



# Expanding the substrate scope of pyrrolysyl-transfer RNA synthetase enzymes to include non- $\alpha$ -amino acids in vitro and in vivo

In the format provided by the  
authors and unedited

## Table of Contents

|                              |    |
|------------------------------|----|
| <u>Materials</u>             | 1  |
| <u>Synthesis Notes</u>       | 3  |
| <u>Supplementary Figures</u> | 9  |
| <u>Supplementary Tables</u>  | 54 |
| <u>References</u>            | 67 |

## **Materials**

Materials were sourced from the following suppliers: **Agilent Technologies** (Santa Clara, CA):

BL21-Gold (DE3) Competent Cells; **Accela Chembio** (San Diego, CA):

(S)-6-((tert-butoxycarbonyl)amino)-2-hydroxyhexanoic acid; **Alfa Aesar** (Ward Hill, MA):

N( $\epsilon$ )-Boc-L-lysine (L-BocK), N( $\epsilon$ )-Boc-D-lysine (D-BocK), 3-Methylbenzyl bromide,

3-(Trifluoromethyl)benzyl bromide, 3-Bromobenzyl bromide, L-lysine monohydrochloride;

**AmericanBio** (Canton, MA): carbenicillin, glycerol, isopropyl  $\beta$ -D-1-thiogalactopyranoside

(IPTG), HEPES, magnesium chloride (1 M solution), spectinomycin sulfate, sodium acetate

buffer (3 M, pH 5.2), EDTA-Na (0.5 M solution, pH 8.0); **BACHEM** (Torrance, CA):

N-methyl-L-phenylalanine, N-formyl-L-phenylalanine; **BioRad** (Hercules, CA): Any kD<sup>TM</sup>

Mini-PROTEAN<sup>®</sup> TGX<sup>™</sup> Precast Protein Gels (product 4569033), 10% Mini-PROTEAN<sup>®</sup>

TBE-Urea Gel (product 4566036), Micro Bio-Spin<sup>™</sup> P-30 Gel Columns, Tris Buffer RNase-free

(product 7326250), Precision Plus Protein<sup>™</sup> Dual Color Standards (product 1610374);

**BioWorld** (Dublin, OH): Luria-Bertani broth (LB), Terrific broth (TB); **Cayman Chemical**

(Ann Arbor, MI):  $\alpha$ -mercapto-benzenepropanoic acid; **Corning** (Kennebunk, ME): 96-well plate

#3904; **Cytiva Life Sciences** (Marlborough, MA): Superdex<sup>®</sup> 75 Increase 10/300 GL column,

HiLoad<sup>®</sup> 16/600 Superdex<sup>®</sup> 200 pg column; **Decon Labs** (King of Prussia, PA): 200 proof

ethanol; **Diversified Biotech** (Dedham, MA): Breathe-Easy<sup>®</sup> membranes; **Echelon Biosciences**

(Salt Lake City, UT): malachite green solution; **Enamine** (Kyiv, Ukraine):

2-hydroxy-3-(3-(trifluoromethyl)phenyl)propanoic acid; **Fisher Scientific** (Pittsburgh, PA): agar,

sodium hydroxide, potassium hydroxide, sodium chloride, potassium chloride, calcium chloride,

dithiothreitol (DTT), 50% polyethylene glycol 3350 solution, acetonitrile Optima<sup>™</sup> LC/MS

Grade, Tris base, ethylenediaminetetraacetic acid (free acid), Pierce<sup>™</sup> 660 nm Protein Assay

Reagent; **Frontier Scientific** (Logan, UT): L-phenylalanine, L-aspartic acid, L-valine, L-tyrosine; **Honeywell** (Charlotte, NC): 1,1,1,3,3,3-Hexafluoro-2-propanol, LC-MS Grade (HFIP); **Integrated DNA Technologies** (Coralville, IA): RF31, RF32, RF33, *Ma-PylT-F*, *Ma-PylT-R*; **Invitrogen** (Waltham, MA): SYBR™ Safe DNA Gel Stain; **J.T.Baker - Avantor** (Radnor, PA): sodium phosphate, chloroform, boric acid, hydrochloric acid, dimethylsulfoxide (DMSO); **MilliporeSigma** (Burlington, MA): β-mercaptoethanol (BME), imidazole, cesium chloride, adenosine 5'-(β,γ-imido)triphosphate lithium salt hydrate (AMP-PNP), ribonuclease A from bovine pancreas (RNase A), acidic phenol (Phenol Saturated Citrate Buffered pH 4.5), ethanol, spermidine, guanosine monophosphate (GMP), bovine serum albumin (BSA), polyethylene glycol 8000, 6-(Boc-amino)hexanoic acid (BocAhx), (*S*)-(-)-3-phenyllactic acid (3-PLA), 2-benzylmalonic acid (2-BMA), 4-(Boc-amino)butyl bromide, diethyl malonate, tetrahydrofuran anhydrous, sodium hydride 60 % dispersion in mineral oil, sodium sulfate anhydrous, diethylether, Thrombin CleanCleave™ Kit, 10 kDa MWCO Amicon® Ultra-15 Centrifugal Filter Unit, Anti-FLAG® M2 Magnetic Beads, basic phenol (Phenol solution equilibrated with 10 mM Tris HCl, pH 8.0, 1 mM EDTA); **MP Biomedicals** (Irvine, CA): D-phenylalanine, glycine hydrochloride; **New England BioLabs** (Ipswich, MA): NdeI restriction enzyme, NEBuilder® HiFi DNA Assembly Master Mix, OneTaq® Quick-Load® 2X Master Mix, nucleotide triphosphate solutions, Low Range ssRNA Ladder, PURExpress® Δ (aa, tRNA) Kit, Q5® High-Fidelity 2X Master Mix, Q5® Site-Directed Mutagenesis Kit; **PepTech** (Bedford, MA): 3-Trifluoromethyl-L-phenylalanine; **Promega** (Madison, WI): RQ1 RNase-Free DNase **Qiagen** (Germantown, MD): Ni-NTA Agarose resin; **Ricca Chemical Company** (Arlington, TX): formic acid LCMS grade, triethylamine (TEA) LCMS grade; **Roche** (Basel, Switzerland): cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail; **Takara Bio** (San Jose,

CA): TALON® Metal Affinity Resin; **Teledyne ISCO** (Lincoln, NE): 65 g RediSep® Disposable Sample Load Cartridge; **Tokyo Chemical Industry** (Portland, OR): 3-phenylpropanoic acid (3-PLA).

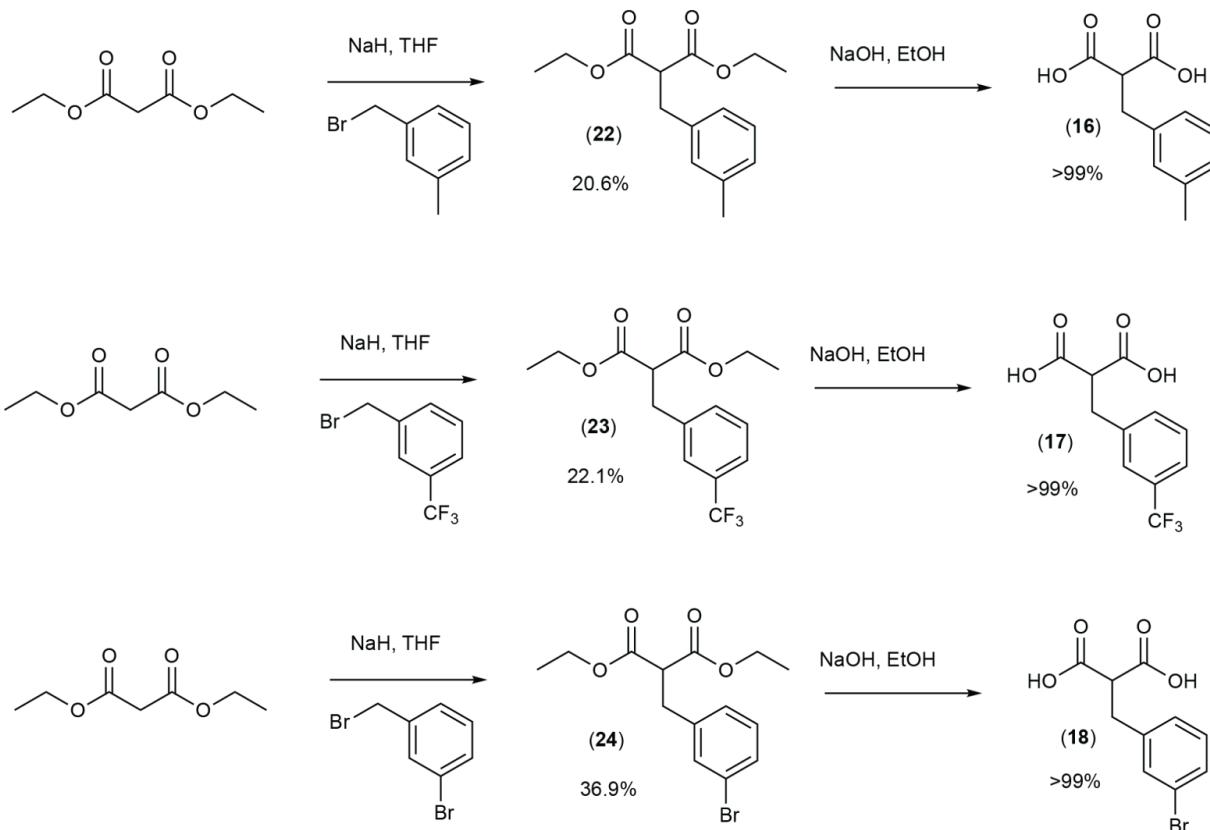
### Synthesis notes

#### **Synthesis of *meta*-substituted 2-benzylmalonates 17-19 and malonate 16.**

**General.** Alkylation reactions to synthesize **22-25** as well as hydrolysis reactions to synthesize **16-19** were based on published methods<sup>1,2</sup>. All reagents and solvents were used as received from commercial suppliers, unless indicated otherwise. Alkylation reactions were carried out with exclusion of air and moisture. Room temperature is considered 20–23 °C. Stirring was achieved with Teflon-coated magnetic stir bars. TLC was performed on glass-backed silica gel plates (median pore size 60 Å) and visualized using UV light at 254 nm or staining with iodine. Column chromatography was performed on an Isco Teledyne Combiflash Nextgen 300+ instrument using pre-packed Redi-sep Gold silica gel cartridges (particle diameter 20-40 µM, pore diameter 60 Å). The eluents are given in brackets. Mass spectrometry was performed on an LTQ FT-ICR mass spectrometer equipped with an electrospray ionization source (Finnigan LTQ FT, Thermo Fisher Scientific, Waltham, MA) operated in either positive or negative ion mode.

<sup>1</sup>H NMR spectra NMR data were acquired at 298 K using a 500 MHz Bruker Avance Neo NMR spectrometer that was equipped with a 5 mm iProbe or a 400 MHz Bruker Avance I spectrometer equipped with a 5 mm BBO Smart Probe. The experiments were conducted using the default Bruker NMR parameters and data was time-averaged until a sufficient level of sensitivity was achieved. <sup>1</sup>H NMR data was calibrated by using the residual peak of the solvent as the internal standard (CDCl<sub>3</sub>: δ<sub>H</sub> = 7.26 ppm; CD<sub>3</sub>OD: δ<sub>H</sub> = 3.31 ppm). All coupling constants are recorded in

Hz. NMR spectra were processed with MestReNova v14.1.2-25024 software using the baseline and phasing correction features. Multiplicities and coupling constants were calculated using the multiplet analysis feature with manual intervention as necessary.



**Diethyl 2-(3-methylbenzyl)malonate (22)** Diethyl malonate (500.57 mg, 3.125 mmol, 1.05 equiv.) was added dropwise to a suspension of 60% NaH on mineral oil (125 mg, 3.125 mmol, 1.05 equiv.) in 6 mL dry THF at 0 °C. After 20 min, 3-methylbenzyl bromide (550.86 mg, 2.97 mmol, 1 equiv.) was added in one portion and the reaction mixture was refluxed overnight. The next day, the reaction was cooled and quenched by the addition of H<sub>2</sub>O. Et<sub>2</sub>O was added and the aqueous layer was extracted three times with Et<sub>2</sub>O. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub> then evaporated to dryness under reduced pressure. The crude product was purified by flash chromatography on SiO<sub>2</sub> [eluent: EtOAc/hexane (5% then 10% then 15% then 20%)] to

obtain pure diethyl 2-(3-methylbenzyl)malonate **22** as a clear liquid. Yield 20.6%.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.19 (t,  $J = 7.8$  Hz, 1H), 7.07 – 7.00 (m, 3H), 4.19 (qd,  $J = 7.1, 2.8$  Hz, 4H), 3.65 (t,  $J = 7.8$  Hz, 1H), 3.20 (d,  $J = 7.8$  Hz, 2H), 2.34 (s, 3H), 1.24 (t,  $J = 7.1$  Hz, 6H). HR-EI-MS [M+H] $^+$ : calculated for  $\text{C}_{15}\text{H}_{21}\text{O}_4^+$ , *m/z* 265.1434, found *m/z* 265.1395.

**2-(3-methylbenzyl)malonic acid (17)** Diethyl 2-(3-methylbenzyl)malonate **22** (100 mg, 0.362 mmol) was dissolved in 1 mL of ethanol then added dropwise to 375  $\mu\text{L}$  of 6.67 M NaOH. The mixture was stirred for 5 h at 60 °C. The solution was then cooled to 0 °C, carefully acidified to pH 1 with 1 N HCl, and extracted with 5 portions of  $\text{Et}_2\text{O}$ . The combined extracts were washed with a saturated aqueous solution of NaCl, dried over  $\text{Na}_2\text{SO}_4$  and evaporated to dryness. 2-(3-methylbenzyl)malonic acid was dissolved in 1:1  $\text{H}_2\text{O}:\text{MeCN}$  and lyophilized to give a white solid. Yield >99%.  $^1\text{H}$  NMR (500 MHz, MeOD)  $\delta$  7.04 (t,  $J = 7.6$  Hz, 1H), 6.95 (s, 1H), 6.94 – 6.88 (m, 2H), 3.49 (t,  $J = 7.8$  Hz, 1H), 3.01 (d,  $J = 7.7$  Hz, 2H), 2.20 (s, 3H). HR-ESI-MS [M-H] $^-$ : calculated for  $\text{C}_{11}\text{H}_{11}\text{O}_4$ , *m/z* 207.0663, found *m/z* 207.0661.

**Diethyl 2-(3-(trifluoromethyl)benzyl)malonate (23)** Diethyl malonate (500.57 mg, 3.125 mmol, 1.05 equiv.) was added dropwise to a suspension of 60% NaH on mineral oil (125 mg, 3.125 mmol, 1.05 equiv.) in 6 mL dry THF at 0°C. After 20 min, 3-(trifluoromethyl)benzyl bromide (711.377 mg, 2.97 mmol, 1 equiv.) was added in one portion and the reaction mixture was refluxed overnight. The next day, the reaction was cooled and quenched by the addition of  $\text{H}_2\text{O}$ .  $\text{Et}_2\text{O}$  was added and the aqueous layer was extracted three times with  $\text{Et}_2\text{O}$ . The combined organic layers were then dried over  $\text{Na}_2\text{SO}_4$  then evaporated to dryness under reduced pressure. The crude product was purified by flash chromatography on  $\text{SiO}_2$  [eluent: EtOAc/hexane (5%

then 10% then 15% then 20%)] to obtain pure diethyl 2-(3-(trifluoromethyl)benzyl)malonate as a clear liquid. Yield 22.1%.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.56 – 7.49 (m, 2H), 7.45 (p,  $J$  = 2.1 Hz, 2H), 4.21 (qd,  $J$  = 7.2, 1.7 Hz, 4H), 3.68 (t,  $J$  = 7.8 Hz, 1H), 3.31 (d,  $J$  = 7.9 Hz, 2H), 1.25 (t,  $J$  = 7.1 Hz, 6H). HR-EI-MS [M+H] $^+$ : calculated for  $\text{C}_{15}\text{H}_{18}\text{F}_3\text{O}_4^+$ ,  $m/z$  319.1152, found  $m/z$  319.1112.

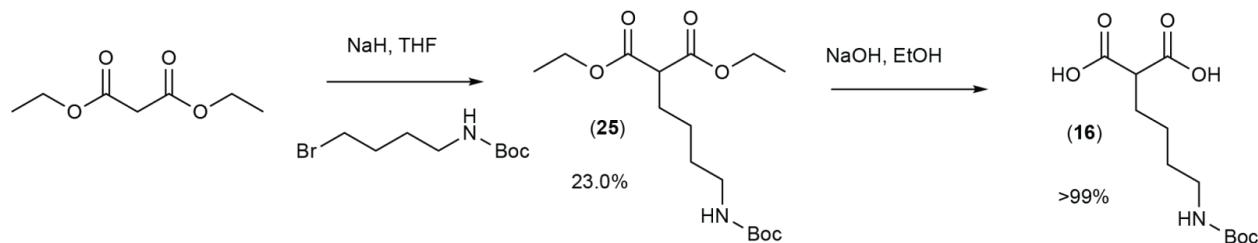
**2-(3-(trifluoromethyl)benzyl)malonic acid (18)** Diethyl 2-(3-(trifluoromethyl)benzyl)malonate (100 mg, 0.314 mmol) was dissolved in 1 mL of ethanol then added dropwise to 375  $\mu\text{L}$  of 6.67 M NaOH. The mixture was stirred for 5 h at 60 °C. The solution was then cooled to 0°C, carefully acidified to pH 1 with 1 N HCl, and extracted with 5 portions of  $\text{Et}_2\text{O}$ . The combined extracts were washed with a saturated solution of NaCl, dried over  $\text{Na}_2\text{SO}_4$ , and evaporated to dryness. 2-(3-(trifluoromethyl)benzyl)malonic acid was dissolved in 1:1  $\text{H}_2\text{O}:\text{MeCN}$  and lyophilized to give a white solid. Yield >99%.  $^1\text{H}$  NMR (500 MHz, MeOD)  $\delta$  7.60 – 7.46 (m, 4H), 3.69 (t,  $J$  = 7.8 Hz, 1H), 3.26 (d,  $J$  = 7.7 Hz, 2H). HR-ESI-MS [M-H] $^-$ : calculated for  $\text{C}_{11}\text{H}_{8}\text{F}_3\text{O}_4$ ,  $m/z$  261.0380, found  $m/z$  261.0377.

**Diethyl 2-(3-bromobenzyl)malonate (24)** Diethyl malonate (500.57 mg, 3.125 mmol, 1.05 equiv.) was added dropwise to a suspension of 60% NaH on mineral oil (125 mg, 3.125 mmol, 1.05 equiv.) in 6 mL dry THF at 0 °C. After 20 min, 3-bromobenzyl bromide (743.91 mg, 2.97 mmol, 1 equiv.) was added in one portion and the reaction mixture was refluxed overnight. The next day, the reaction was cooled and quenched by the addition of  $\text{H}_2\text{O}$ .  $\text{Et}_2\text{O}$  was added and the aqueous layer was extracted three times with  $\text{Et}_2\text{O}$ . The combined organic layers were then dried over  $\text{Na}_2\text{SO}_4$  then evaporated to dryness under reduced pressure. The crude product was purified

by flash chromatography on SiO<sub>2</sub> [eluent: EtOAc/hexane (5% then 10% then 15% then 20%)] to obtain pure diethyl 2-(3-bromobenzyl)malonate as a clear liquid. Yield 22.1%. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.39 – 7.32 (m, 2H), 7.19 – 7.10 (m, 2H), 4.17 (qd, J = 7.1, 1.3 Hz, 4H), 3.61 (t, J = 7.8 Hz, 1H), 3.18 (d, J = 7.9 Hz, 2H), 1.22 (t, J = 7.1 Hz, 6H). HR-EI-MS [M+H]<sup>+</sup>: calculated for C<sub>14</sub>H<sub>18</sub>BrO<sub>4</sub>, *m/z* 329.0383, found *m/z* 329.0343.

**2-(3-bromobenzyl)malonic acid (19)** Diethyl 2-(3-bromobenzyl)malonate **24** (100 mg, 0.305 mmol) was dissolved in 1 mL of ethanol then added dropwise to 375 μL of 6.67 M NaOH. The mixture was stirred for 5 h at 60 °C. The solution was then cooled to 0 °C, carefully acidified to pH 1 with 1 N HCl, and extracted with 5 portions of Et<sub>2</sub>O. The combined extracts were washed with a saturated solution of NaCl, dried over Na<sub>2</sub>SO<sub>4</sub>, and evaporated to dryness.

2-(3-bromobenzyl)malonic acid was dissolved in 1:1 H<sub>2</sub>O:MeCN and lyophilized to give a white solid. Yield >99%. <sup>1</sup>H NMR (500 MHz, MeOD) δ 7.33 (t, J = 1.9 Hz, 1H), 7.26 (dt, J = 7.7, 1.7 Hz, 1H), 7.16 – 7.06 (m, 2H), 3.52 (t, J = 7.8 Hz, 1H), 3.04 (d, J = 7.8 Hz, 2H). HR-ESI-MS [M-H]<sup>-</sup>: calculated for C<sub>10</sub>H<sub>8</sub>BrO<sub>4</sub>, *m/z* 270.9611, found *m/z* 270.9609.



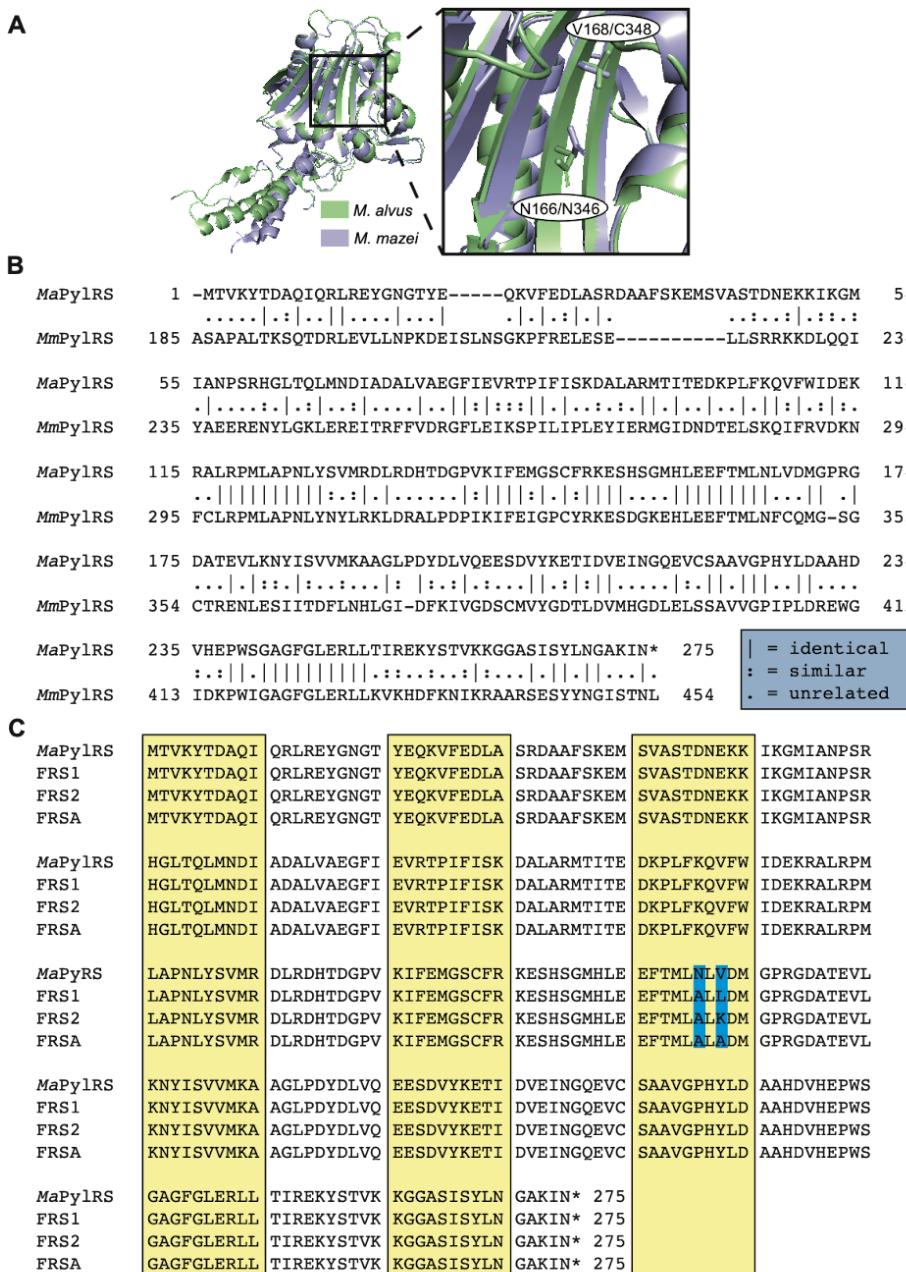
**Diethyl 2-((tert-butoxycarbonyl)amino)butylmalonate (25)** Diethyl malonate (166.86 mg, 1.04 mmol, 1.05 equiv.) was added dropwise to a suspension of 60% NaH on mineral oil (41.67 mg, 1.04 mmol, 1.05 equiv.) in 6 mL dry THF at 0 °C. After 20 min, 4-(Boc-amino)butyl

bromide (250.17 mg, 0.99 mmol, 1 equiv.) was added in one portion and the reaction mixture was refluxed overnight. The next day, the reaction was cooled and quenched by the addition of H<sub>2</sub>O. Et<sub>2</sub>O was added and the aqueous layer was extracted three times with Et<sub>2</sub>O. The combined organic layers were then dried over Na<sub>2</sub>SO<sub>4</sub> then evaporated to dryness under reduced pressure. The crude product was purified by flash chromatography on SiO<sub>2</sub> [eluent: EtOAc/hexane (5% then 10% then 15% then 20%)] to obtain pure diethyl 2-(4-((tert-butoxycarbonyl)amino)butyl)malonate as a clear liquid. Yield 22.1%. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 4.44 (s, 1H), 4.13 (qd, J = 7.1, 2.0 Hz, 4H), 3.24 (t, J = 7.5 Hz, 1H), 3.04 (q, J = 6.7 Hz, 2H), 1.87 – 1.79 (m, 2H), 1.44 (p, J = 7.3 Hz, 2H), 1.37 (s, 9H), 1.28 (tt, J = 10.5, 6.3 Hz, 2H), 1.20 (t, J = 7.1 Hz, 6H).

