
<td>no</td>
<td>40%</td>
</tr>
<tr>
<td>4</td>
<td>n = 3</td>
<td>H₂O/tBuOH 1/1</td>
<td>1.5 eq</td>
<td>60%–70%</td>
</tr>
</tbody>
</table>

Yields of CM and reactions were improved up in aqueous media (H₂O/tBuOH 1/1) with Hoveyda Grubbs II catalyst (0.05 eq–0.1 eq), allyl alcohol (10 eq) and addition of MgCl₂ (1.5 eq) at 32 °C for 2 to 6 h. Surprisingly the yields were inferior when the reactions were run in CH₂Cl₂ (approximately 10%–40% compared to 55%–60% isolated yield in aqueous media), showing the need of MgCl₂ to prevent non-productive chelation.

Supplemental Figures

Figure S1. Mass analysis of TTL(221Pek) after 6xHis tag removal conjugated with HA-4-SH.

The expected mass is 31071 Da, the difference of 17 Da can be explained by a hydroxylation or oxidation during the analysis. The lower-mass peak is caused by unspecific activity of the TEV protease used to remove the 6xHis tag prior to the decoration reaction.

Supplemental References