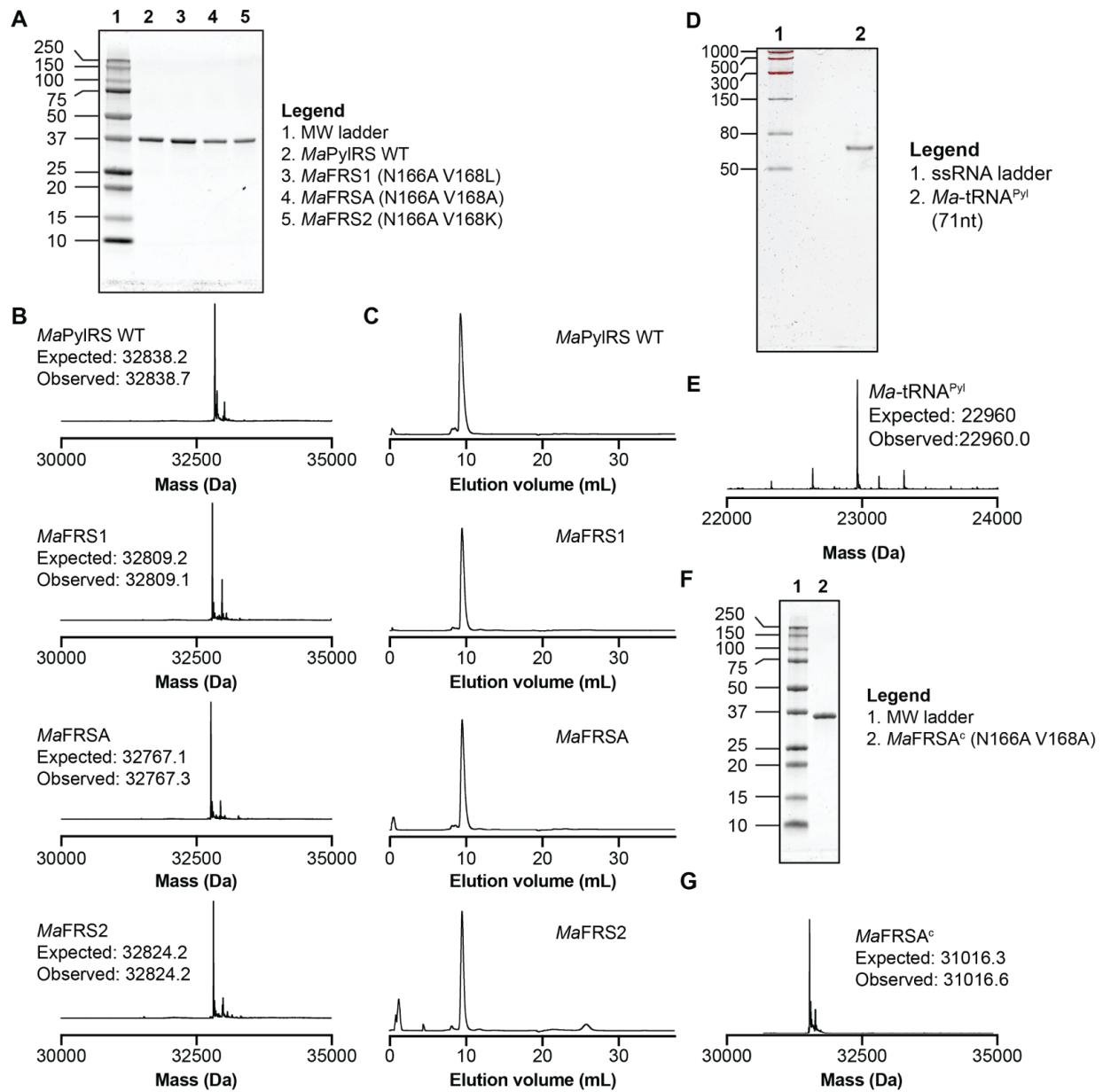
### **2-(4-((tert-butoxycarbonyl)amino)butyl)malonic acid (16) Diethyl**

2-(4-((tert-butoxycarbonyl)amino)butyl)malonate (72 mg, 0.22 mmol) was dissolved in 1 mL of ethanol then added dropwise to 375 μL of 6.67 M NaOH. The mixture was stirred for 5 h at 60 °C. The solution was then cooled to 0 °C, carefully acidified to pH 1 with 1 N HCl, and extracted with 5 portions of Et<sub>2</sub>O. The combined extracts were washed with a saturated solution of NaCl, dried over Na<sub>2</sub>SO<sub>4</sub>, and evaporated to dryness. 2-(4-((tert-butoxycarbonyl)amino)butyl)malonic acid was dissolved in 1:1 H<sub>2</sub>O:MeCN and lyophilized to give a clear serum. Yield >99%. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 6.29 (s, 1H), 4.57 (s, 1H), 3.35 (t, J = 7.0 Hz, 1H), 3.03 (t, J = 6.7 Hz, 2H), 1.92 – 1.85 (m, 2H), 1.43 (p, J = 6.7 Hz, 2H), 1.36 (s, 11H). HR-ESI-MS [M-H]<sup>-</sup>: calculated for C<sub>12</sub>H<sub>20</sub>O<sub>6</sub>N<sub>1</sub> *m/z* 274.1296, found *m/z* 274.1295.

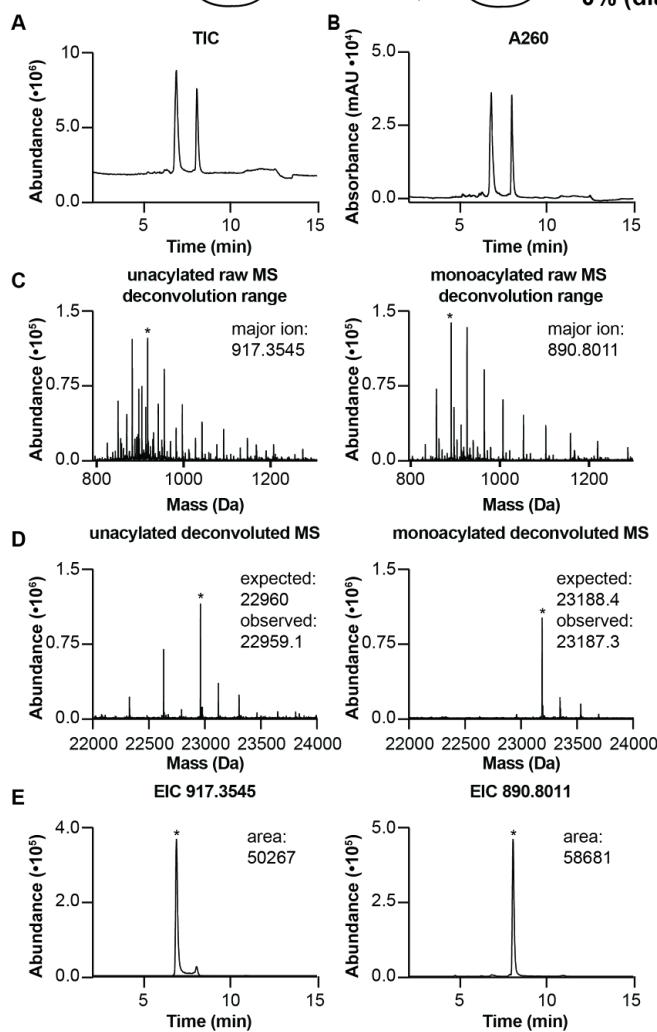
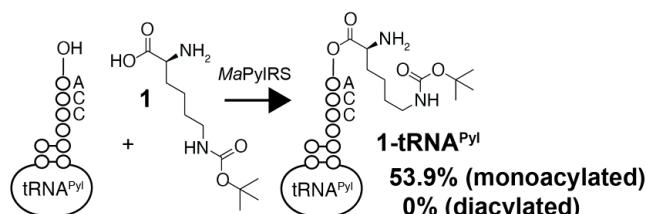
## Supplementary Figures



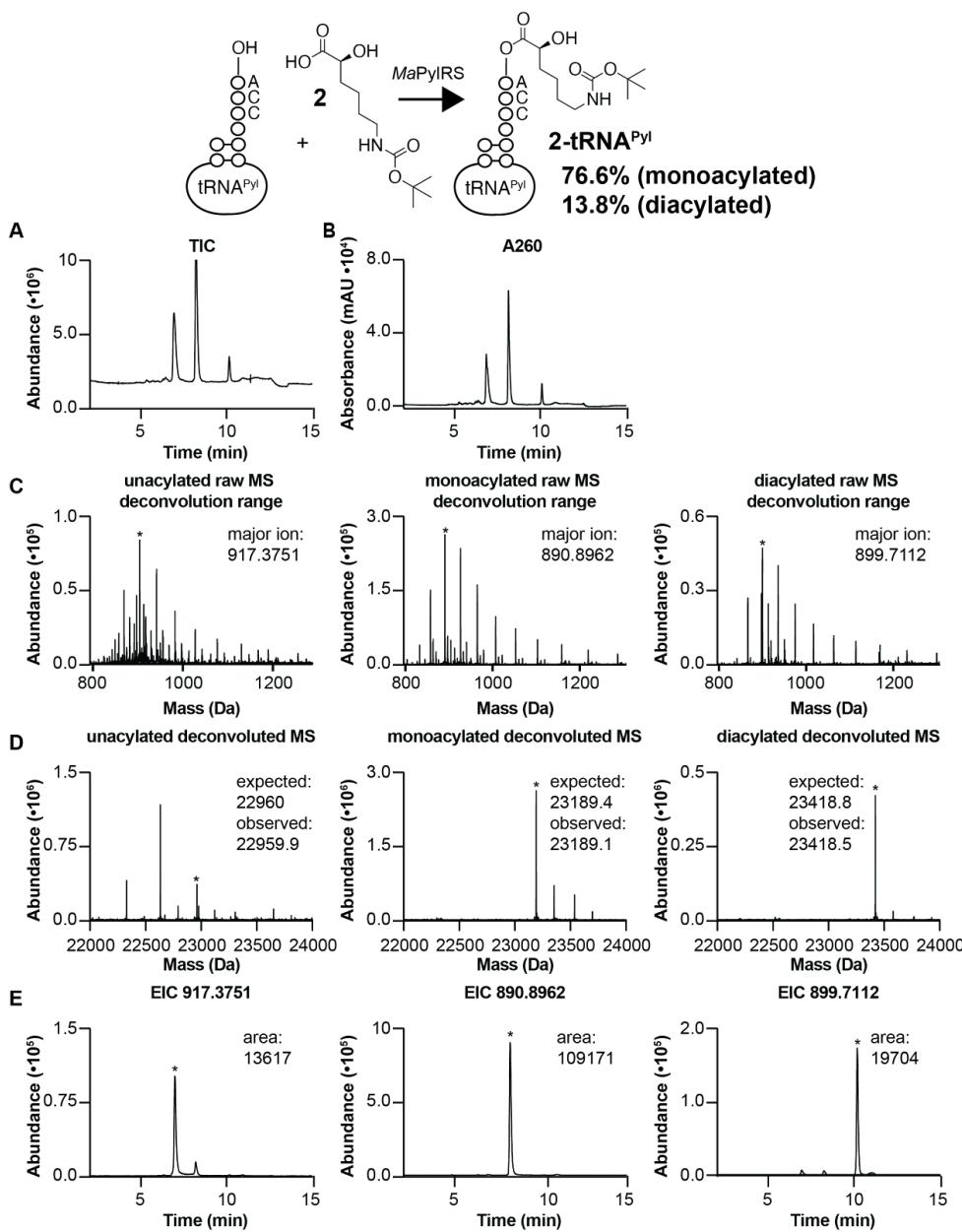
**Supplementary Figure 1.** **a**, Structural alignment of the *M. mazei* PylRS (*MmPylRS*) catalytic domain (PDB 2ZCE) and *M. alvus* PylRS (*MaPylRS*) (PDB 6JP2). The two active site residues substituted in FRS1, FRS2, and FRSA are shown explicitly. **b**, Sequence alignment of *MmPylRS* and *MaPylRS* using the EMBOSS Needle software<sup>3</sup>. **c**, Sequences of the four enzymes used in this study with differences highlighted in blue.



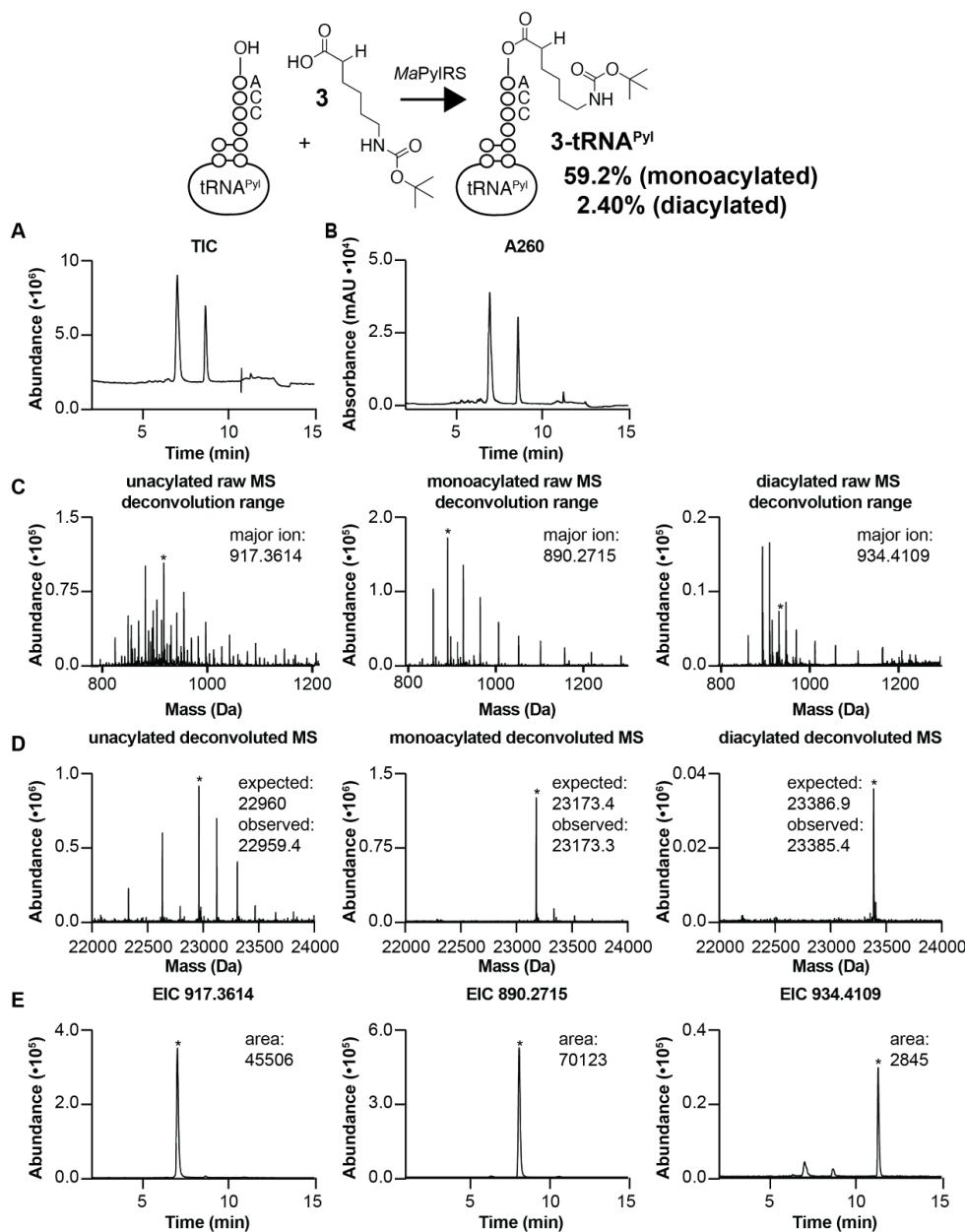
**Supplementary Figure 2.** **a**, SDS-PAGE; **b**, LC-MS; and **c**, analytical FPLC chromatograms of purified *MaPylRS*, *MaFRS1*, *MaFRS2*, and *MaFRSA* used in biochemical experiments. **d**, Urea-PAGE; and **e**, LC-MS analysis of *Ma-tRNA<sup>Pyl</sup>*. **f**, SDS-PAGE and **g**, LC-MS analysis of *MaFRSA* used for crystallography. We note that the *MaFRSA* in panel **b** is extended by an N-terminal His-tag and linker (GSSHHHHHHSSGLVPRGSH-), whereas the *MaFRSA* used for crystallography in panel **g** only contains an N-terminal GSH scar. The molecular weight standards used in lane 1 of panels **a** and **f** units are shown in units of kDa; the units of the standards used in panel **b** are nucleotides (nt).



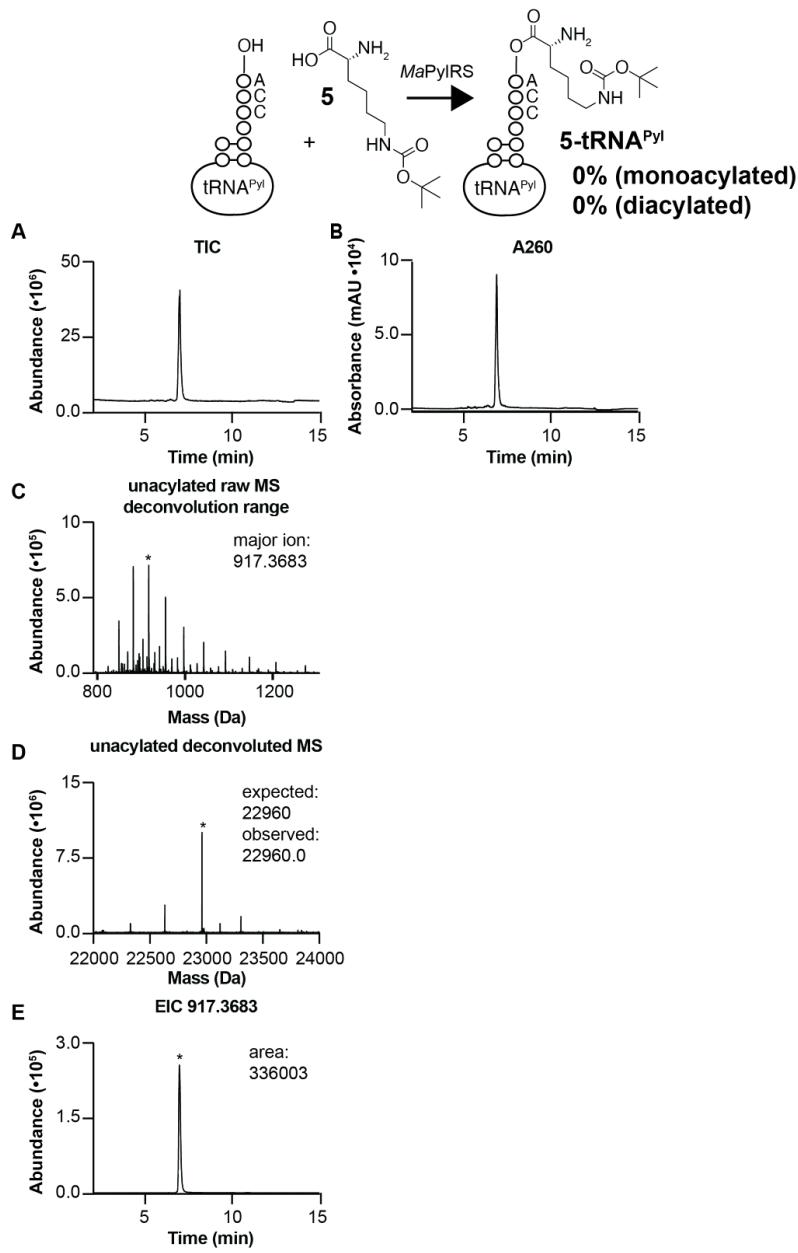
**Supplementary Figure 3.** Analysis of tRNA acylation product mixtures obtained using *MaPylRS*, *Ma-tRNA<sup>Pyl</sup>* and monomer **1** as described. **a**, Total ion count and **b**, UV absorbance (260 nm) as a function of elution time. **c**, The raw MS deconvolution range represents the subset of the raw MS data used to determine the deconvoluted mass spectrum of each tRNA species (unacylated or monoacylated). The major ion identified with an asterisk is the most abundant charge state of the tRNA species used for quantification. **d**, Deconvoluted mass spectra generated from the data in **(c)**. **e**, Extracted ion chromatograms of the major ions of each tRNA species. The peak corresponding to each tRNA species is noted with an asterisk. The EICs were integrated and the area under the curve ( $A$ ) was used to determine the overall tRNA acylation yield according to the equation:  $\text{yield} = [(A_{\text{mono-acylated}} + A_{\text{di-acylated}})/(A_{\text{unacylated}} + A_{\text{mono-acylated}} + A_{\text{di-acylated}})]$ . The yield shown in this figure is from a representative sample. Average yields from three technical replicates are displayed in Supplementary Table 3.



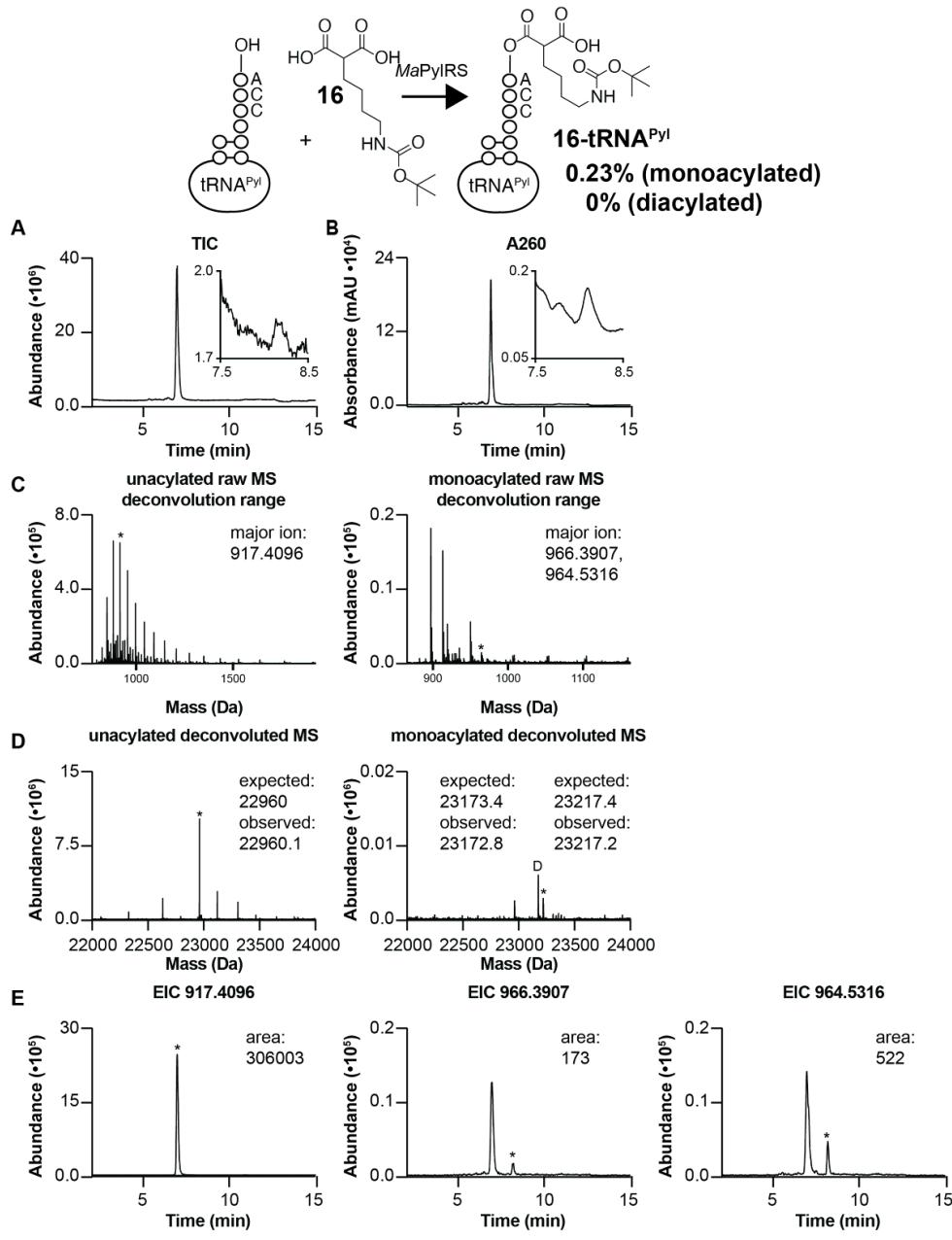
**Supplementary Figure 4.** Analysis of tRNA product mixtures obtained using *MaPylRS*, *Ma-tRNA<sup>Pyl</sup>* and monomer **2** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



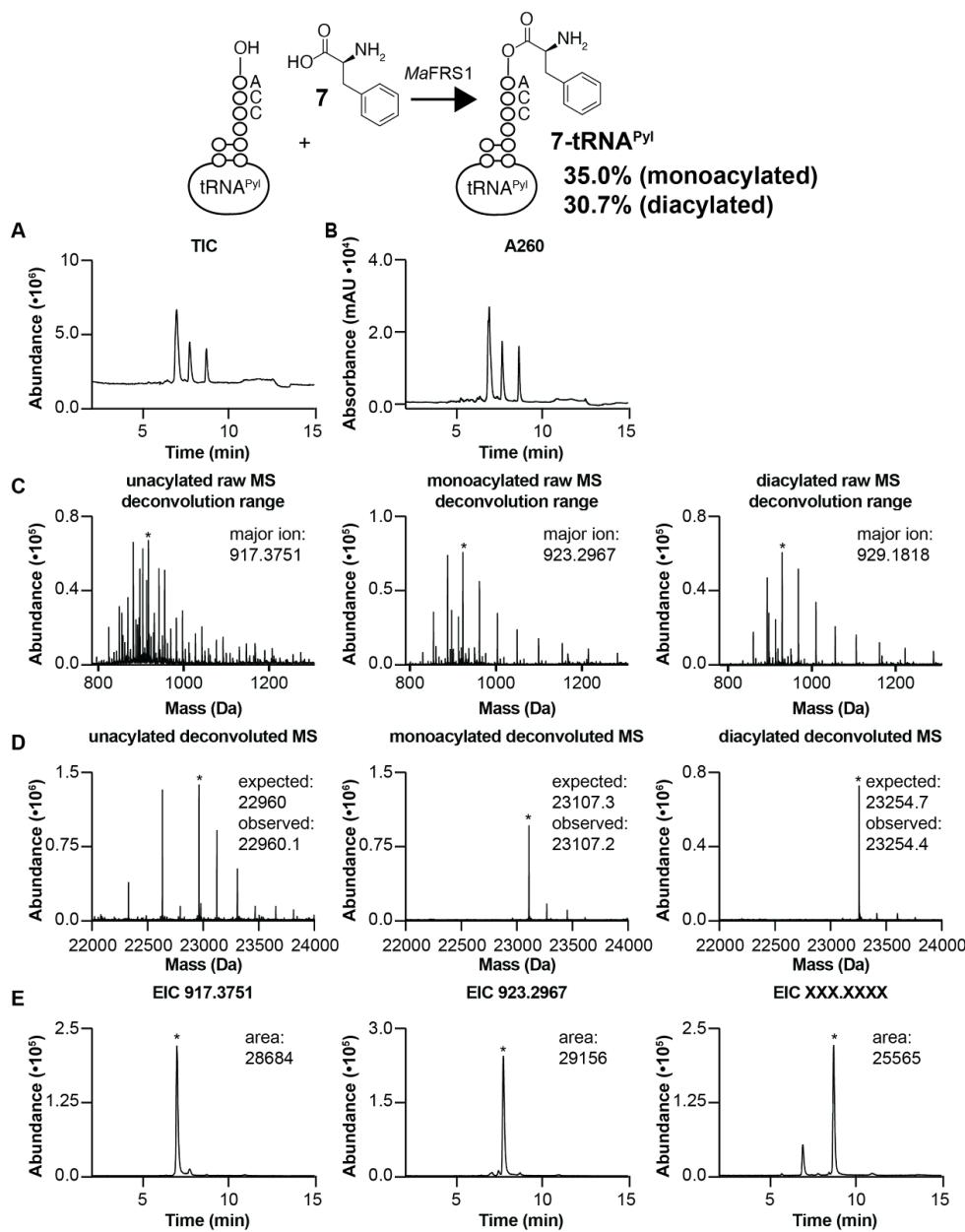
**Supplementary Figure 5.** Analysis of tRNA product mixtures obtained using *MaPylRS*, *Ma-tRNA<sup>Pyl</sup>* and monomer **3** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



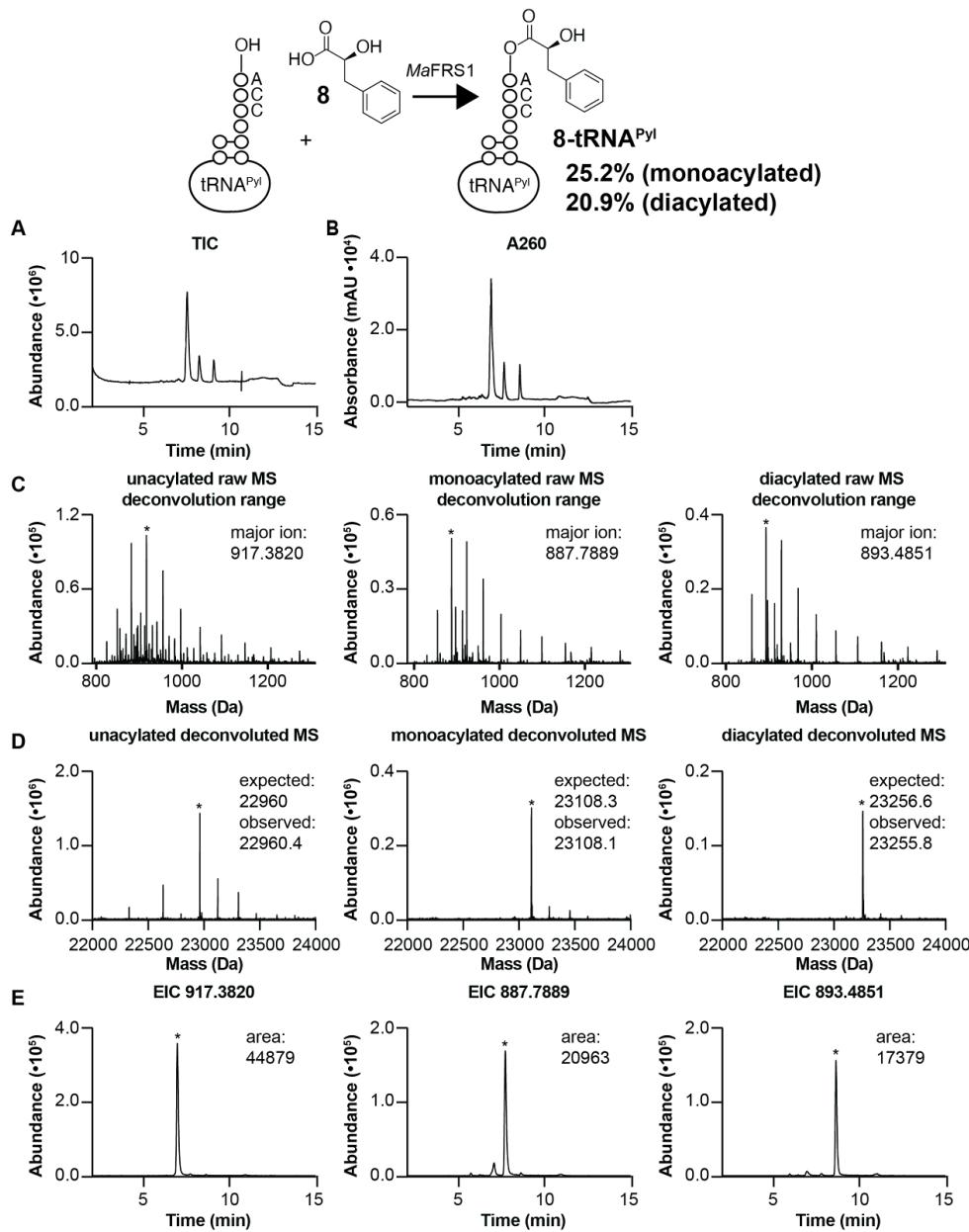
**Supplementary Figure 6.** Analysis of tRNA product mixtures obtained using *MaPylRS*, *Ma-tRNA*<sup>Pyl</sup> and monomer **5** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e.



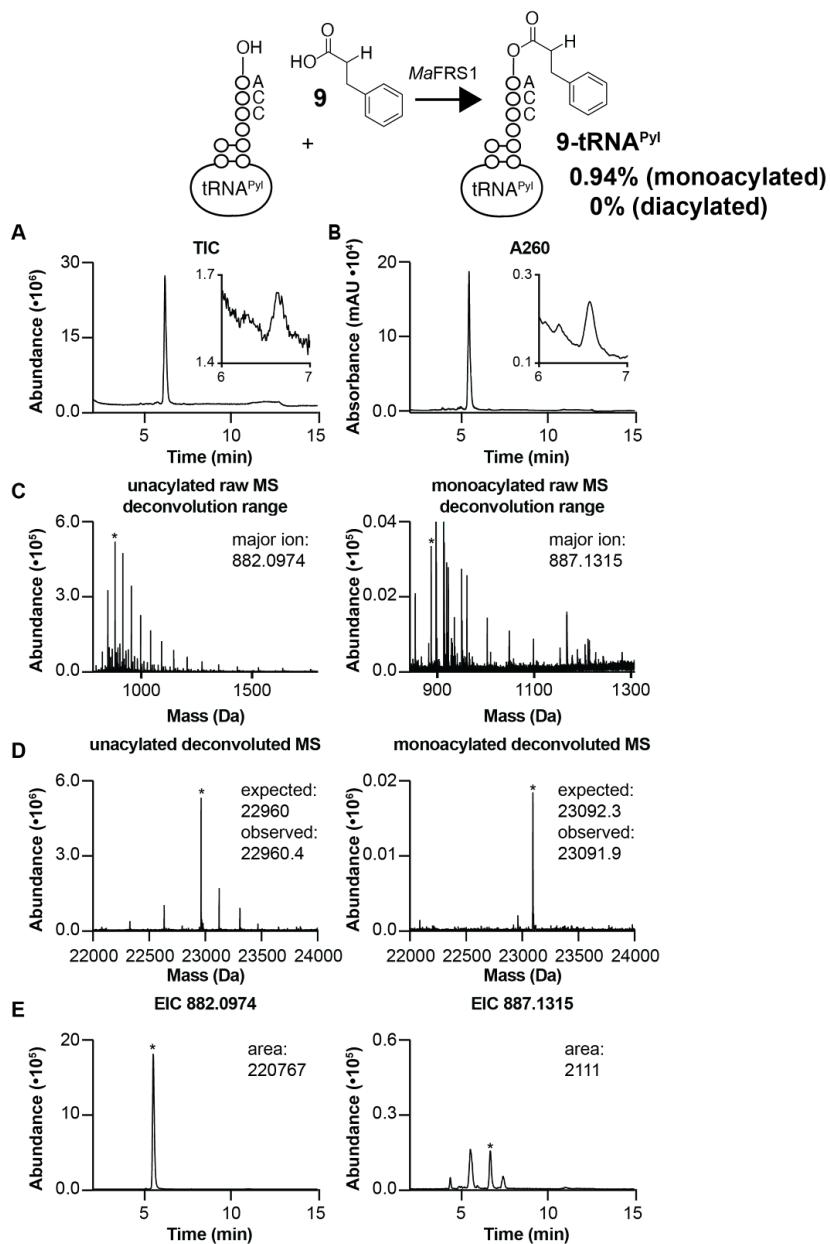
**Supplementary Figure 7.** Analysis of tRNA product mixtures obtained using *MaPylRS*, *Ma-tRNA<sup>Pyl</sup>* and monomer **16** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**. In panel **c**, the major ion for both the base mass and the decarboxylation product are listed. In panel **d**, the decarboxylation product mass is denoted by a D. The areas under the curve in panel **e** for the base and decarboxylation product masses were combined to calculate the overall acylation yield.



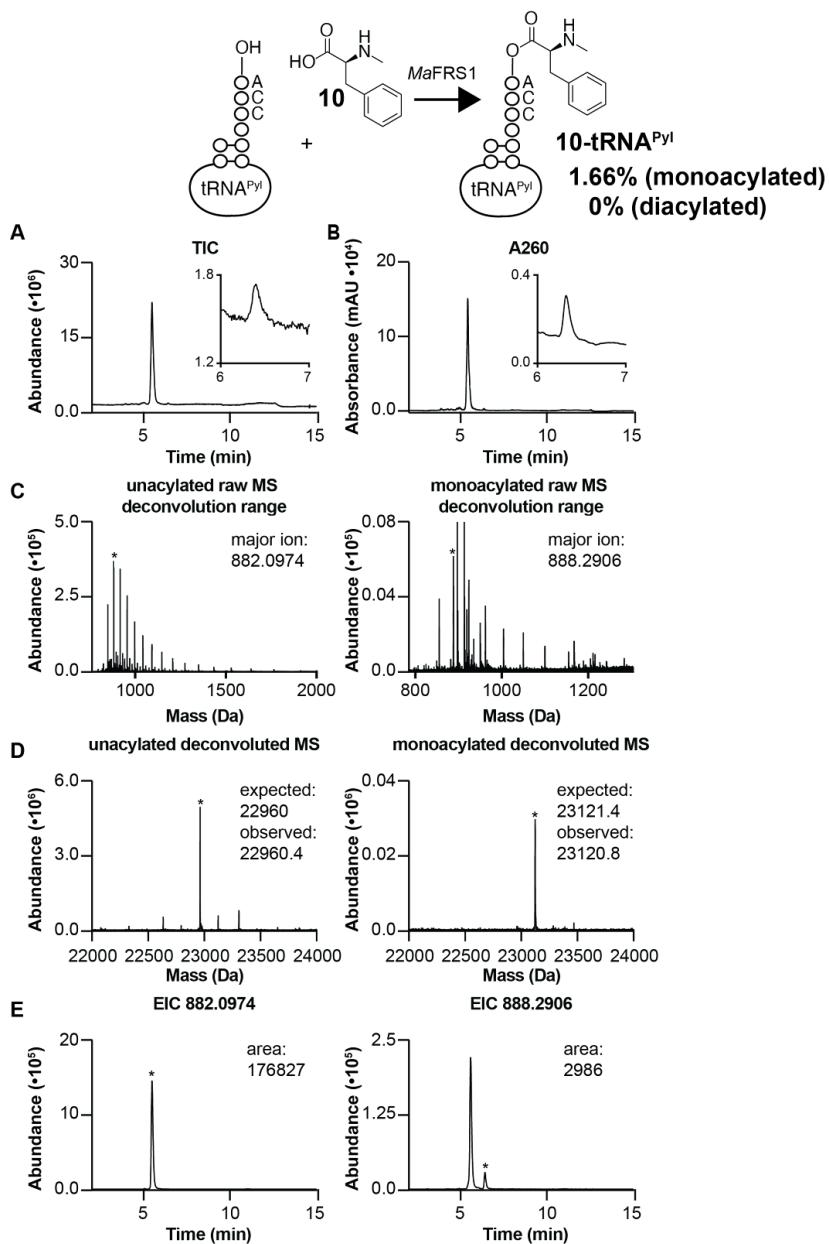
**Supplementary Figure 8.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **7** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



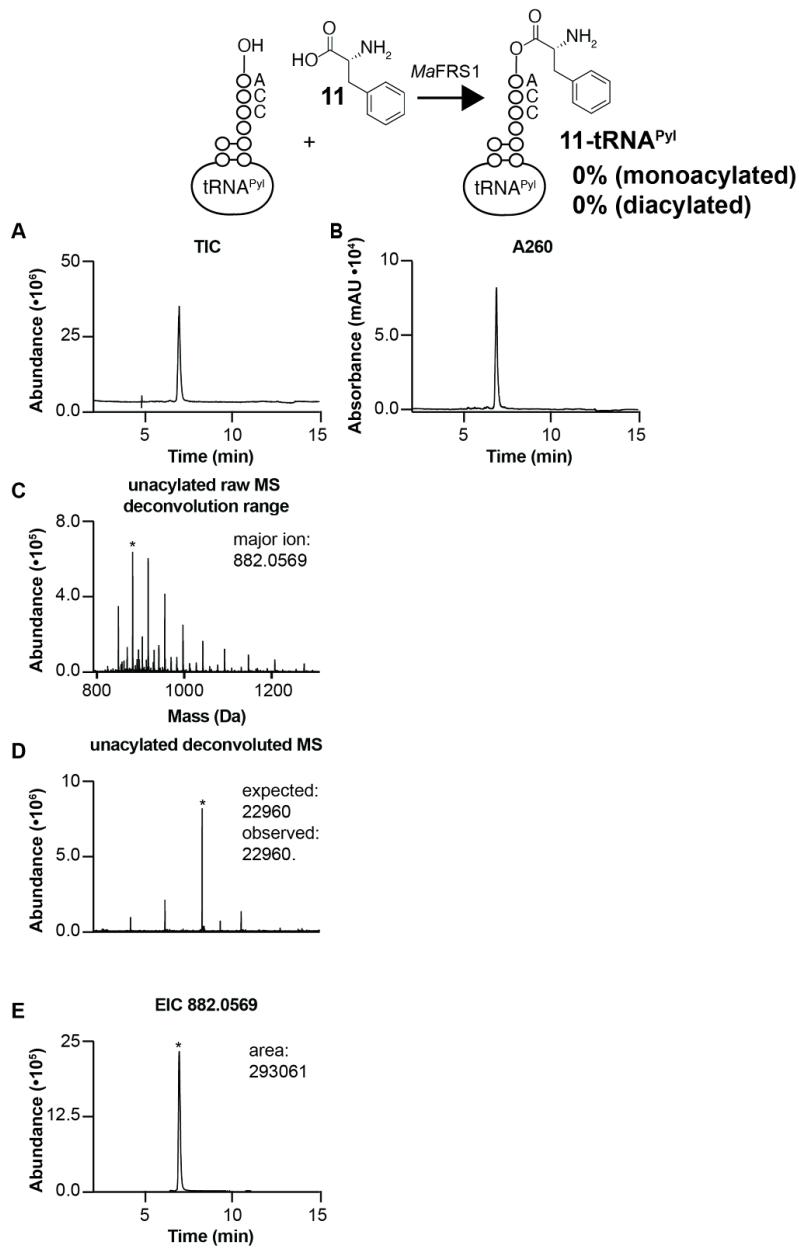
**Supplementary Figure 9.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **8** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e.



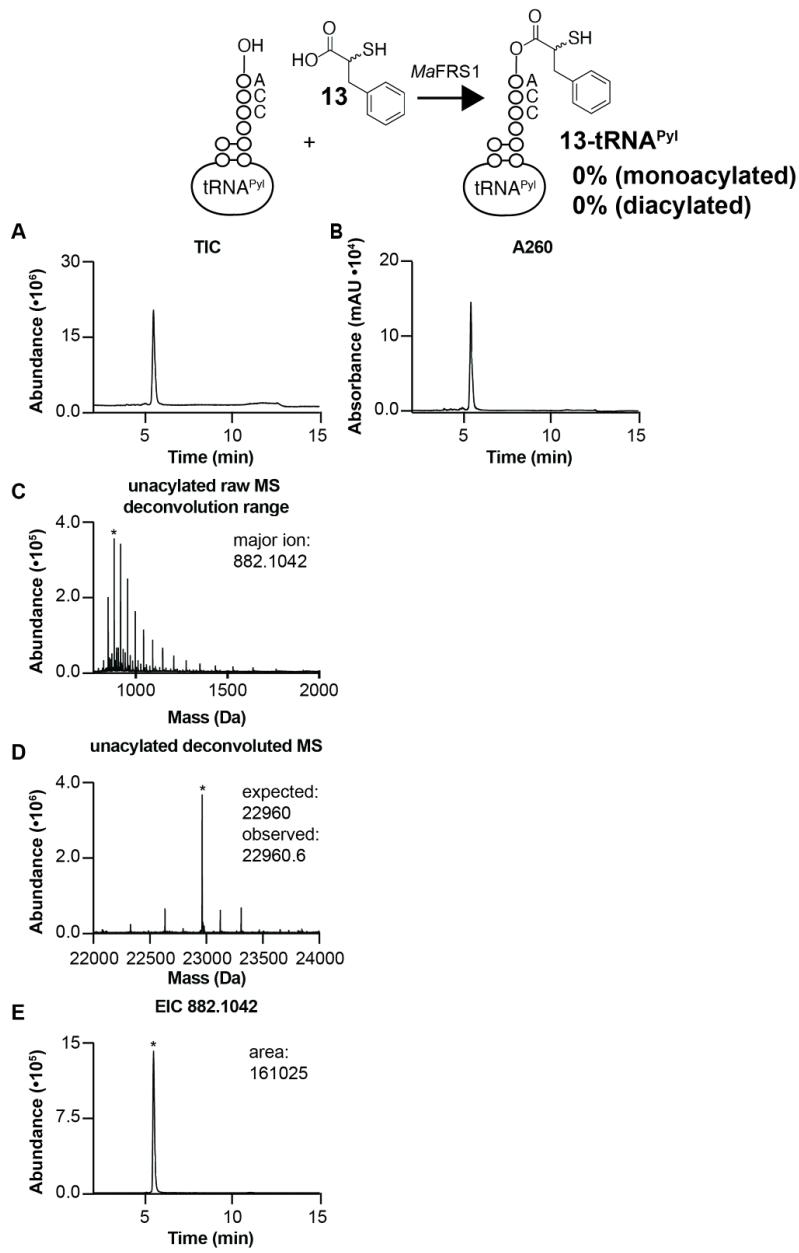
**Supplementary Figure 10.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **9** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e.



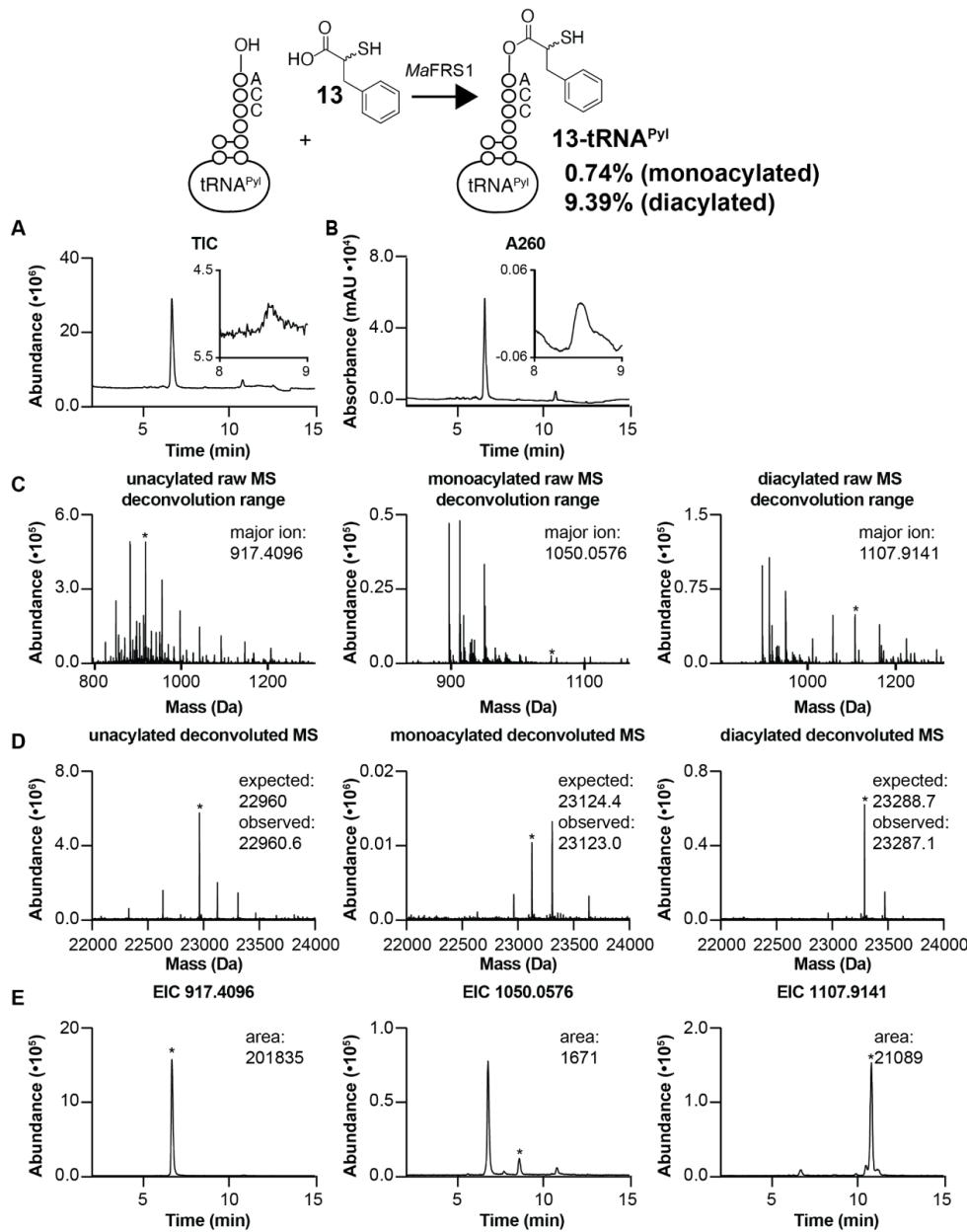
**Supplementary Figure 11.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA*<sup>Pyl</sup> and monomer **10** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e.



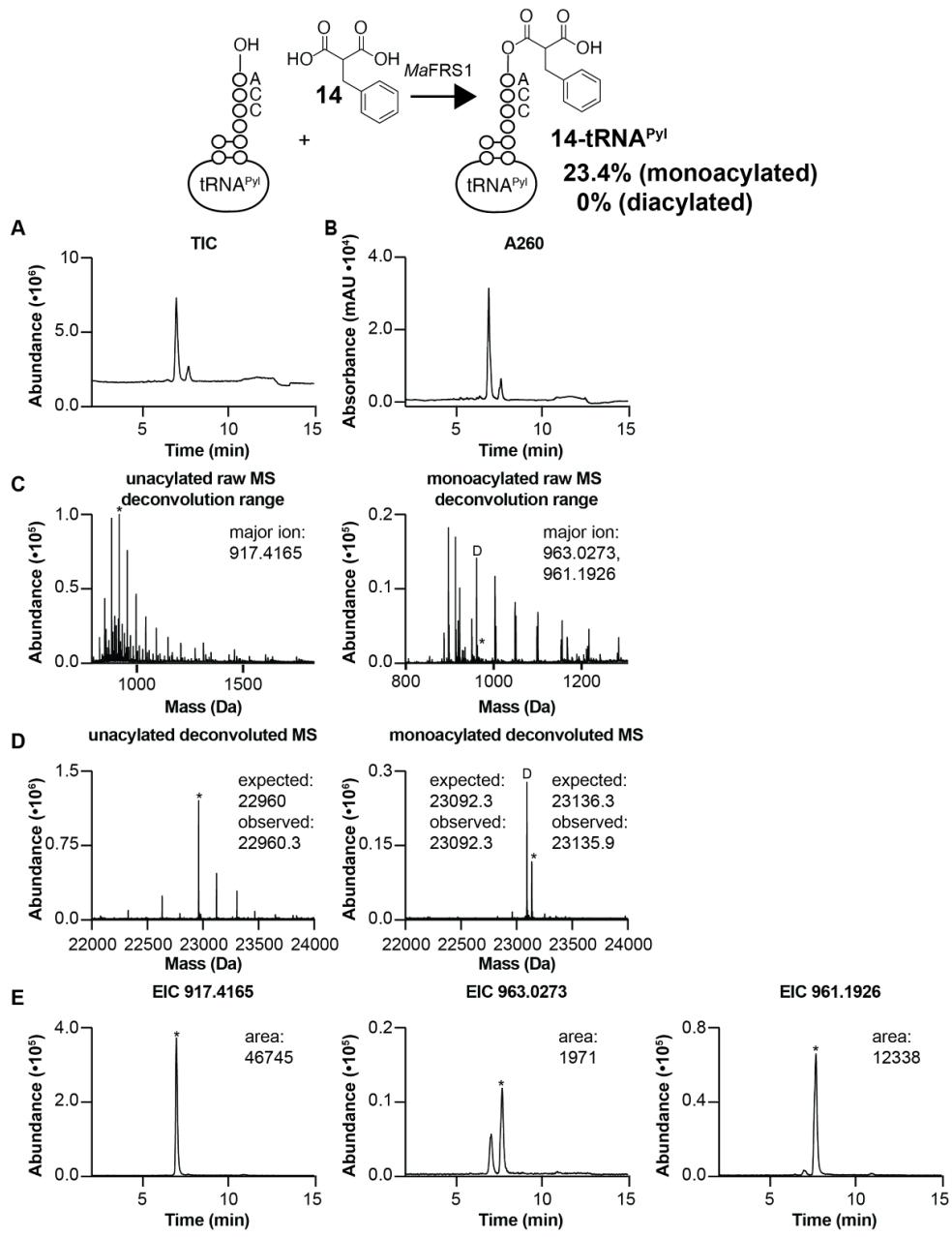
**Supplementary Figure 12.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **11** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



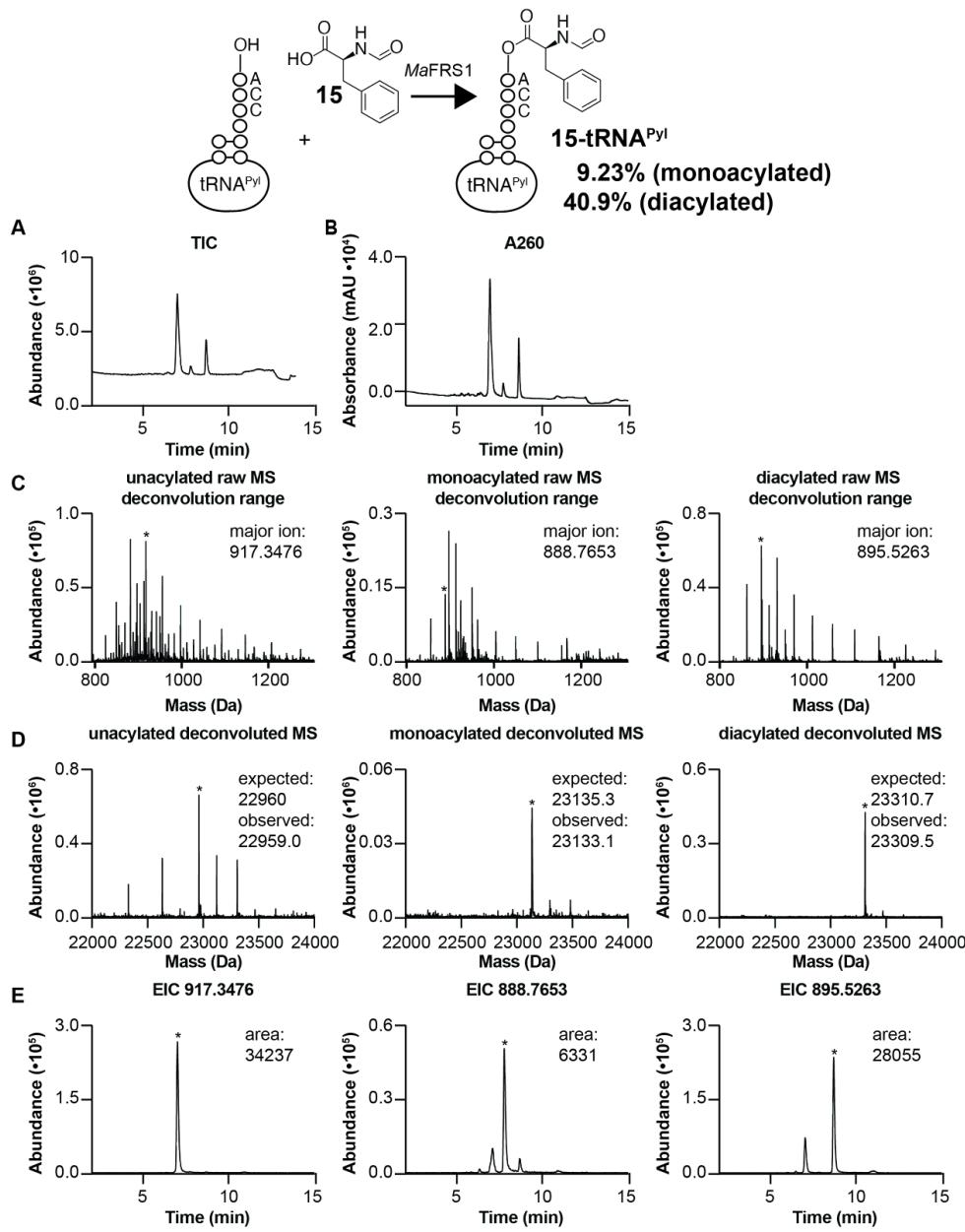
**Supplementary Figure 13.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **13** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



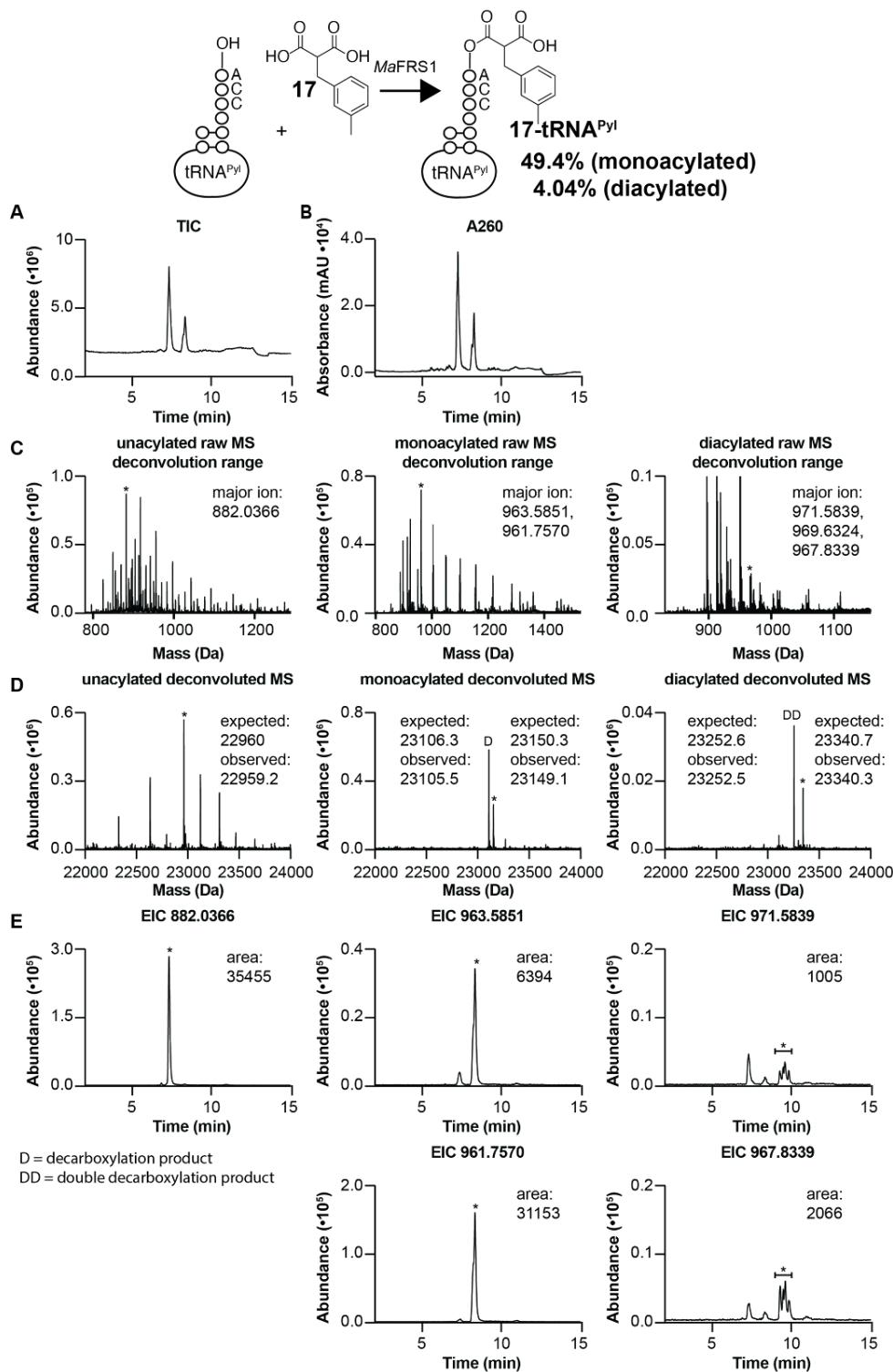
**Supplementary Figure 14.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **13** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e. In this case, acylation was performed using *MaFRS1*:*Ma-tRNA<sup>Pyl</sup>* ratio of 1:2.



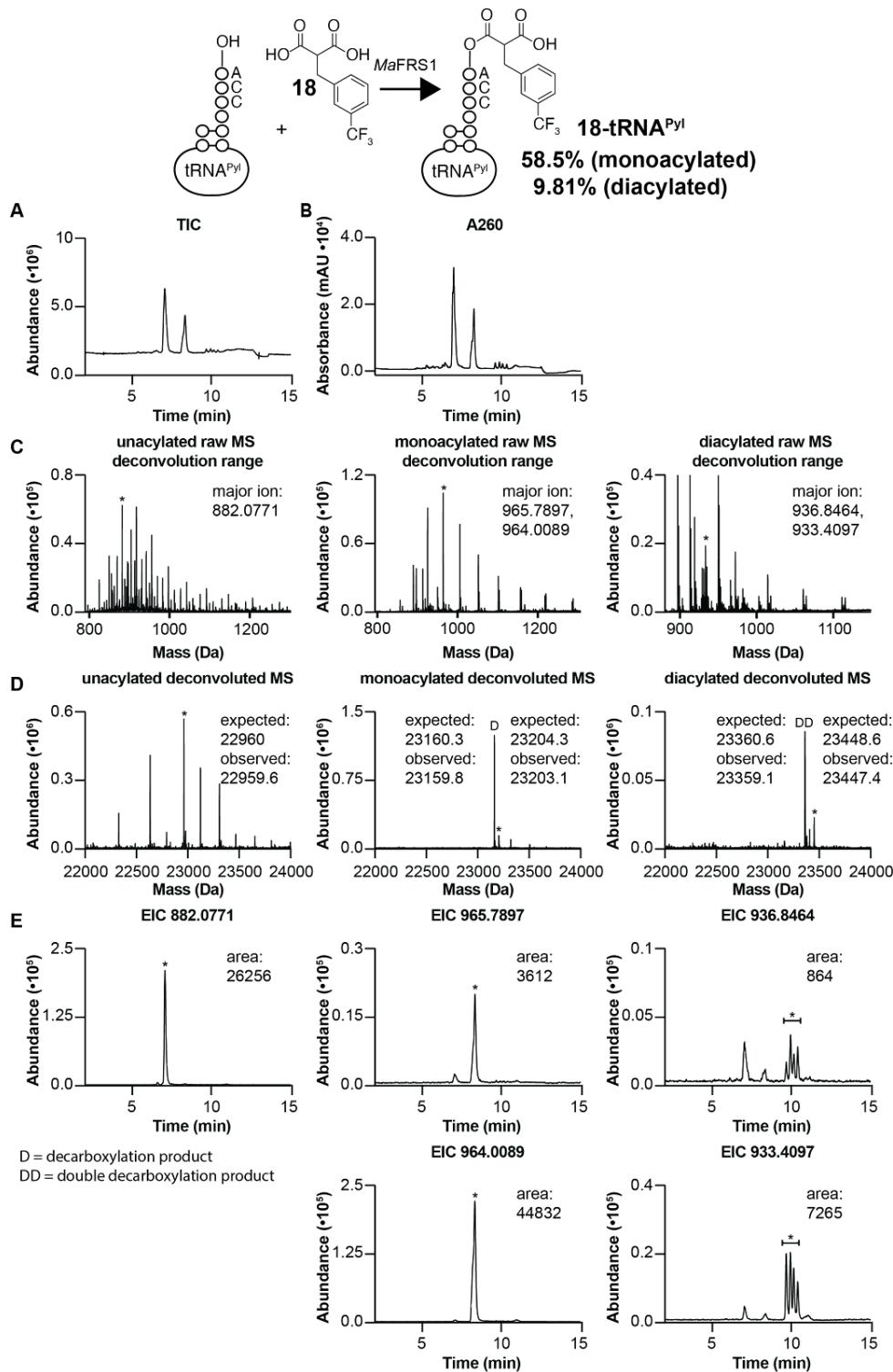
**Supplementary Figure 15.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **14** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e** and the legend of Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



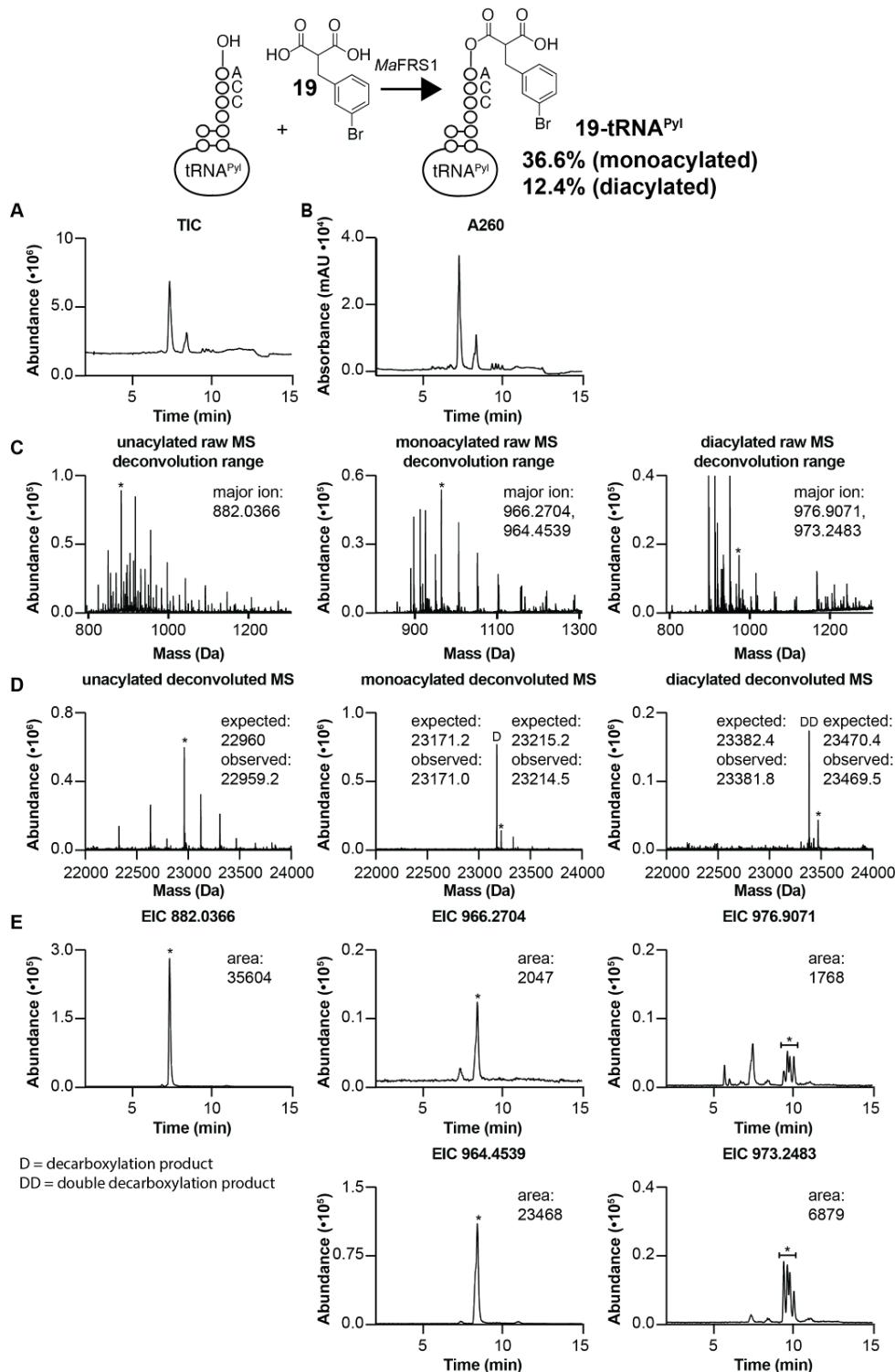
**Supplementary Figure 16.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **15** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



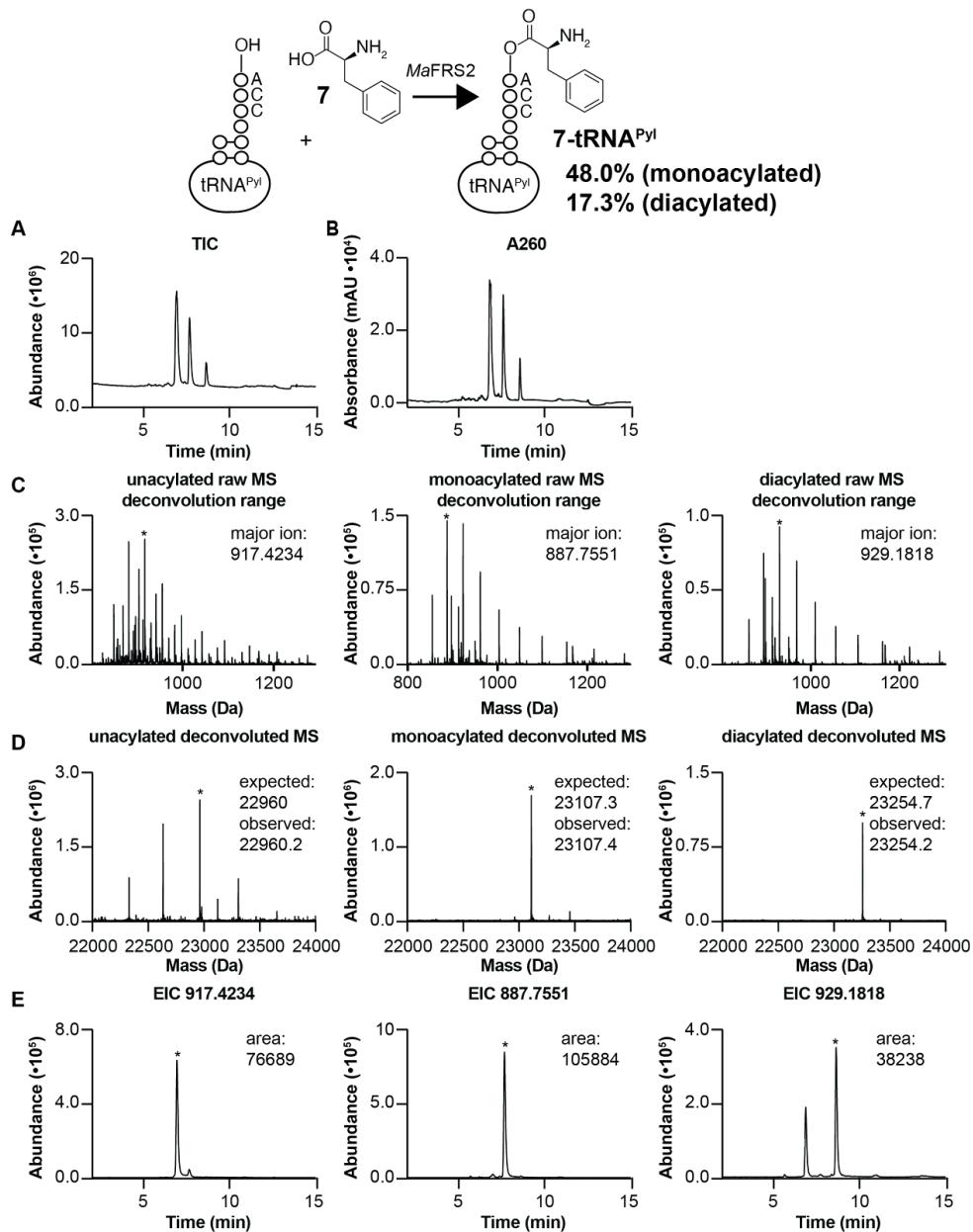
**Supplementary Figure 17.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **17** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e** and the legend of Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



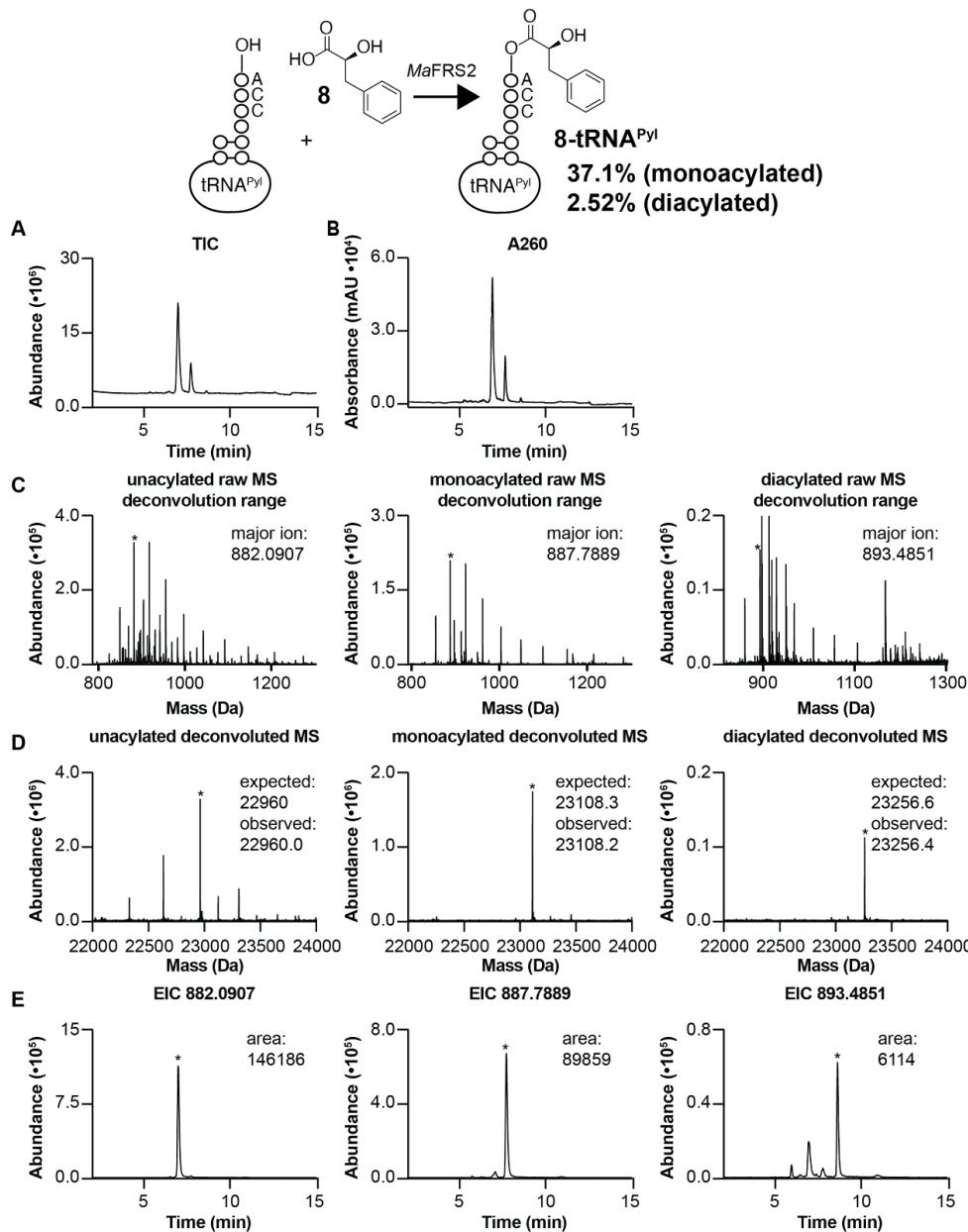
**Supplementary Figure 18.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **18** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e** and the legend of Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



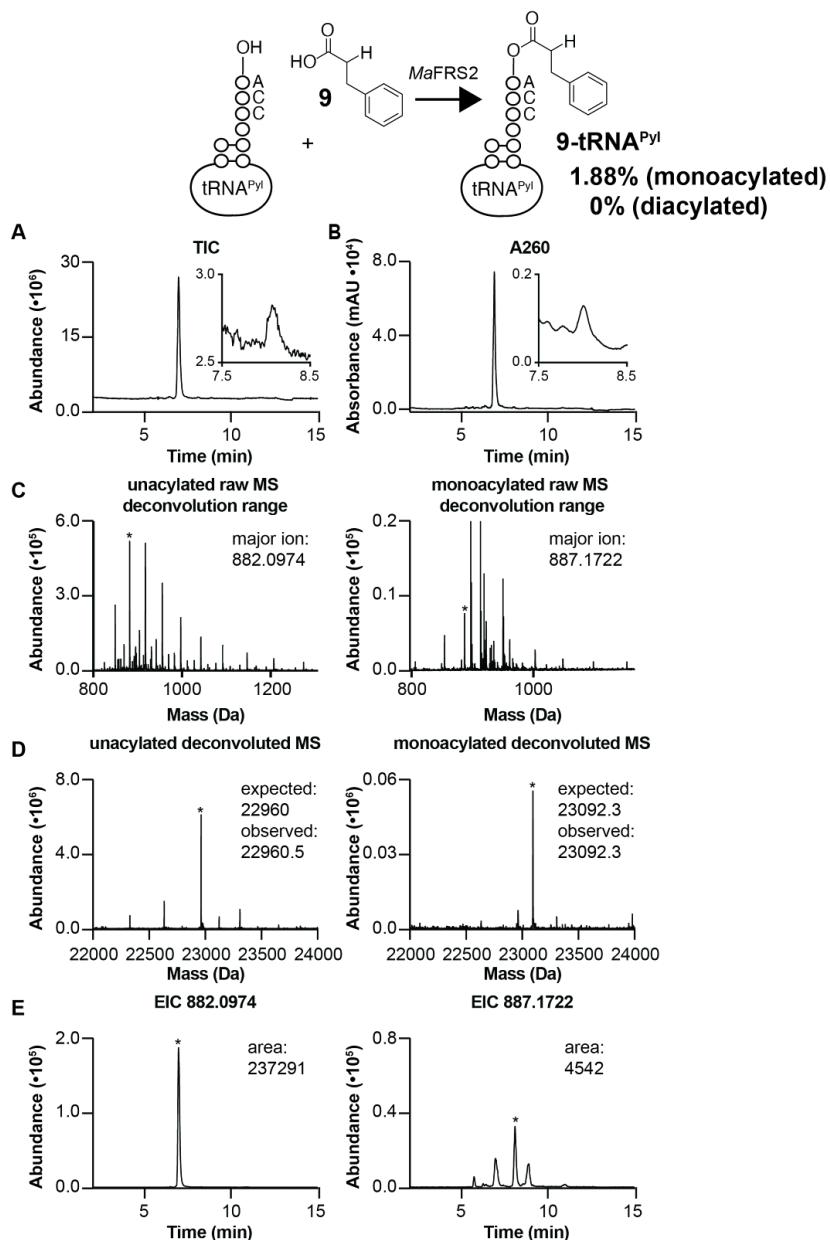
**Supplementary Figure 19.** Analysis of tRNA product mixtures obtained using *MaFRS1*, *Ma-tRNA<sup>Pyl</sup>* and monomer **19** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e** and a note on the decarboxylation products observed in the mass spectra.



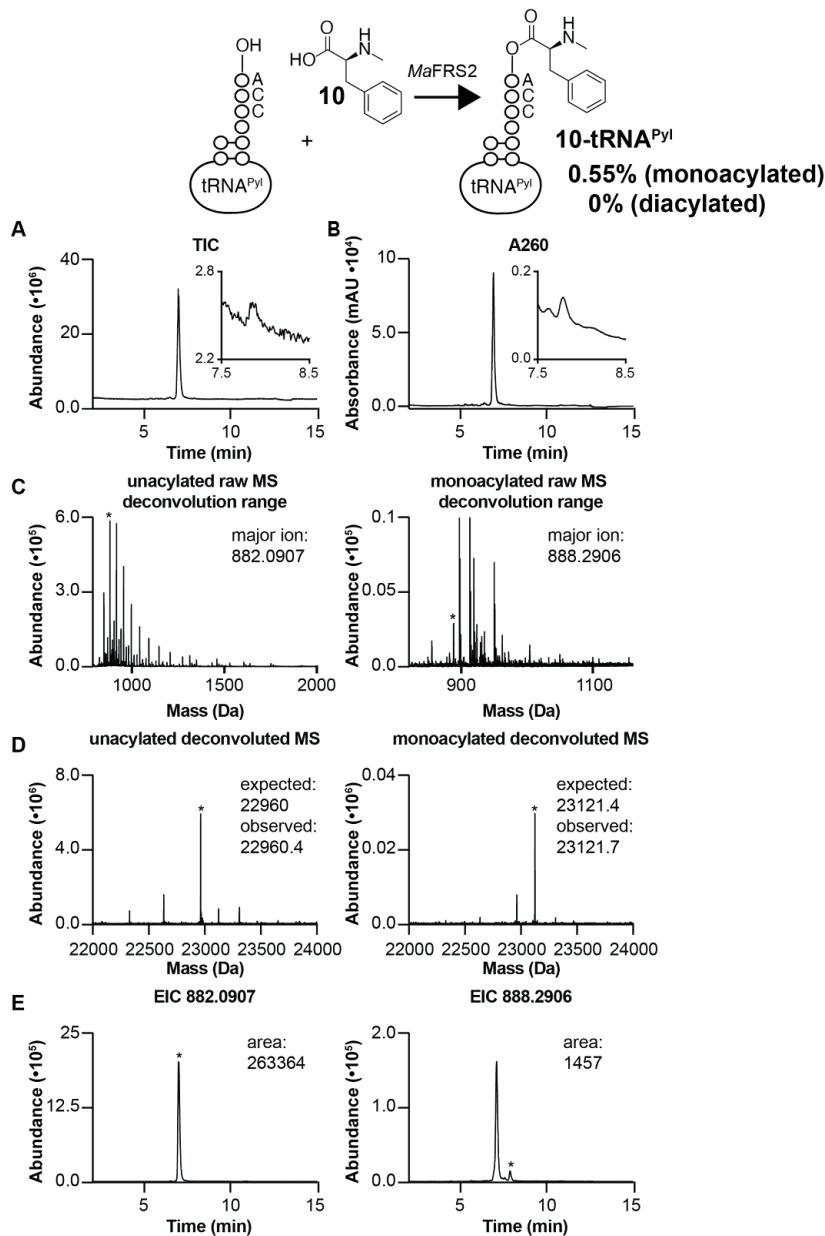
**Supplementary Figure 20.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **7** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e.



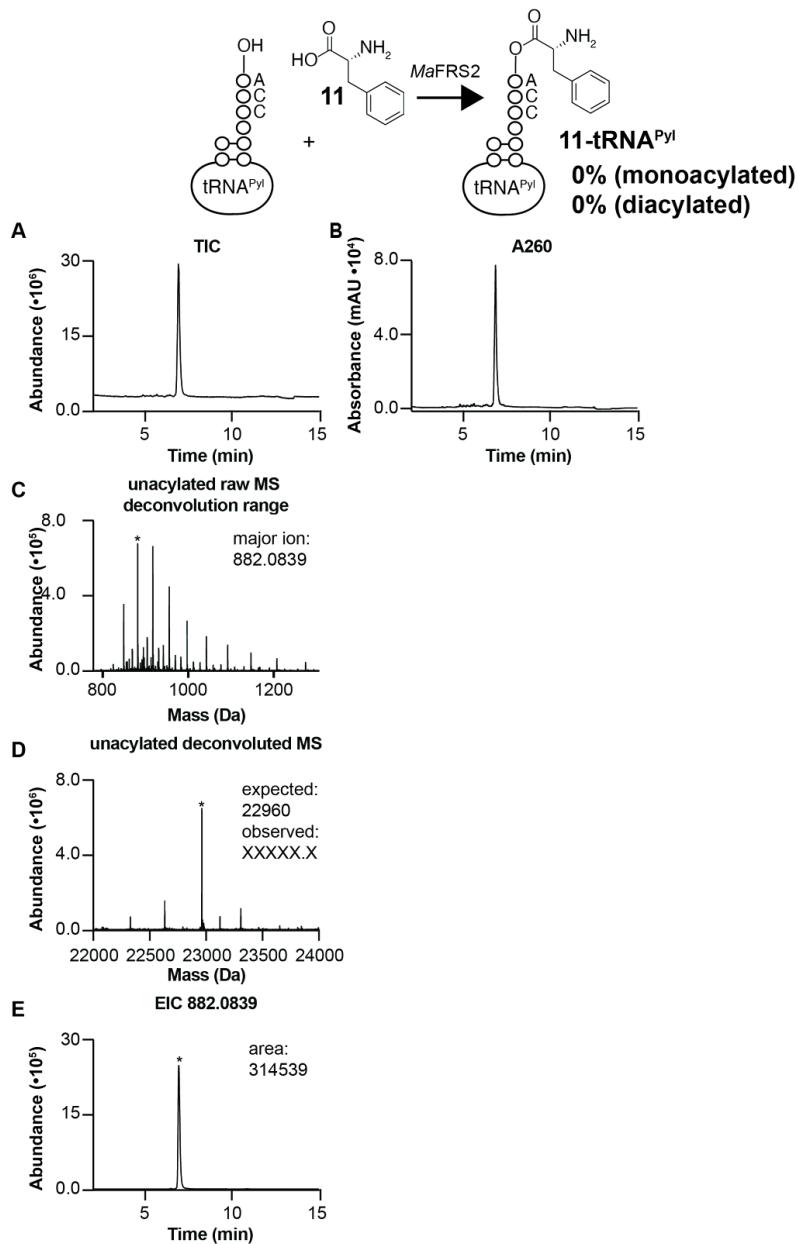
**Supplementary Figure 21.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **8** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e.



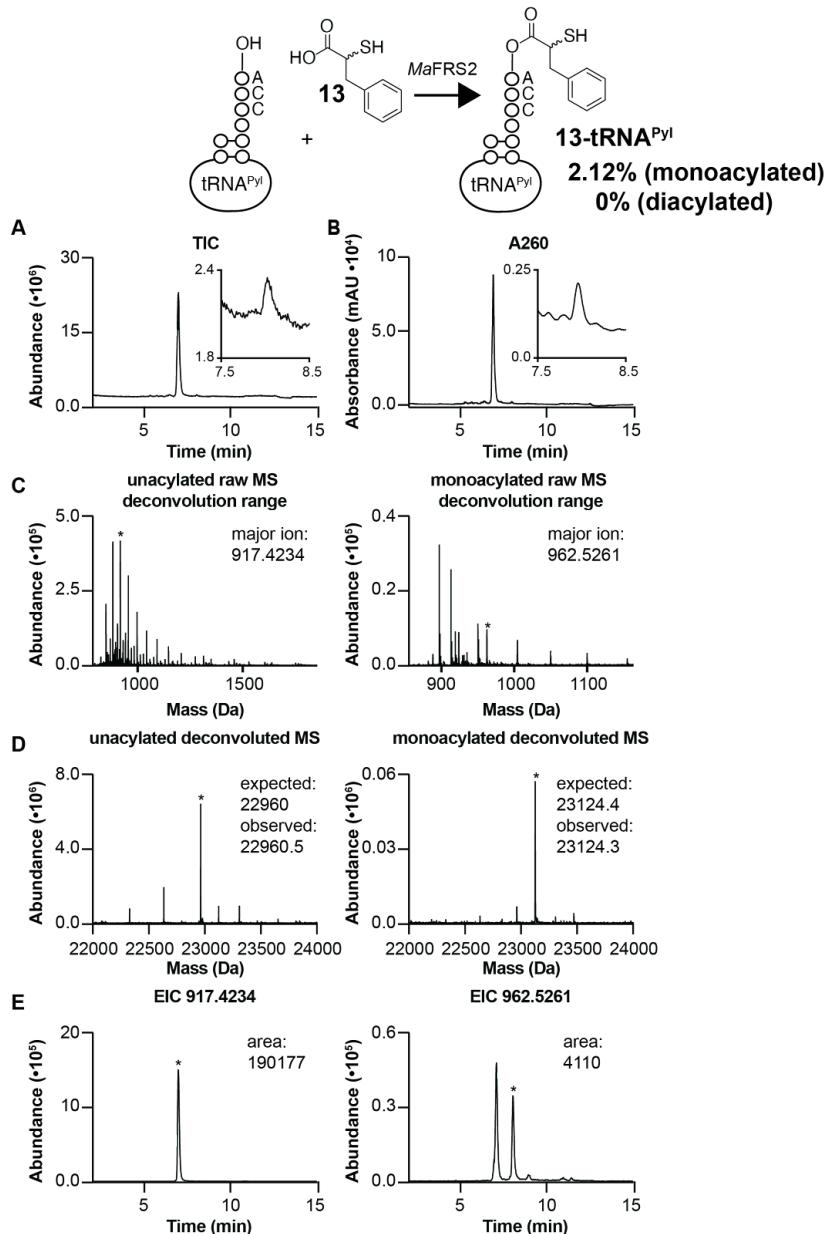
**Supplementary Figure 22.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **9** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e.



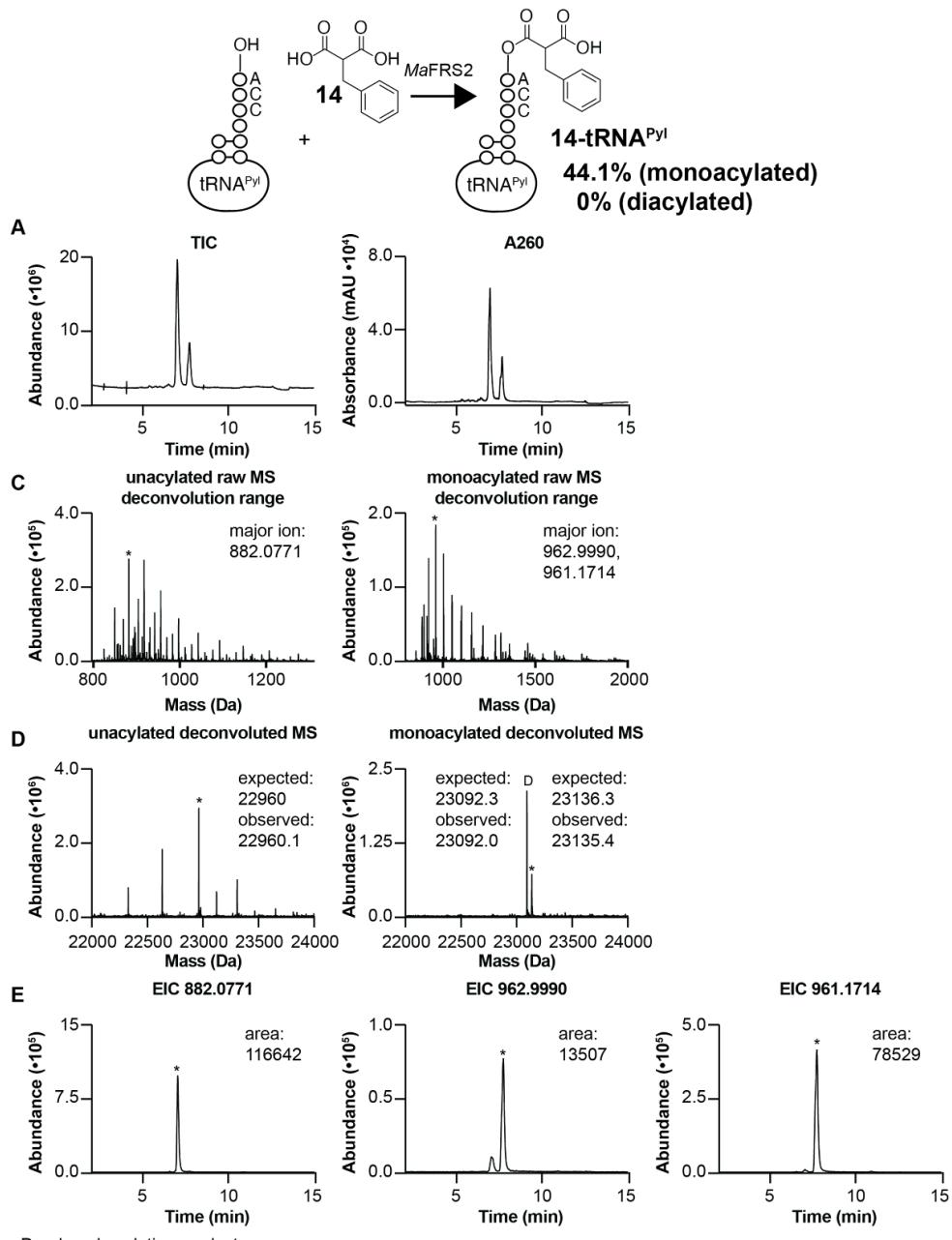
**Supplementary Figure 23.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **10** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



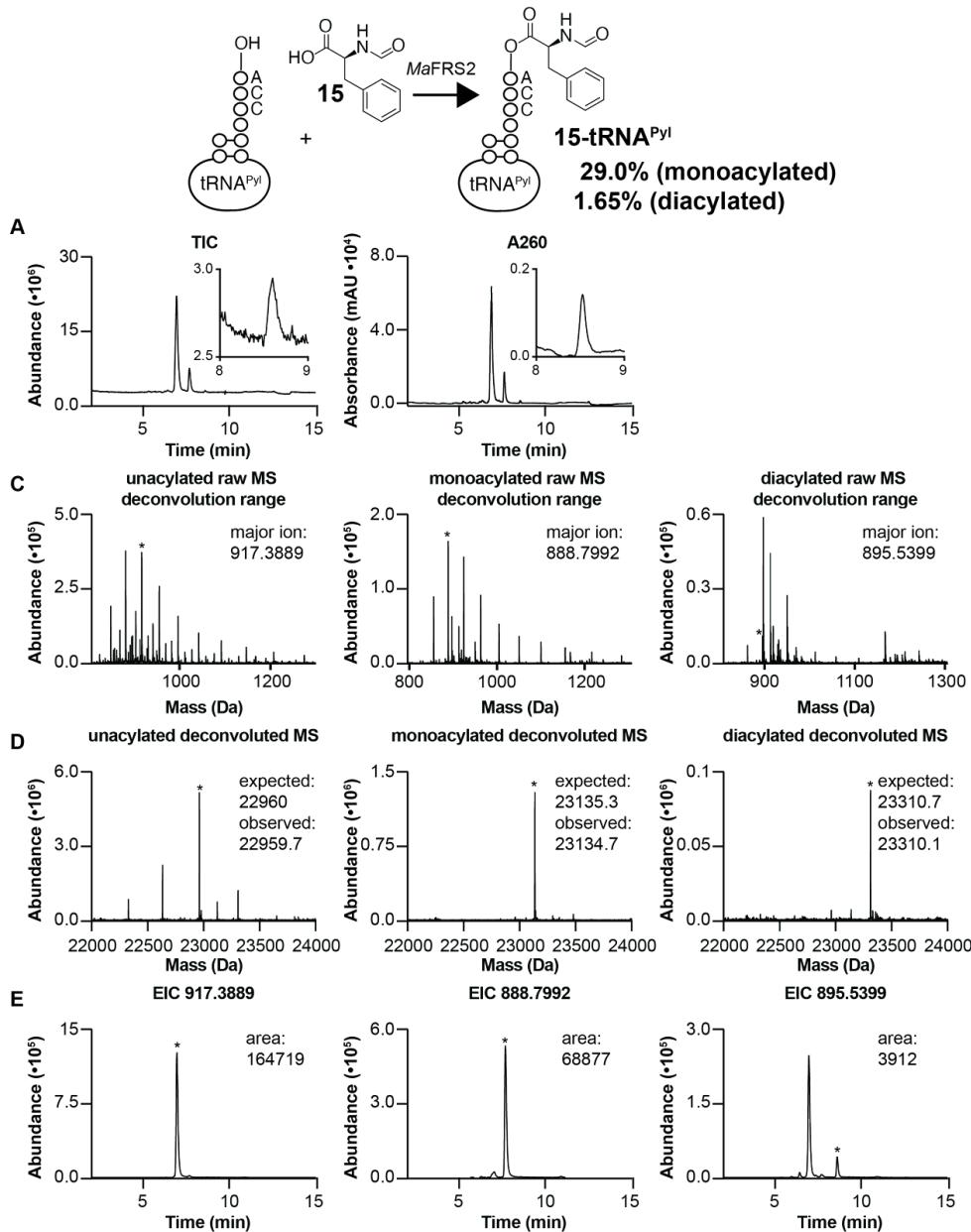
**Supplementary Figure 24.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **11** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



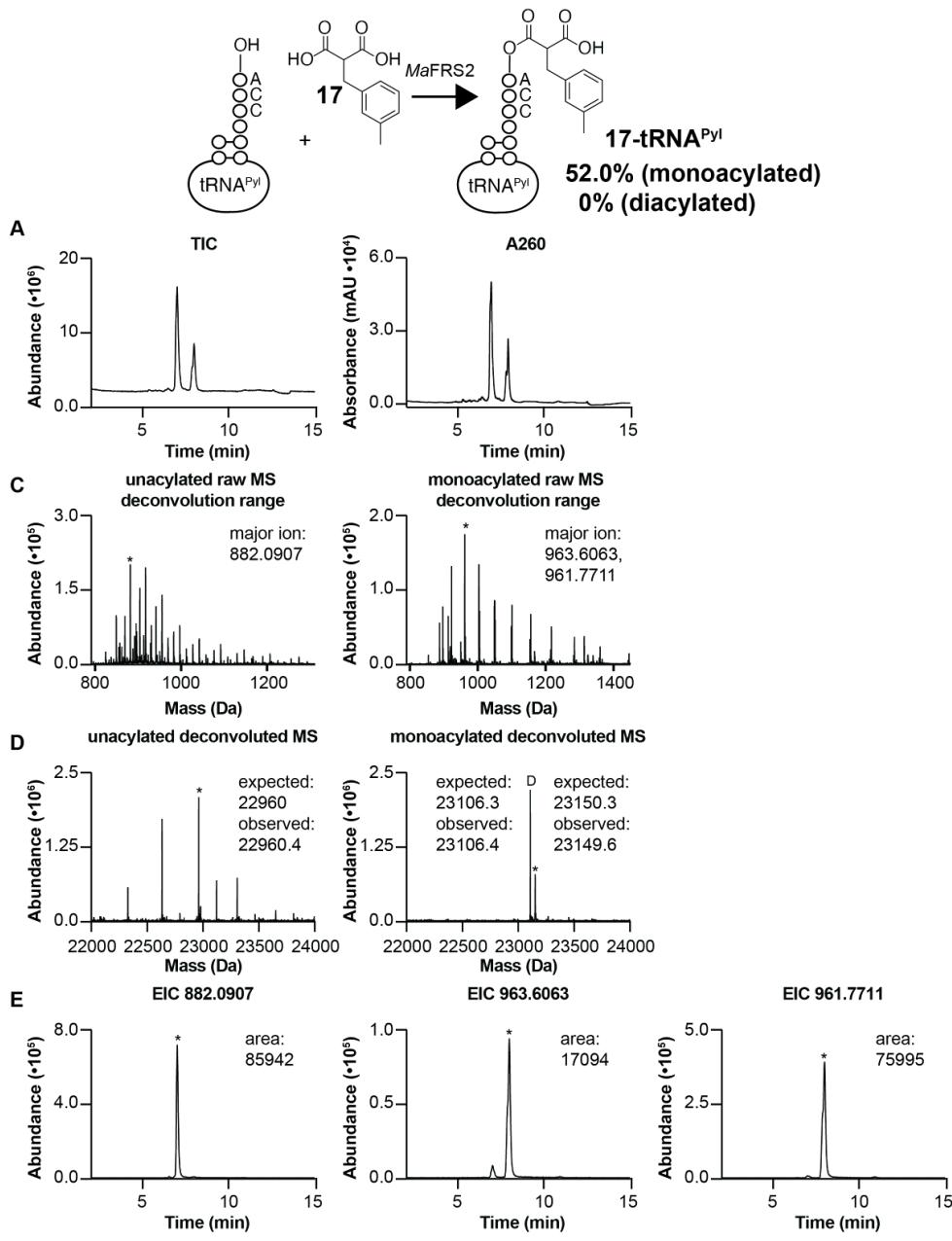
**Supplementary Figure 25.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **13** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



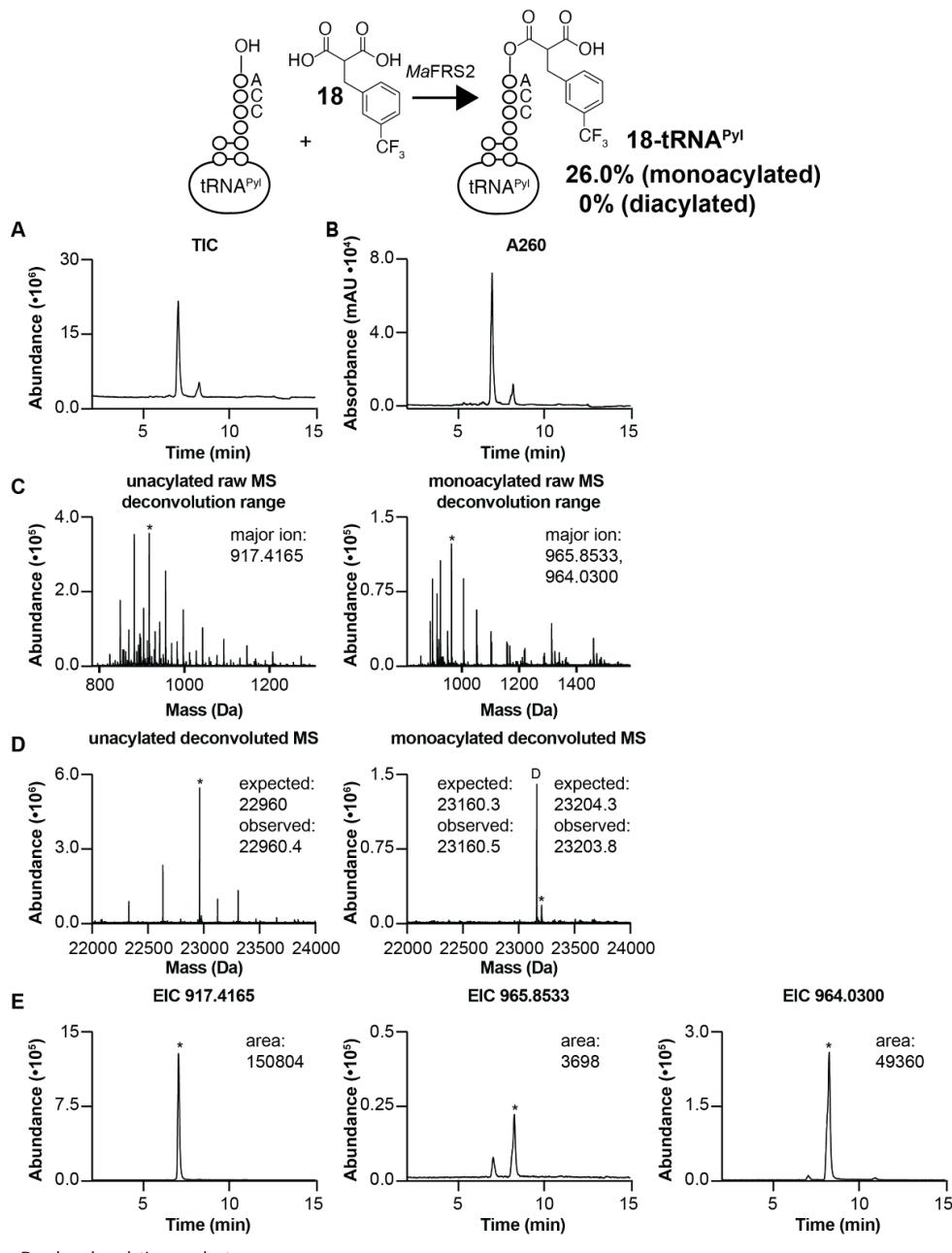
**Supplementary Figure 26.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **14** as described. Please refer to the legends for Supplementary Fig. 3 for descriptions of panels **a-e** and Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



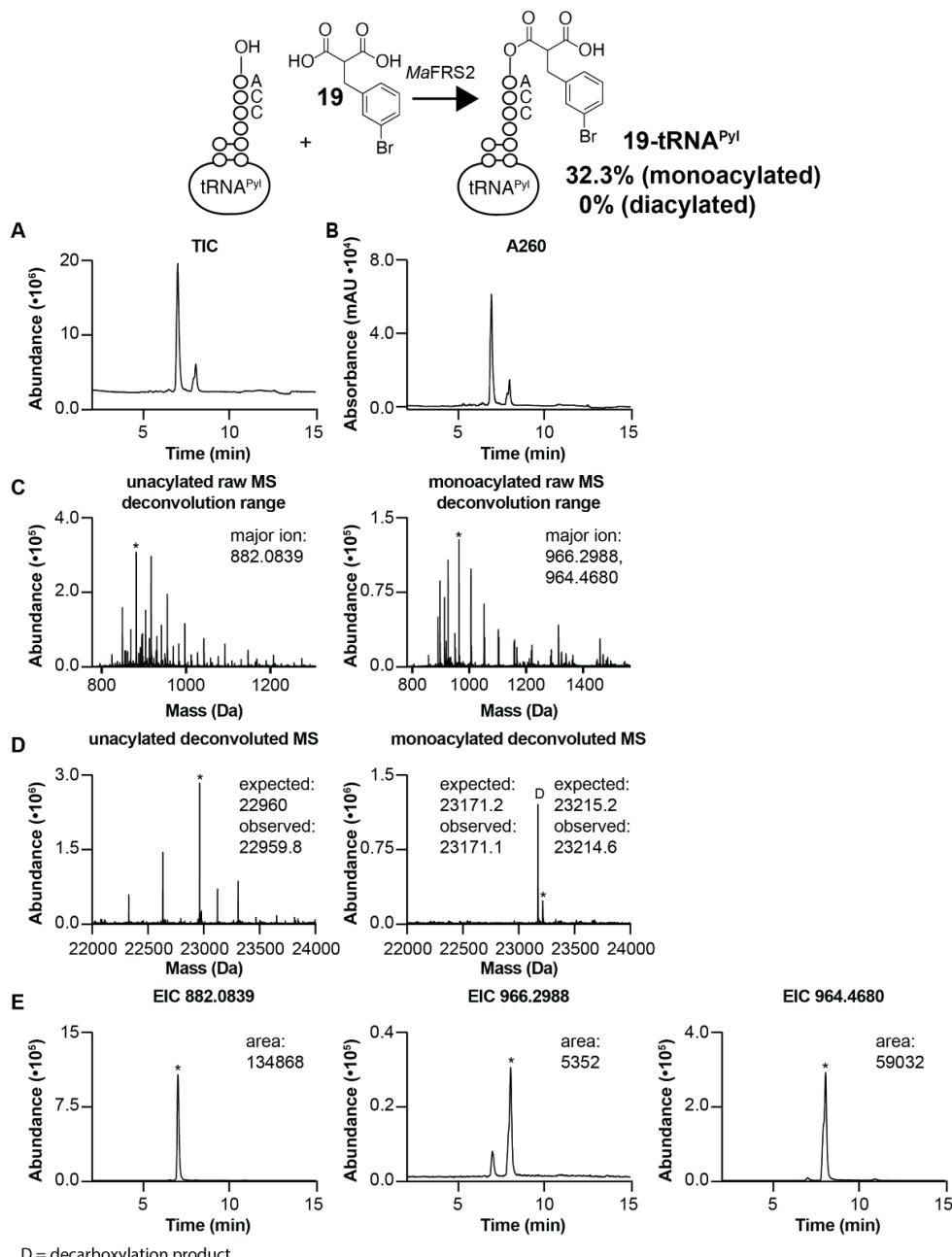
**Supplementary Figure 27.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **15** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels **a-e**.



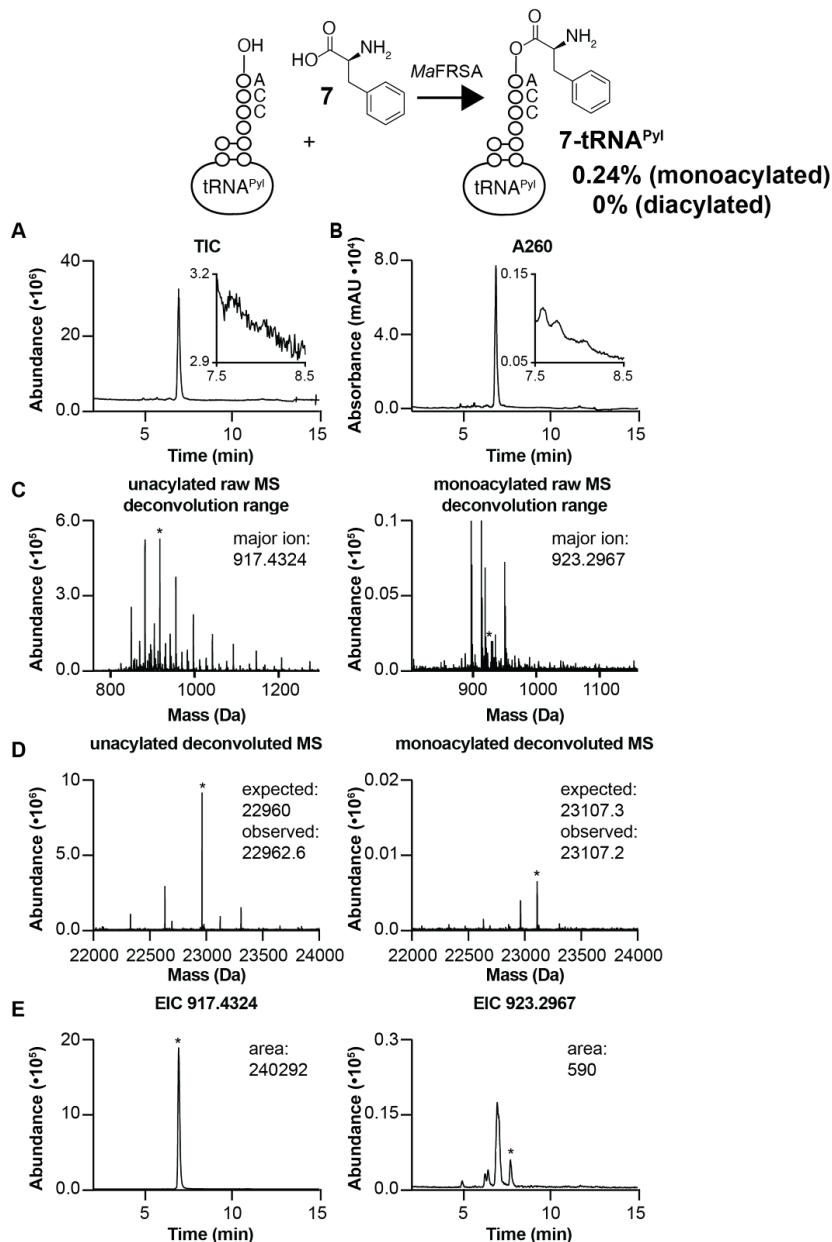
**Supplementary Figure 28.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **17** as described. Please refer to the legends for Supplementary Fig. 3 for descriptions of panels **a-e** and Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



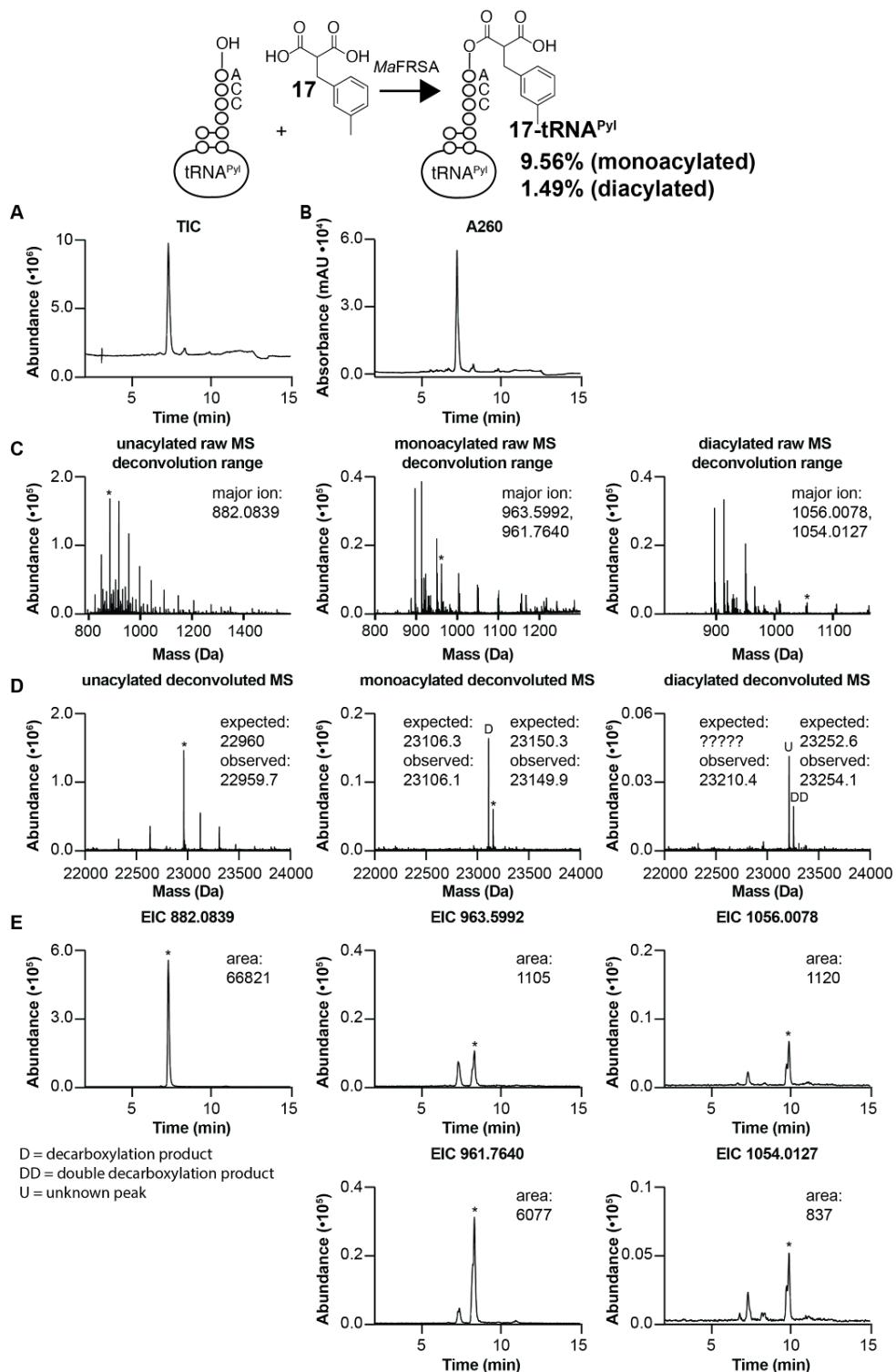
**Supplementary Figure 29.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **18** as described. Please refer to the legends for Supplementary Fig. 3 for descriptions of panels a-e and Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



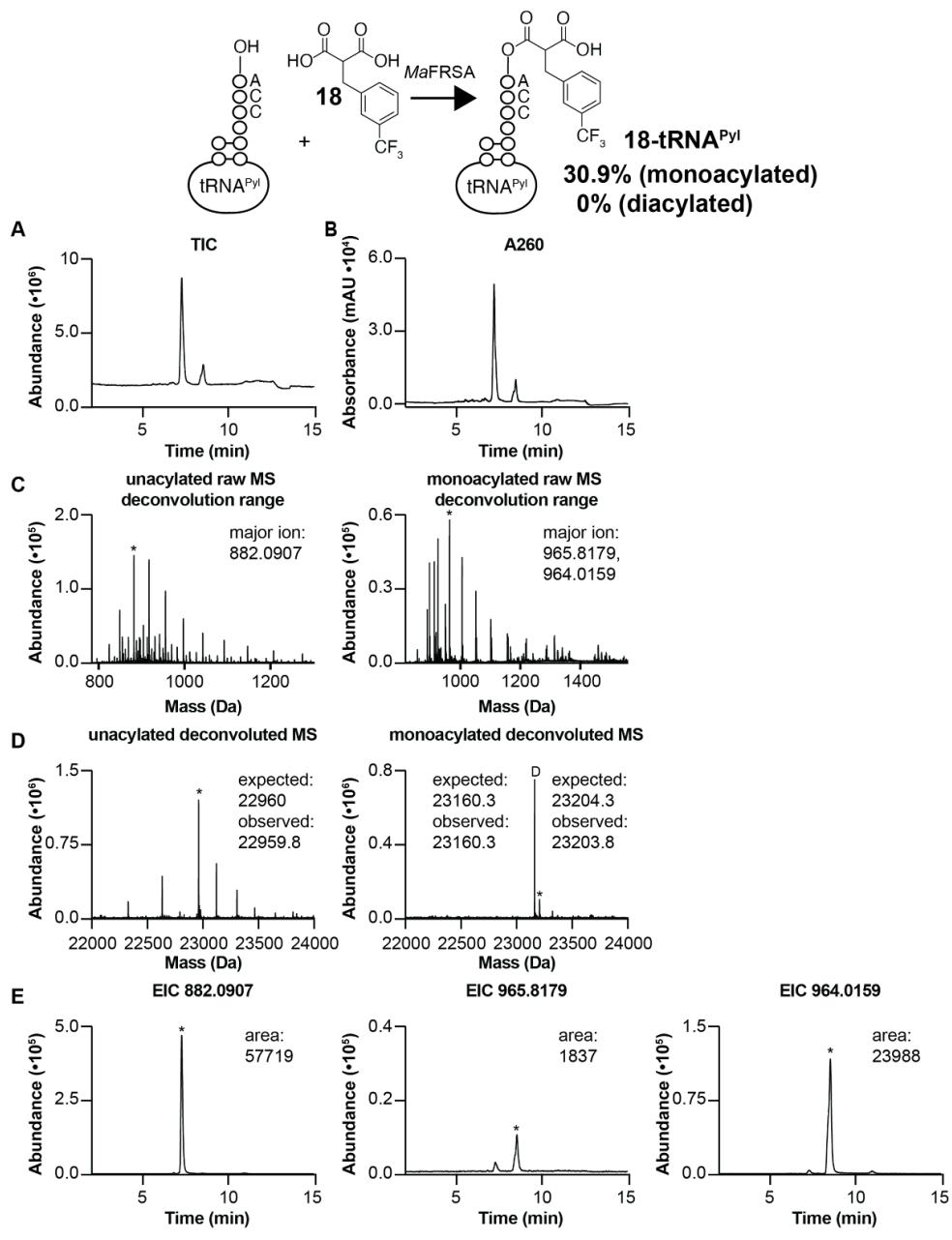
**Supplementary Figure 30.** Analysis of tRNA product mixtures obtained using *MaFRS2*, *Ma-tRNA<sup>Pyl</sup>* and monomer **19** as described. Please refer to the legends for Supplementary Fig. 3 for descriptions of panels **a-e** and Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



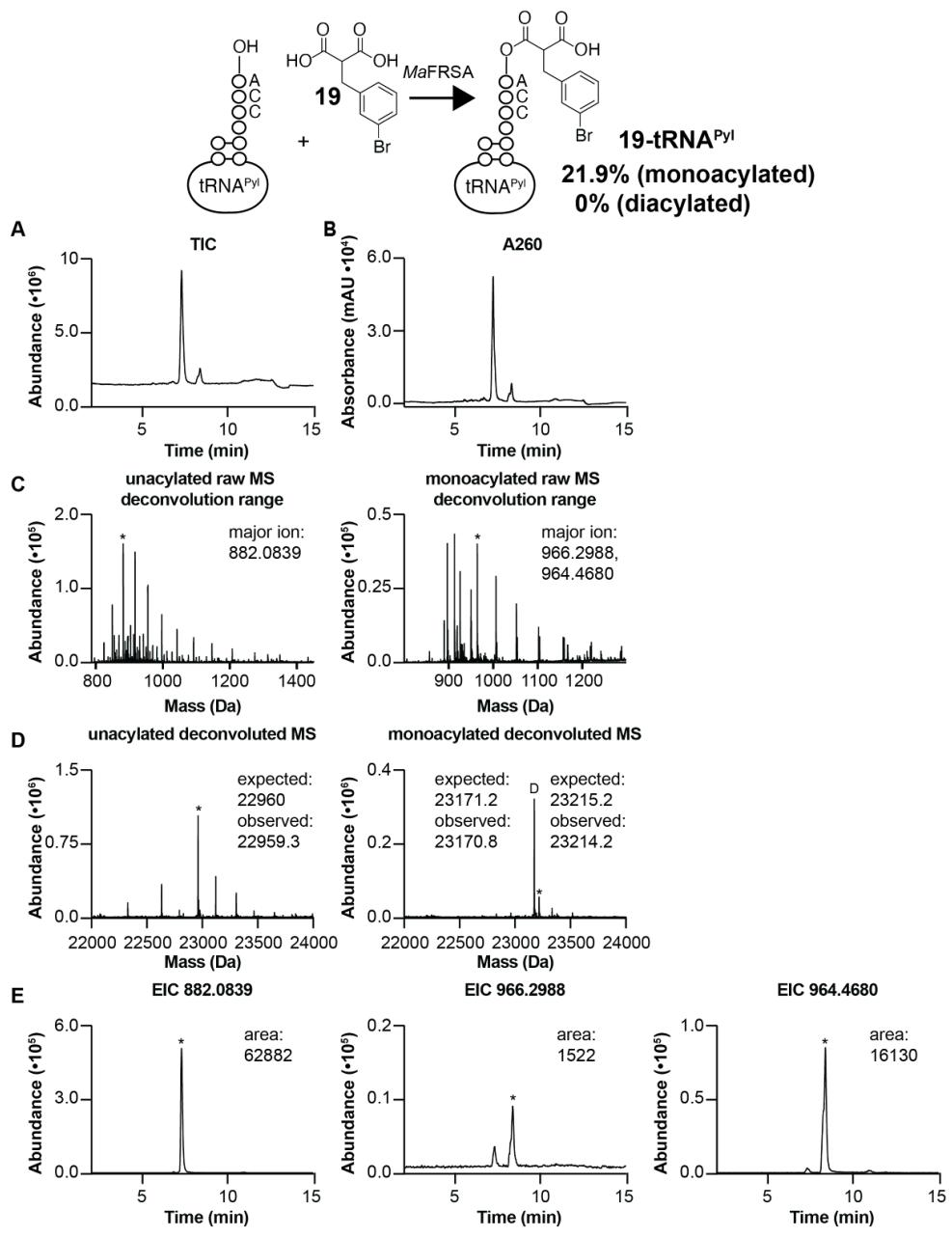
**Supplementary Figure 31.** Analysis of tRNA product mixtures obtained using *MaFRSA*, *Ma-tRNA<sup>Pyl</sup>* and monomer **7** as described. Please refer to the legend for Supplementary Fig. 3 for descriptions of panels a-e.



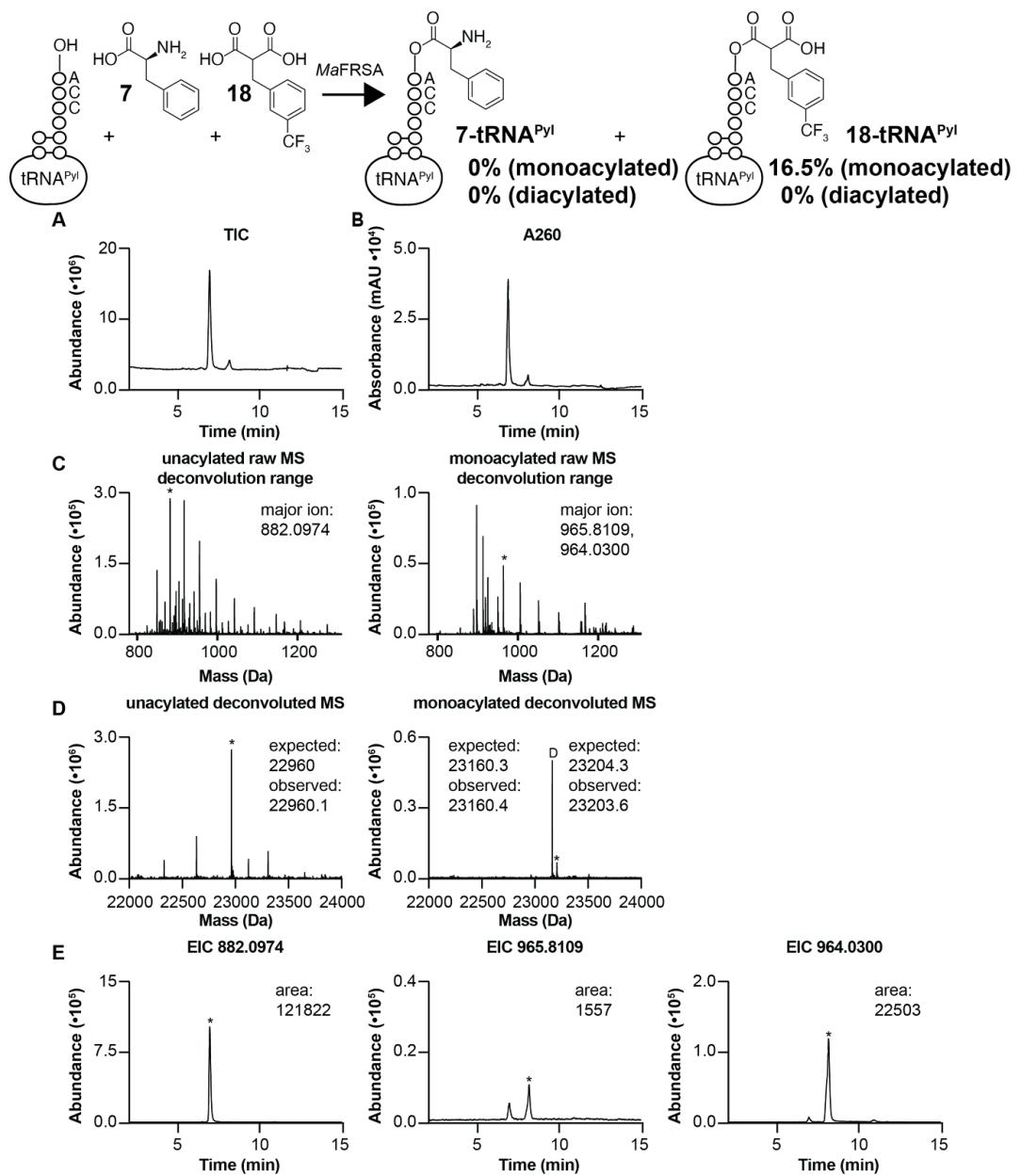
**Supplementary Figure 32.** Analysis of tRNA product mixtures obtained using *MaFRSA*, *Ma-tRNA<sup>Pyl</sup>* and monomer **17** as described. Please refer to the legends for Supplementary Fig. 3 for descriptions of panels **a-e** and Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



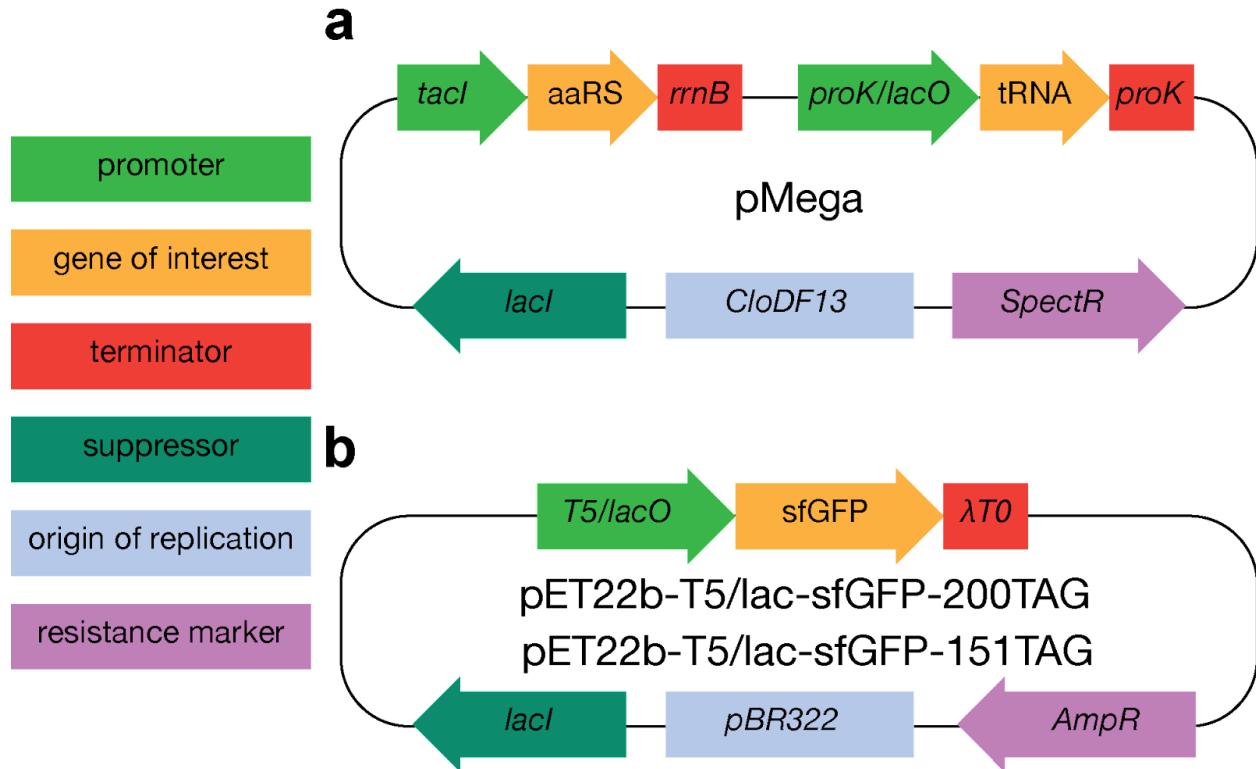
**Supplementary Figure 33.** Analysis of tRNA product mixtures obtained using *MaFRSA*, *Ma-tRNA<sup>Pyl</sup>* and monomer **18** as described. Please refer to the legends for Supplementary Fig. 3 for descriptions of panels **a-e** and Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.



**Supplementary Figure 34.** Analysis of tRNA product mixtures obtained using *MaFRSA*, *Ma-tRNA<sup>Pyl</sup>* and monomer **19** as described. Please refer to the legends for Supplementary Fig. 3 for descriptions of panels **a-e** and Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra.

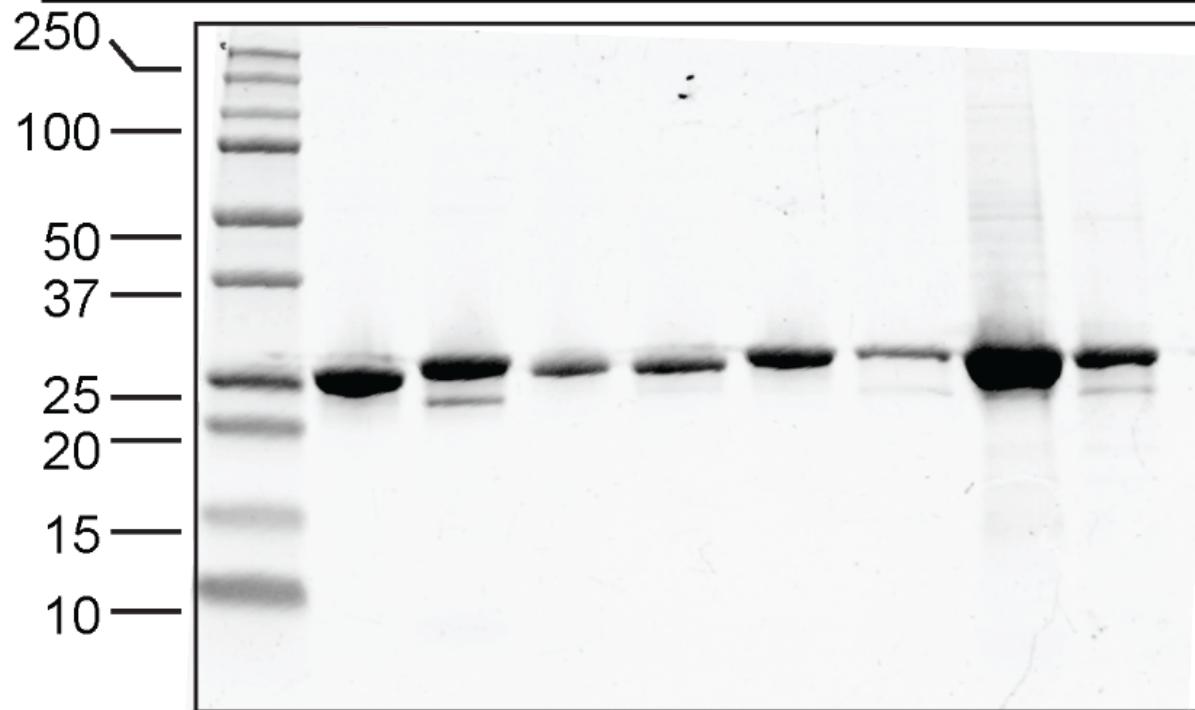


**Supplementary Figure 35.** Analysis of tRNA product mixtures obtained using *MaFRSA*, *Ma-tRNA<sup>Pyl</sup>* and monomers **7** and **18** as described. Please refer to the legends for Supplementary Fig. 3 for descriptions of panels **a-e** and Supplementary Fig. 7 for a note on the decarboxylation products observed in the mass spectra. Only the acylation product of **18** and *Ma-tRNA<sup>Pyl</sup>* is observed.

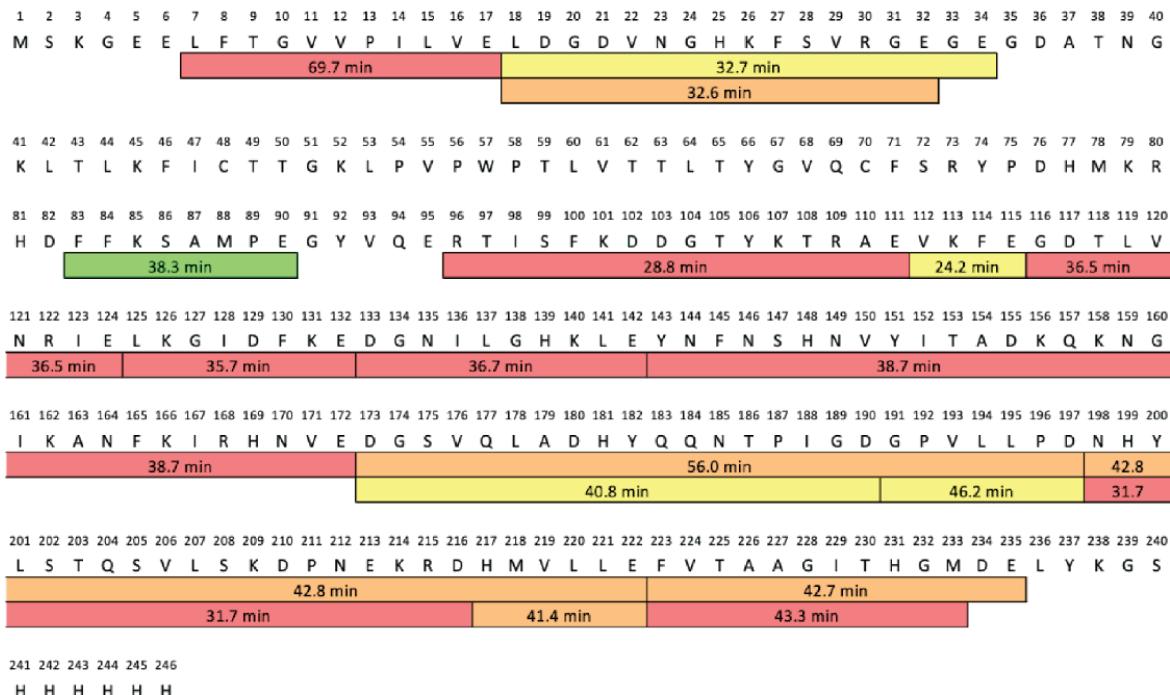
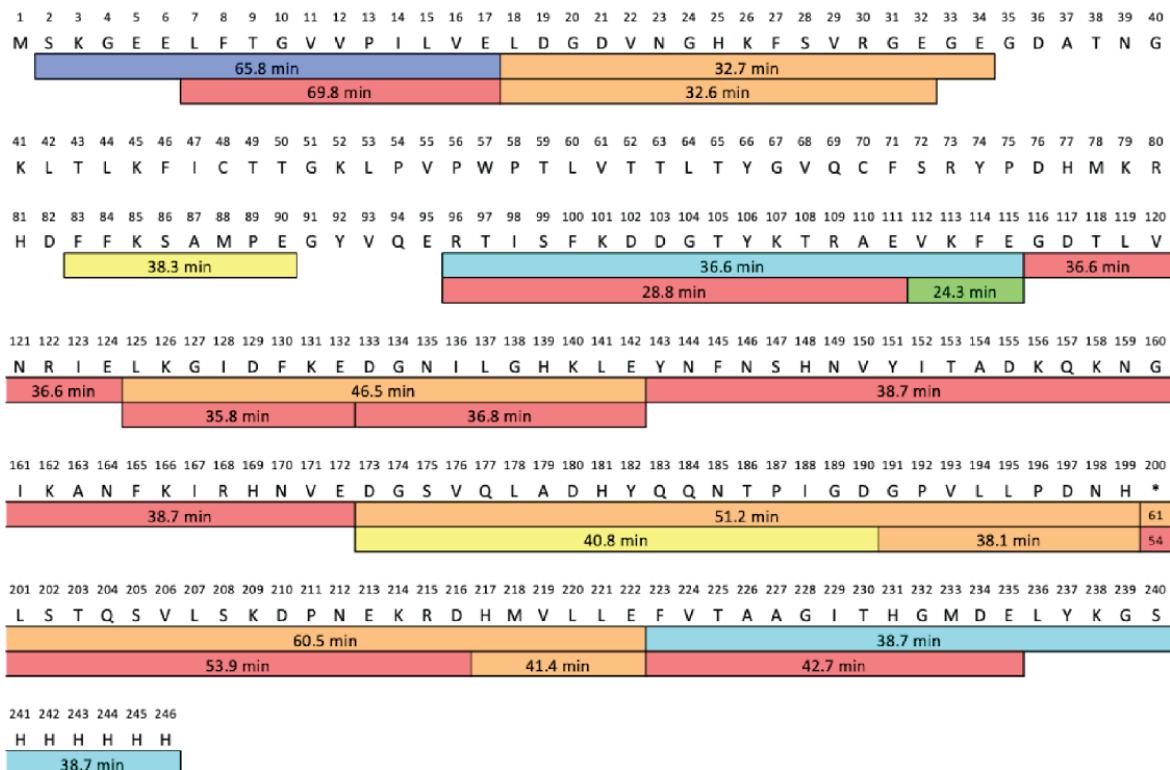


**Supplementary Figure 36.** Annotated maps of plasmids used for *in vivo* expression of sfGFP (Fig. 6c-e). **a**, pMega plasmids used for *MaPyIRS* or *MaFRSA* expression. **b**, Reporter plasmids used for expression of sfGFP with a TAG stop codon at position 200 or 151.

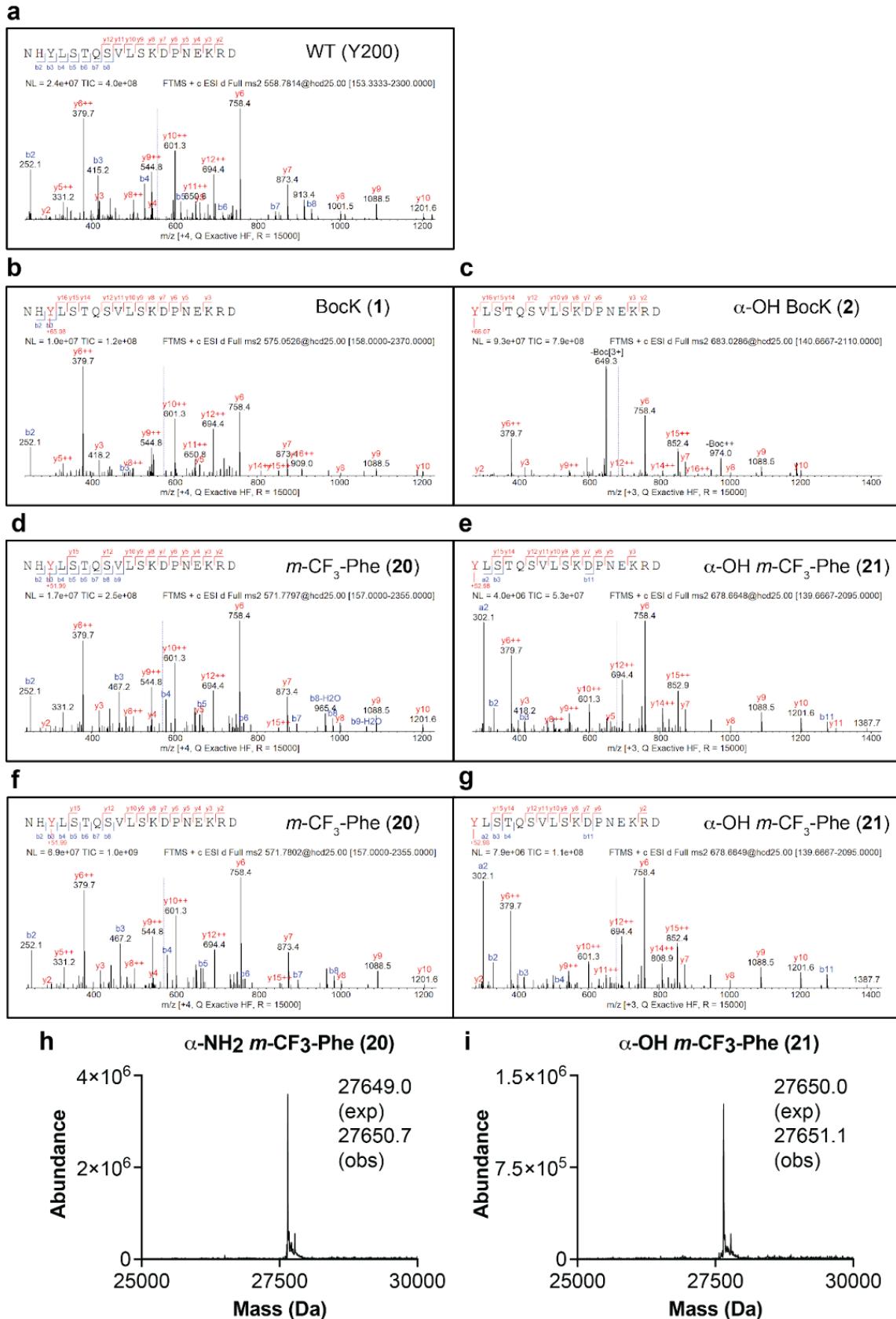
| monomer | <b>1</b> | <b>2</b> | <b>20</b> | <b>21</b> | <b>20</b> | <b>21</b> | <b>20</b> | <b>21</b> |
|---------|----------|----------|-----------|-----------|-----------|-----------|-----------|-----------|
| aaRS    | Pyl      | Pyl      | FA        | FA        | FA        | FA        | FA        | FA        |
| cells   | WT       | WT       | WT        | WT        | Δ         | Δ         | WT        | WT        |
| media   | TB       | TB       | TB        | TB        | TB        | TB        | DM        | DM        |



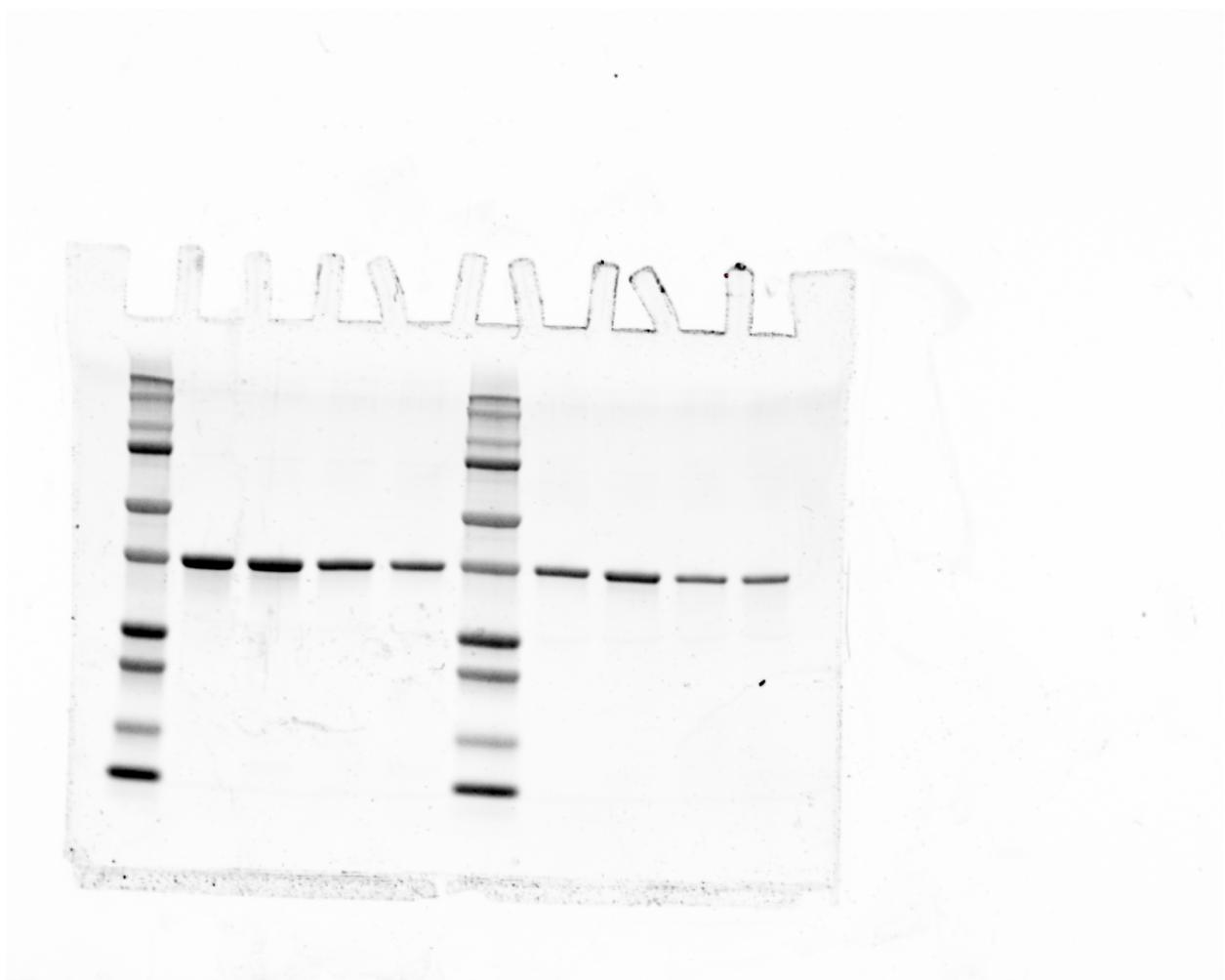
**Supplementary Figure 37.** Denaturing gel analysis of sfGFP variants. SDS-PAGE analysis of sfGFP variants expressed in DH10B or DH10B  $\Delta aspC \Delta tyrB$  in the presence of monomers **1**, **2**, **20**, or **21**. Abbreviations: “Pyl” = *MaPylRS*, “FA” = *MaFRSA*, “WT” = DH10b, “Δ” = DH10B  $\Delta aspC \Delta tyrB$ , “TB” = terrific broth, “DM” = defined media, +/- “OH-” = basic treatment or neutral control. Molecular weight ladder masses indicated at left in kDa.

**a****b**

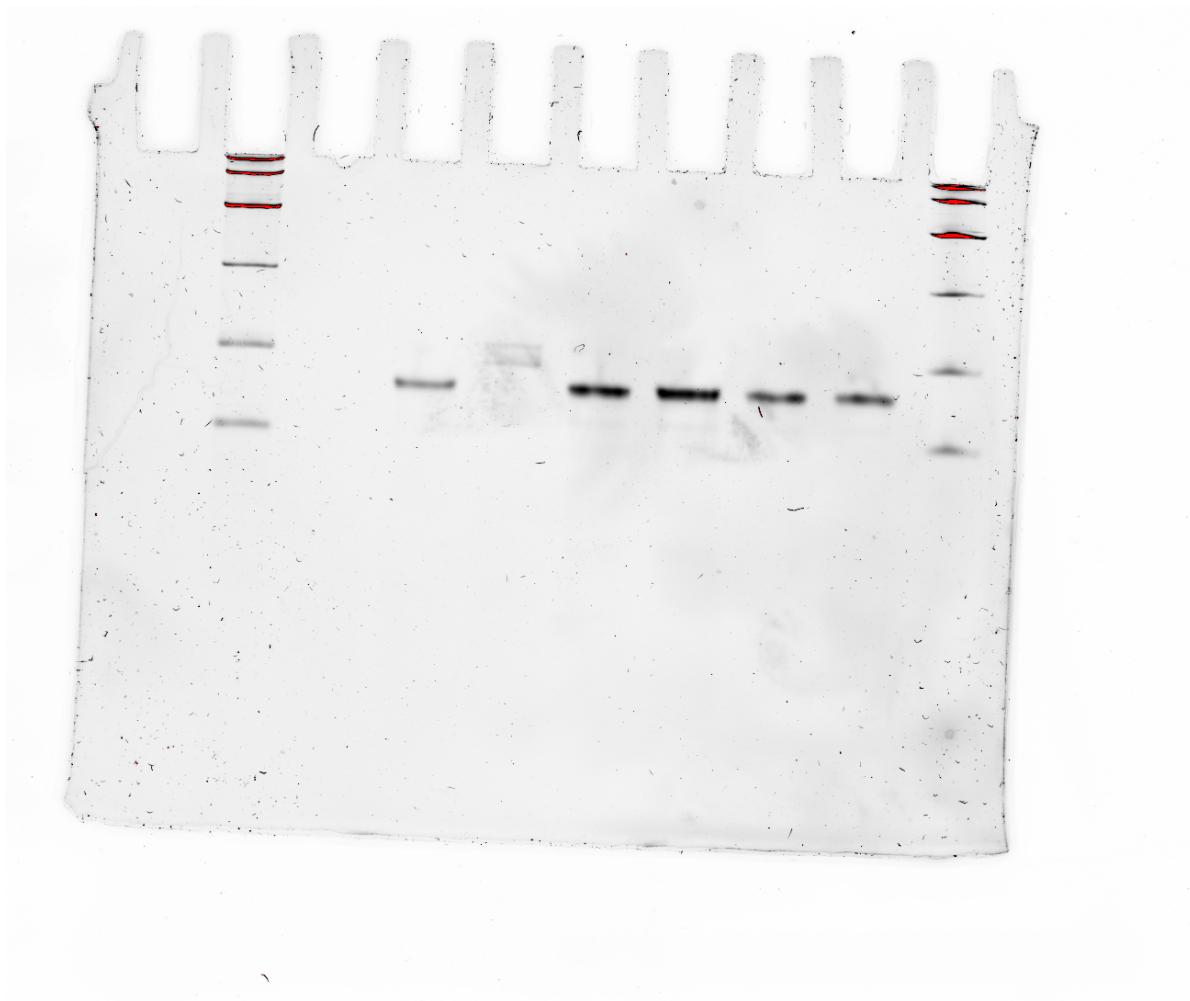
**Supplementary Figure 38.** Sequence of sfGFP illustrating the peptide fragments obtained after digestion with GluC and their retention times. **a**, Fragments expected when sfGFP contains Y, BocK (**1**), or *meta*-trifluoromethyl-L-phenylalanine (**20**) at position 200. Digestion with Glu-C generates two overlapping peptides containing position 200, those encompassing residues 198-216 and 198-222. Both were used to quantify the composition at position 200. **b**, Fragments expected when sfGFP contains  $\alpha$ -OH BocK (**2**) or  $\alpha$ -OH *meta*-trifluoromethyl phenylalanine (**21**) at position 200. In these cases, digestion with GluC exhibited additional cleavages at the proposed ester bond (presumably due to ester hydrolysis during work-up), generating two additional peptides containing position 200, encompassing residues 200-216 and 200-222. These peptides, in addition to those encompassing residues 198-216 and 198-222, were used to quantify the composition at residue 200. Colors from red to blue represent decreasing signal intensity. Retention times are indicated in the boxes illustrating the observed peptide fragments.



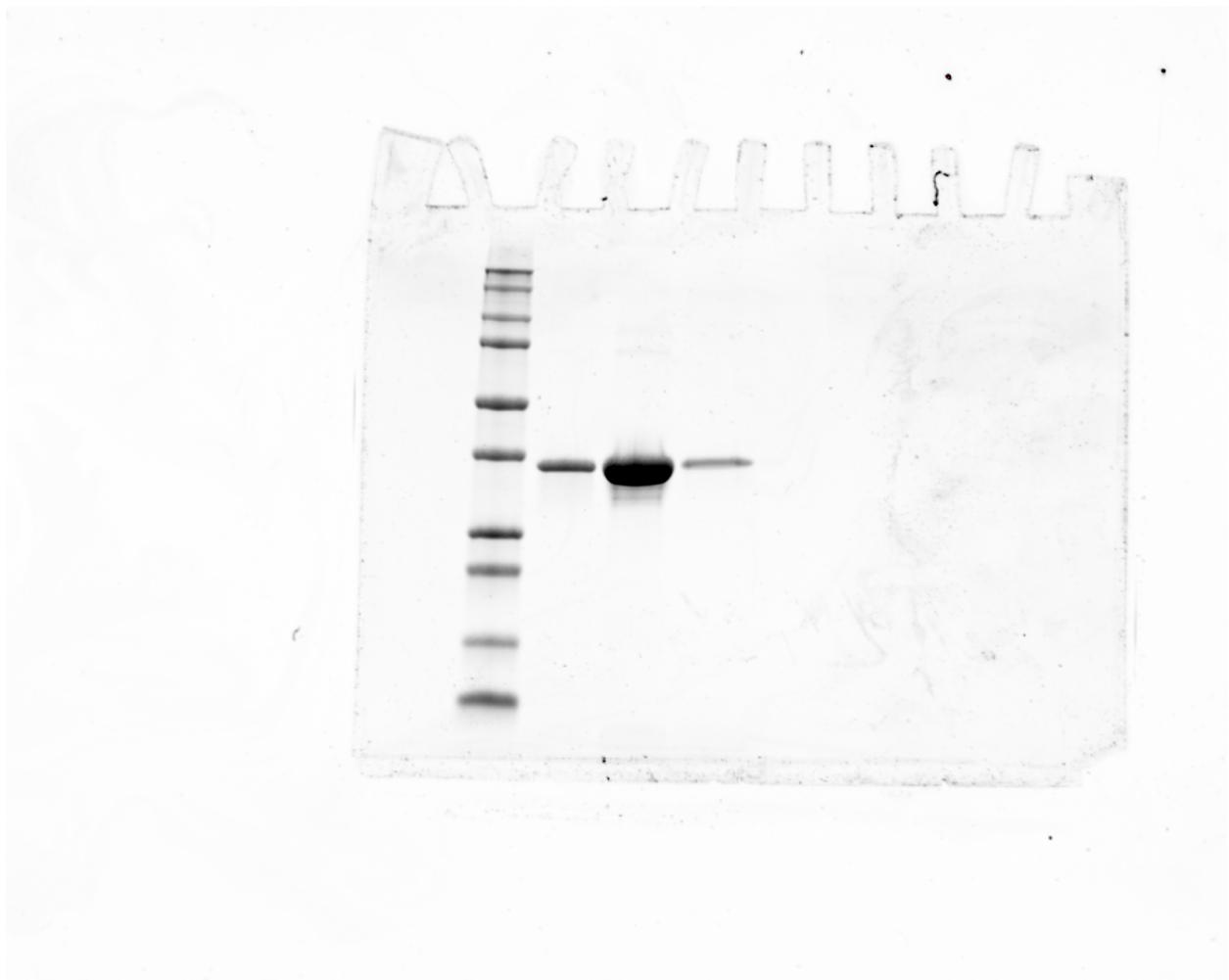
**Supplementary Figure 39.** Mass spectrometry confirms the presence of an ester at position 200 of sfGFP. MS/MS identification of peptide 198-216, sequence: NHXLSTQSVLSKDPNEKRD from sfGFP expressed in DH10B cells containing **a**, tyrosine (WT); **b**, BocK (**1**); or **c**,  $\alpha$ -OH BocK (**2**) at position 200. MS/MS identification of peptide 200-216, sequence: XLSTQSVLSKDPNEKRD resulting from sfGFP expressed in DH10B (**d**, **e**) or DH10B  $\Delta$ aspC  $\Delta$ tyrB (**f**, **g**) containing *meta*-trifluoromethyl-L-phenylalanine (**20**) (**d**, **f**) or  $\alpha$ -OH *meta*-trifluoromethyl-L-phenylalanine (**21**) (**e**, **g**) at position 200. Peptides were generated by endoproteinase Glu-C digestion of sfGFP samples expressed with each indicated substrate. For fragment assignments, position 200 was considered as a tyrosine (in red) modified to have the correct mass. Intact protein mass spectra of sfGFP variants purified from DH10B  $\Delta$ aspC  $\Delta$ tyrB cells co-expressing MaFRSA in the presence of 1 mM *meta*-trifluoromethyl phenylalanine (**20**, **h**), or  $\alpha$ -OH *meta*-trifluoromethyl phenylalanine (**21**, **i**).



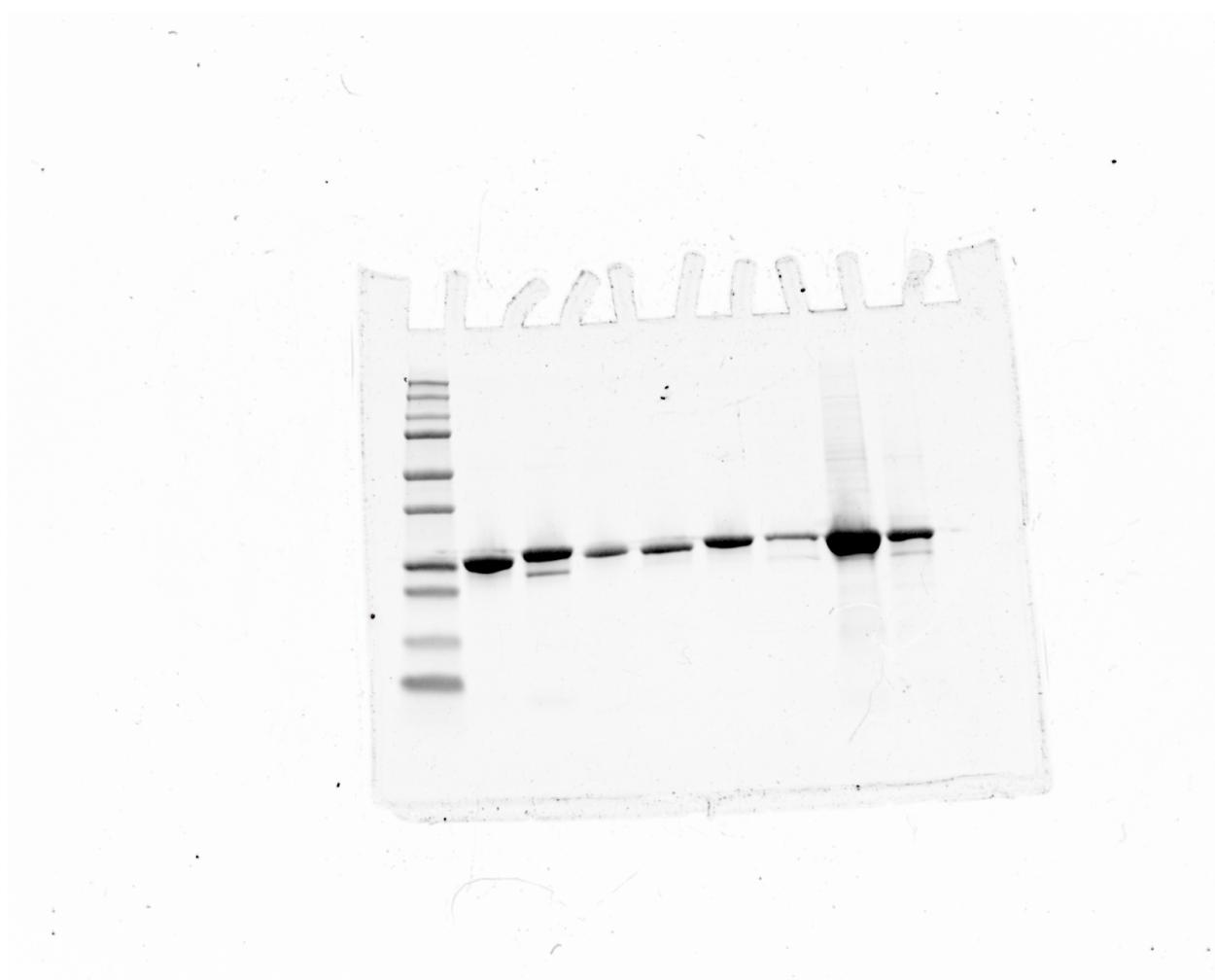
**Supplementary Figure 40.** Uncropped gel from Supplementary Fig. 2a.



**Supplementary Figure 41.** Uncropped gel from Supplementary Fig. 2d.



**Supplementary Figure 42.** Uncropped gel from Supplementary Fig. 2f.



**Supplementary Figure 43.** Uncropped gel from Supplementary Fig. 37.

## **Supplementary Tables**

**Supplementary Table 1.** Oligonucleotides used in this study

| Name    | Sequence (5'-3')  |
|---------|---|
| pET32a  | ATTTGTTAACCTAACAGAAGGAGATACATATGGCAGCAGCCATCATCAT   |
| MaPylRS | CATCATCACAGCAGCGGCCTGGTGCCGCGCAGCCATATGACGGTAAA   |
| gBlock  | TACACTGACGCACAGATCCAGCGCCTCGCGAATATGGAAATGGCACGTATG<br>AACAGAAAGTGTGCAAGATTGGCTCGCGACGCAGCCTTAGCAAAG<br>AAATGAGTGTGCCTCAACCGACAATGAGAAAAAAATTAAGGGCATGATTG<br>CCAACCCGTCACGTCATGGACTTACGCAACTTATGAACGACATTGCCGACGC<br>ATTAGTCGCTGAGGGATTATCGAGGTCCGACGCCAATCTTATCTCAAAA<br>GACGCGCTGCCGTATGACGATTACAGAAGACAAGCCCCTGTTCAAGCAA<br>GTATTCTGGATCGACGAGAACGCGTGCCTACGCCAATGTTGGCTCCAAATT<br>TATATTCCGTTATGCGTGATTGCGTGACCACACCGACGGCCAGTGAAGAT<br>TTTCGAGATGGGAGCTGTTCGCAAGGAAAGTCACAGTGGCATGCATT<br>GGAGGAGTTCACGATGCTAACCTGTGGATATGGACCGCGTGGTATGC<br>GACAGAGGTTTAAAAAATTACATTAGTGTGATGAAAGCAGCGGGATTG<br>CCCGATTATGATTAGTCCAGGAAGAGAGTGACGTCTACAAAGAAACTATCG<br>ATGTTGAGATTAACGGCAAGAAGTATGTAGCGCTGCTCGGACCCATTA<br>TCTGGATGCTGCCATGATGTGATGAACCTGGTCTGGTCTGGTTCGGT<br>TTGGAGCGCTTATTAACCATTGAGAAATATTCCACAGTAAAGAAAGGGG<br>GGGCAAGTATCTCGTACCTGAACGGTGCAAAATTAACATATGCACCA<br>TCACCACCATCATTCTCTGG |
| pET32a  | ATTTGTTAACCTAACAGAAGGAGATACATATGGCAGCAGCCATCATCAT   |

|                                       |   |
|---------------------------------------|---|
| <i>MaFRS1</i><br>gBlock               | CATCATCACAGCAGCGGCCTGGTGCCGCGCGCAGCCATATGACGGTGAAA<br>TACACTGACGCACAGATCCAGCGCCTCGCGAATATGGAATGGCACGTATG<br>AACAGAAAGTGTTCGAAGATTGGCTTCGCGCAGCAGCCTTAGCAAAG<br>AAATGAGTGTGCCTCAACCGACAATGAGAAAAAAATTAAGGGCATGATTG<br>CCAACCCGTCACGTCATGGACTTACGCAACTTATGAACGACATTGCCGACGC<br>ATTAGTCGCTGAGGGATTATCGAGGTCCGCACGCCAATCTTATCTCAAAA<br>GACGCGCTTGCCCGTATGACGATTACAGAAGACAAGCCCCTGTTCAAGCAA<br>GTATTCTGGATCGACGAGAAGCGTGCCTACGCCAATGTTGGCTCCAATT<br>TATATTCCGTTATGCGTGATTGCGTGACCACACCGACGGCCCAGTGAAGAT<br>TTTCGAGATGGGGAGCTGTTTCGCAAGGAAAGTCACAGTGGCATGCATT<br>GGAGGAGTTACGATGCTGGCGCTCTGGATATGGGACCGCGTGGTATG<br>GACAGAGGTTTAAAAAATTACATTAGTGTGATGAAAGCAGCGGGATTG<br>CCCGATTATGATTAGTCCAGGAAGAGAGTGACGTCTACAAAGAAACTATCG<br>ATGTTGAGATTAACGGGCAAGAAGTATGTAGCGCTGCTCGGACCCATTA<br>TCTGGATGCTGCCATGATGTGCATGAACCTGGTCTGGTCTGGTTTCGGT<br>TTGGAGCGCTTATTAACCATTCTGAGAAATATTCCACAGTAAAGAAAGGGG<br>GGGCAAGTATCTCGTACCTGAACGGTGCAAAATTAACATATGCACCA<br>TCACCACCATCATTCTCTGG |
| <i>MaFRSA</i><br>fusion tag<br>gblock | TGTGAGCGGATAACAATTCCCCTCTAGAAATAATTGTTAACTTAAAGAAG<br>GAGATATACAAT <u>GGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGG</u><br><u>TGCCGCGCGCAGCCATATGACGGTGAAATACACTGACGCACAGATCCAGC</u><br>GCCTCGCGAATATGGGAATGGCACCGTA   |
| pMega                                 | CAATTTCACAAAGGAGGTGCGGCCGCATGACAGTCAAATATACCGACGCC  |

|                |  |
|----------------|--|
| <i>MaPyIRS</i> | AGATTCA CGCTT GCGC GAGTACGGCAACGGTACGTATGAGCAAAAGGTTT  |
| gBlock         | TTGAAGACCTTGCTAGTCGCGACGCCGTT CAGCAAGGAAATGAGTGTG<br>CCTCCACAGATAATGAAAAGAAAATCAAAGGCATGATGCCAACCCCTCTC<br>GTCACGGCTTGACCCAGCTGATGAACGATATTGCAGATGCGTTGGTAGCTGA<br>GGGGTTCAT TGAGGTGCGTACT CCTATTTCATTAGTAAGGACGCCCTGCA<br>CGCATGACTATTACGGAGGATAAACCTTGTCAAACAGGTCTCTGGATCG<br>ATGAGAAGCGTGC ACTGCGCCCCATGCTGGCACCCAAC TTATAGTGTCA<br>GCGCGATTACGCGATCACACTGACGGTCCGGTAAAATT TGAGATGGC<br>TCATGCTCCGCAAAGAGTCGCATTCTGGTATGCACCTGGAAGAGTTACCA<br>TGTAAACCTGGTAGACATGGGCCACGCCGGACGCTACGGAGGTCTGA<br>AGAACTACATTAGCGTCGTGATGAAAGCAGCCGGTTACCCGATTACGATCT<br>TGTT CAGGAAGAGAGTGACGTTACAAAGAAACAATCGATGTTGAAATCAA<br>CGGTCAAGAAGTTGTT CAGCGGCCGTGGGGCCGCATTATTAGACGCTGC<br>CCACGATGTACATGAGCCGTGGAGTGGTGCAGGCTCGGTCTGAGCGCTT<br>GCTGACCATT CGTGAGAAATATAGCACTGTAAAAAAGGTGGCGCCTCCAT<br>CTCTTATCTTAATGGAGCTAAATCAACTAAGCGGCCGCGTTAACGGTCT<br>CCAGCTTGGCTGTTGGCGGATGAGAGAAGATTTCAGCCTGATACAGATT<br>AAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTGCGCTGGCGCAGT<br>AGCGCGGTGGTCCCACCTGACCCATGCCGAACTCAGAAGTGAACGCCGT<br>AGCGCCGATGGTAGTGTGGGGTCTCCCCTGCGAGAGTAGGAACTGCCAG<br>GCATCAAATAAACGAAAGGCTCAGTCGAAAGACTGGCCTGTTGTGAG<br>CTCCCGGT CATCAATCATCCCCATAATCCTGTTAGCCTGCAGGTAATTCCGC<br>TTCGCAACATGTGAGCACC GGTTATTGACTACCGGAAGCAGTGTGACCGT<br>GTGCTTCTCAAATGCC TGAGGCCAGTTGCTCAGGCTCTCCCGTGGAGGTA |

|                  |   |
|------------------|---|
|                  | ATAATTGACGATATGATCAGTGCACGGCTAACTAAGCGGCCTGCTGACTTTCTGCCGATAAAAGGCATTTGCTATTAAGGGATTGACGAGGGCGTATCTGCAGTAAGATAATTGTGAGCGGATAACAATTAGCAGACAAGATGGGTCCCTTATCATGGCAACCATCTGAACGGGGACGGTCCGGCACCAGGGTCTCTAACACCTAGCCAGCGGGTTCGACGCCCGGTCTCGCAAATTGAAAGCCTGCTCAACGAGCAGGCTTTTGATGCTCGAGCAGCTCAGGGTCGAA |
| RF31             | GCTGGCGCTTAAGGATATGGGA  |
| RF32             | ATCGTGAACTCCTCCAAATG  |
| RF33             | GCTGGCGCTTGCAGGATATGGGA   |
| RF48             | TGGTATGCACCTGGAAGAGTTACCATGTTAGCGCTGGCGGACATGGGCCACGCGGGGA  |
| RF49             | TCCCCGCGTGGCCCCATGTCCGCCAGCGCTAACATGGTAAACTCTTCCAGGTGCATACCA  |
| RF61             | GACATGGGGCACGC  |
| RF62             | TAACATGGTAAACTCTTCCAGGTGCA  |
| CS43             | AGACAACCATTAGCTGTCGACACAATC   |
| CS44             | GGTAAAAGGACAGGGCCA  |
| T7 F             | TAATACGACTCACTATAGGG  |
| T7 R             | GCTAGTTATTGCTCAGCGG   |
| <i>Ma-PyIT-F</i> | CTAATACGACTCACTATAGGGGACGGTCCGGCACCAGGGTCTCTAAA   |

|                                      |   |
|--------------------------------------|---|
|                                      | ACCTAGCCA   |
| <i>Ma-PylT-R</i>                     | TmGGCGAGAGACCAGGGCGTCGAACCCGCTGGCTAGGTTAGAGACCC<br>GCTGGTCGCCG                    |
| <i>Ma-tRNA<sup>Pyl</sup></i>         | GGGGGACGGGUCCGGCGACCAGCGGGUCUAAAACCUAGCCAGCGGGGU<br>UCGACGCCCGGUCUCUCGCCA         |
| <i>Ma-PylT-AC</i><br>C-F             | CTAATACGACTCACTATAAGGGGACGGTCCGGCGACCAGCGGGTCTACCAA<br>ACCTAGCCA                  |
| <i>Ma-PylT-AC</i><br>C-R             | TmGGCGAGAGACCAGGGCGTCGAACCCGCTGGCTAGGTTGGTAGACCC<br>GCTGGTCGCCG                   |
| <i>Ma-tRNA<sup>Pyl-</sup></i><br>ACC | GGGGGACGGTCCGGCGACCAGCGGGTCTACCAAACCTAGCCAGCGGGGTT<br>CGACGCCCGGTCTCTCGCCA        |
| MGV-flag-1                           | GCGAATTAAATACGACTCACTATAAGGGTTAACTTTAACAAAGGAGAAAAACATG<br>GGTGTGACTACAAGGACGACGA |
| MGV-flag-2                           | AAACCCCTCCGTTAGAGAGGGGTATGCTAGTTACTGTCGTCGTCCT<br>TGTAGTCGACACCCATGTTTC           |

\*mG represents 2'-O-methyl-deoxymethylguanosine

**Supplementary Table 2.** Expected exact masses of acyl-adenosine nucleosides extracted in LC-HRMS analysis of acyl-tRNA products digested by RNase A.

| Substrate used   | Exact mass |
|--|------------|
| $\text{N}^{\varepsilon}\text{-}(\text{tert-butoxycarbonyl})\text{-L-lysine}$ ( <b>1</b> )        | 496.25142  |
| (S)-6-((tert-butoxycarbonyl)amino)-2-hydroxyhexanoic acid ( <b>2</b> )                           | 497.23544  |
| 6-((tert-butoxycarbonyl)amino)hexanoic acid ( <b>3</b> )   | 481.24052  |
| $\text{N}^{\varepsilon}\text{-}(\text{tert-butoxycarbonyl})\text{-D-lysine}$ (D-BocK, <b>5</b> ) | 496.25142  |
| 2-(4-((tert-butoxycarbonyl)amino)butyl)malonic acid ( <b>16</b> )                                | 525.23035  |
| L-phenylalanine ( <b>7</b> )   | 415.17244  |
| (S)-2-hydroxy-3-phenylpropanoic acid ( <b>8</b> )  | 416.15646  |
| 3-phenylpropanoic acid ( <b>9</b> )  | 400.16155  |
| N-methyl-L-phenylalanine ( <b>10</b> )   | 429.18809  |
| D-phenylalanine ( <b>11</b> )  | 415.17244  |
| 2-mercaptopropanoic acid ( <b>13</b> )   | 432.13362  |
| 2-benzylmalonic acid ( <b>14</b> )   | 444.15137  |
| N-formyl-L-phenylalanine ( <b>15</b> )   | 443.16736  |
| 2-(3-methylbenzyl)malonic acid ( <b>17</b> )   | 458.16702  |

|   |           |
|---|-----------|
| 3-( <i>meta</i> -tolyl)propanoic acid (decarboxylation product of <b>17</b> )       | 414.17720 |
| 2-(3-trifluoromethylbenzyl)malonic acid ( <b>18</b> )                               | 512.13876 |
| 3-(3-(trifluoromethyl)phenyl)propanoic acid (decarboxylation product of <b>18</b> ) | 468.14893 |
| 2-(3-bromobenzyl)malonic acid ( <b>19</b> )   | 522.06189 |
| 3-(3-bromophenyl)propanoic acid (decarboxylation product of <b>19</b> )             | 478.07206 |

**Supplementary Table 3.** Yields of mono- and diacylated *Ma*-tRNA<sup>PyL</sup> calculated from intact tRNA analysis as shown in Supplementary Figs. 3-35 and performed as described in Section IV. The number on top in each box represents the average of 3 technical replicates, while the number on the bottom is the standard deviation. ND = not determined.

| Enzyme    | <i>MaPylRS</i> |              | <i>MaFRS1</i> |             | <i>MaFRS2</i> |              | <i>MaFRSA</i> |        |
|-----------|----------------|--------------|---------------|-------------|---------------|--------------|---------------|--------|
| Substrate | mono           | di           | mono          | di          | mono          | di           | mono          | di     |
| <b>1</b>  | 53.8<br>0.3    | 0.0<br>0.0   | ND            | ND          | ND            | ND           | ND            | ND     |
| <b>2</b>  | 76.8<br>0.3    | 13.7<br>0.2  | ND            | ND          | ND            | ND           | ND            | ND     |
| <b>3</b>  | 59.1<br>0.5    | 2.51<br>0.09 | ND            | ND          | ND            | ND           | ND            | ND     |
| <b>5</b>  | 0<br>0         | 0<br>0       | ND            | ND          | ND            | ND           | ND            | ND     |
| <b>7</b>  | ND             | ND           | 36.2<br>1.1   | 29.9<br>0.7 | 47.7<br>0.3   | 17.7<br>0.3  | 0.24<br>0.01  | 0<br>0 |
| <b>8</b>  | ND             | ND           | 26.4<br>1.0   | 21.0<br>0.3 | 36.5<br>0.5   | 2.78<br>0.23 | ND            | ND     |
| <b>9</b>  | ND             | ND           | 0.97<br>0.03  | 0<br>0      | 1.89<br>0.02  | 0<br>0       | ND            | ND     |
| <b>10</b> | ND             | ND           | 1.73          | 0           | 0.55          | 0            | ND            | ND     |

|           |              |            |              |              |              |              |              |              |
|-----------|--------------|------------|--------------|--------------|--------------|--------------|--------------|--------------|
|           |              |            | 0.06         | 0            | 0.02         | 0            |              |              |
| <b>11</b> | ND           | ND         | 0<br>0       | 0<br>0       | 0<br>0       | 0<br>0       | ND           | ND           |
| <b>13</b> | ND           | ND         | 0*<br>0      | 0*<br>0      | 2.08<br>0.06 | 0<br>0       | ND           | ND           |
| <b>14</b> | ND           | ND         | 24.4<br>0.8  | 0<br>0       | 43.7<br>0.5  | 0<br>0       | ND           | ND           |
| <b>15</b> | ND           | ND         | 9.52<br>0.28 | 40.5<br>0.5  | 29.3<br>0.3  | 1.80<br>0.15 | ND           | ND           |
| <b>16</b> | 0.25<br>0.02 | 0.0<br>0.0 | ND           | ND           | ND           | ND           | ND           | ND           |
| <b>17</b> | ND           | ND         | 49.8<br>0.7  | 3.64<br>0.35 | 52.0<br>0.1  | 0<br>0       | 9.48<br>0.07 | 1.54<br>0.10 |
| <b>18</b> | ND           | ND         | 59.3<br>1.0  | 9.24<br>0.60 | 26.1<br>0.4  | 0<br>0       | 31.4<br>0.7  | 0<br>0       |
| <b>19</b> | ND           | ND         | 37.4<br>0.7  | 10.8<br>1.4  | 32.7<br>0.7  | 0<br>0       | 22.3<br>0.4  | 0<br>0       |

\*No product was observed when **13** was incubated with 2.5  $\mu$ M *MaFRS1* and 25  $\mu$ M *Ma-tRNA<sup>Pyl</sup>*. Increasing [*MaFRS1*] to 12.5  $\mu$ M led to  $0.72 \pm 0.02\%$  and  $9.72 \pm 0.31\%$  of mono- and diacylated tRNA, respectively.

**Supplementary Table 4.** Structure refinement statistics.

| <i>MaFRSA, meta-CF<sub>3</sub>-2-BMA, AMP-PNP</i> |                                    |
|---|------------------------------------|
| Wavelength (Å)                                    | 1.11583                            |
| Resolution range (Å)                              | 45.37 - 1.803 (1.867 - 1.803)      |
| Space group                                       | I 4                                |
| Unit cell dimensions (Å)                          | 108.958 108.958 112.26 90° 90° 90° |
| Total reflections                                 | 115764 (9210)                      |
| Unique reflections                                | 57972 (4659)                       |
| Multiplicity                                      | 2.0 (2.0)                          |
| Completeness (%)                                  | 95.02 (66.69)                      |
| Mean I/sigma (I)                                  | 18.75 (0.43)                       |
| Wilson B-factor                                   | 39.02                              |
| R-merge   | 0.02241 (1.976)                    |
| R-meas  | 0.03169 (2.794)                    |
| R-pim   | 0.02241 (1.976)                    |
| CC1/2   | 1 (0.106)                          |
| CC*   | 1 (0.437)                          |
| Reflections used in refinement                    | 57275 (4008)                       |
| Reflections used for R-free                       | 2924 (215)                         |
| R-work  | 0.1800 (0.4105)                    |

|                              |                 |
|------------------------------|-----------------|
| R-free                       | 0.2115 (0.4402) |
| CC(work)                     | 0.969 (0.410)   |
| CC(free)                     | 0.955 (0.270)   |
| Number of non-hydrogen atoms | 4931            |
| macromolecules               | 4367            |
| ligands                      | 111             |
| solvent                      | 453             |
| Protein residues             | 553             |
| RMS(bonds) (Å)               | 0.011           |
| RMS(angles) (°)              | 1.1             |
| Ramachandran favored (%)     | 98              |
| Ramachandran allowed (%)     | 2               |
| Ramachandran outliers (%)    | 0               |
| Rotamer outliers (%)         | 0.64            |
| Clashscore                   | 5.2             |
| Average B-factor             | 47.8            |
| macromolecules               | 47.42           |
| ligands                      | 55.19           |
| solvent                      | 49.59           |

\* Parentheses indicate values for last resolution shell

**Supplementary Table 5.** LC-MS/MS analysis of sfGFP samples generated in DH10B and DH10B  $\Delta aspC \Delta tyrB$ . Sample numbers refer to column values shown in Fig. 6e. All values are in %.

| Sample                | <i>i</i> | <i>ii</i> | <i>iii</i> | <i>iv</i> | <i>v</i> | <i>vi</i> | <i>vii</i> | <i>viii</i> | <i>ix</i> |
|-----------------------|----------|-----------|------------|-----------|----------|-----------|------------|-------------|-----------|
| variant/<br>substrate | WT       | 1         | 2          | 20        | 21       | 20        | 21         | 20          | 21        |
| <b>Y</b>              | 100.00   | 0.36      | 0.05       | 0.28      | 0.59     | 0.15      | 0.95       | 0.60        | 1.57      |
| <b>1</b>              | 0.00     | 94.56     | 0.07       | 0.00      | 0.00     | 0.00      | 0.00       | 0.00        | 0.00      |
| <b>2</b>              | 0.00     | 4.44      | 99.82      | 0.00      | 0.00     | 0.00      | 0.00       | 0.00        | 0.00      |
| <b>20</b>             | 0.00     | 0.00      | 0.00       | 98.73     | 85.35    | 97.14     | 51.54      | 98.19       | 52.31     |
| <b>21</b>             | 0.00     | 0.00      | 0.00       | 0.32      | 11.23    | 0.63      | 38.48      | 0.65        | 44.10     |
| <b>A</b>              | 0.00     | 0.02      | 0.00       | 0.03      | 0.17     | 0.14      | 0.76       | 0.02        | 0.00      |
| <b>F</b>              | 0.00     | 0.14      | 0.01       | 0.34      | 0.61     | 0.15      | 0.20       | 0.34        | 0.27      |
| <b>K</b>              | 0.00     | 0.32      | 0.01       | 0.02      | 0.13     | 0.14      | 0.90       | 0.01        | 0.53      |
| <b>Q</b>              | 0.00     | 0.08      | 0.03       | 0.11      | 0.81     | 0.62      | 3.49       | 0.07        | 0.07      |
| <b>W</b>              | 0.00     | 0.08      | 0.00       | 0.17      | 1.11     | 1.02      | 3.68       | 0.12        | 0.48      |

References:

1. Herold, S., Bafaluy, D. & Muñiz, K. Anodic benzylic C(sp<sub>3</sub>)–H amination: unified access to pyrrolidines and piperidines. *Green Chem.* **20**, 3191–3196 (2018).
2. Matveeva, E. D. *et al.* Syntheses of Compounds Active toward Glutamate Receptors: II. Synthesis of Spiro Hydantoins of the Indan Series. *Russ. J. Org. Chem.* **38**, 1769–1774 (2002).
3. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).